

Constraining Operations

A Phase-based Account of Improper Movement
and Anaphoric Binding

Monika Bader

A thesis submitted for the degree of Philosophiae Doctor

University of Tromsø

Faculty of Humanities, Social Sciences and Education
Center for Advanced Study in Theoretical Linguistics

May 2011

Contents

Acknowledgements	vii
Abbreviations	ix
1 Introduction	1
1.1 The empirical domain	1
1.2 Theoretical background	4
1.3 Organization of the dissertation	7
I Improper Movement	9
2 When does a phrase move (im)properly?	11
2.1 Expanding the movement type inventory	13
2.1.1 On the status of A- vs A'-distinction	13
2.1.2 Exploring feeding/bleeding relations between movement operations	16
2.2 Non-identity cases	24
2.3 Clause-boundedness	34
2.4 Summary and conclusions	36
3 Previous analyses of improper movement	37
3.1 The binding-theoretic approach	37
3.2 Müller and Sternefeld (1993)	38
3.3 Williams (2003)	41
3.4 Obata and Epstein (2008)	47
3.5 Abels (2008b)	50
3.6 Summary	53
4 Relating improper movement to the functional sequence	55
4.1 The GBOIM is too weak	57
4.2 Strengthening the GBOIM	63
4.2.1 Enter locality	63
4.2.2 Non-identity cases again	65

4.3	ECM constructions	68
4.3.1	On the status of EPP in non-finite TP	70
4.3.2	Interim summary	87
4.4	Long movement out of a CP	87
4.5	Conclusions	90
5	Deriving improper movement from the functional sequence	93
5.1	Svenonius (2004)	94
5.2	The proposal	100
5.2.1	Outline of the proposal	100
5.2.2	Long movement out of a CP	106
5.2.3	Non-identity cases	108
5.2.4	Middle field scrambling	112
5.2.5	Symmetric feeding relations	115
5.3	Summary	129
II	Anaphora	131
6	Setting the stage	133
7	Phases as local binding domains	137
7.1	Introduction	137
7.2	Laying out the hypothesis	138
7.3	Binding in finite and non-finite clauses	142
7.4	Non-local binding	146
7.5	Binding into picture-NPs	150
7.6	Zooming in on non-local binding	161
7.6.1	Refining the typology of binding relations	162
7.6.2	Variable binding vs coreference	163
7.6.3	A fresh look at nonlocal binding contexts	168
7.6.4	Further considerations	175
7.6.5	Interim summary	177
7.7	The ban on A'-binders	177
7.7.1	Local binding relations	178
7.7.2	Non-local binding relations	184
7.7.3	Long-distance scrambling	189
7.8	Binding into PPs	190
7.8.1	Functional Ps	191
7.8.2	Spatial Ps	193
7.8.3	Interim summary	199
7.9	Hicks (2009)	200
7.10	Summary	205

8	Encoding binding relations	207
8.1	Previous proposals	208
8.1.1	Binding relations derived by Merge/Move	208
8.1.2	Binding relations derived by Agree	211
8.2	Developing an analysis	216
8.2.1	On lexical insertion	216
8.2.2	On the nature of the feature(s) involved	218
8.2.3	Some comments on the feature system assumed	226
8.2.4	Minimality violations	229
8.2.5	Encoding a three-way contrast in the anaphoric system: A case study of Dutch	231
8.2.6	Summary	248
8.3	The (non)complementarity issue	250
8.3.1	Non-local binding configurations	250
8.3.2	Reflexive local binding configurations	252
8.3.3	Non-reflexive local binding configurations	253
8.4	Summary and some consequences	266
9	Conclusion	271

Acknowledgements

The process of dissertation writing, in its final stages, is a process of gradual estrangement from the rest of the world. First you stop participating in seminars and other academic activities. You stop writing e-mails to your friends back home. Then even your friends who are not far away get less and less of your attention. You say no to skiing tours, hytte weekends, sauna evenings, cinema going, and all the other things you used to enjoy doing. At some point, even a short lunch with your fellow colleagues seems like too much of a distraction. When you reach that point where you don't have time even for your closest family members, you know that the end is near. Writing acknowledgements feels a bit like reconnecting with the world. As you sit down and think of all the people who have contributed in some way or another and helped you accomplish this important feat, you are reminded that, despite what you might have felt in the last couple of months, dissertation writing is far from being a solitary process.

First and foremost, the person to thank is my supervisor, Peter Svenonius. Who knows what I would have ended up doing and what my life would have turned out to be, if I hadn't met Peter at some point (more precisely in July 2001 at a summer school in Niš) and heard about this wonderful place called Tromsø and the opportunities it has to offer for aspiring linguists. Of course, Peter has played a central role in the development of this dissertation. His insightful comments, constructive criticism, patient guidance, and encouragement have been truly invaluable.

I am also thankful to the broader CASTL linguistic community for making Tromsø such an inspiring place to work at. I remember when I was about to travel to Tromsø for the first time, I had to stand on a chair to show the woman at the travel agency where Tromsø is on this enormous map of Europe. "Why would you want to go there?!"- she asked quite shocked. The answer- because it has a great linguistic department- didn't particularly impress her. But for anyone wishing to study linguistics, the collection of people in Tromsø at that time would surely have been enough of a reason to venture so far north. In the years I have since then spent here, the department kept adding excellent people to its ranks (both professors and students) and has grown into the CASTL it is today. I was fortunate to have a series of great teachers (Gillian, Tarald, Michal, Peter, Klaus, to name a few) that were instrumental in shaping the way I think about the linguistic issues. In this context, I have to especially mention Michal Starke, who was my supervisor for a part of my study period,

and from whom I have learned so much both in classes and in our meetings. Special thanks to Klaus whose seminars were (obviously) a great inspiration to me and who provided the necessary encouragement when it was needed the most. I thank also the administrative staff at CASTL for all the assistance and the Faculty of Humanities for providing excellent working conditions.

I am incredibly grateful to my present and past fellow students, a truly exceptional group of people. They have made studying in Tromsø both intellectually challenging and extremely pleasurable. Special thanks to Kaori, Marleen and Yulia for our numerous joint lunches and to Pavel I. for much needed help with Latex. I am also grateful to everyone who provided me with grammaticality judgements.

I am fortunate to have so many wonderful friends who have made Tromsø into so much more than just a place of study. For all the non-linguistic activities, I am grateful to Kristine, Sylvia, Madeleine, Bjørn, Peter M., Naoyuki, Dragana, Minjeong, PJ, Klaus and Luisa, Islam and Cvijeta, Laila, Andrea and especially Marleen, Kaori and Tom, Marina and Pavel C., Isabelle and Pierre. Thank you all! Special thanks to Dragana for providing the crash space when it was needed. Finally, I have to single out a couple of people who have practically been like a family to me. Zhenja, with whom I have shared a flat in the first year as a PhD and so much more in all the years that we have known each other. Christine and Lars. Finally and most of all, Yulia and Adnan.

I also have to thank my family: my parents and my brother. Though far away, they have been a constant source of support and love. I wouldn't be what I am and where I am today, if it wasn't for them. Hvala mama i tata!

Finally, how to thank people that mean everything to you: Mateja, for putting a smile on my face every day, and Luci, for everything that he is. Luci has been so much a part of this process, and I don't mean just as a source of emotional support. I honestly don't know how many times I have pestered him for judgements. He got so immersed into the whole thing that he started noticing himself when something of linguistic interest would come up. So one evening we were making dinner and he just happened to produce the following sentence: "Tortellini weiss ich genau wann ich das erste Mal gegessen hab'." Then he stopped, turned to me and said: "That is something for you!" Indeed. Anticipating the discussion to come: topicalization across a wh-island.

Abbreviations used in glosses

ABL	ablative
ACC	accusative
aux	auxiliary
COMP	complementizer
DAT	dative
GEN	genitive
INF	infinitive
INSTR	instrumental
m	masculine
NOM	nominative
PRT	particle
Q	question particle
REFL	reflexive
sg	singular
SUBJ	subjunctive
TOP	topic marker

Chapter 1

Introduction

1.1 The empirical domain

This dissertation offers a novel angle on uniformity of movement and binding dependencies, arguing that common syntactic factors constrain both movement operations and the relation between anaphoric expressions and their antecedents.

The starting point is a long noted puzzle regarding the possible sequencing of movement operations. Consider the following example of what Chomsky (1973) calls improper passivization:

- (1) a. John asked what to read.
- b. *What was asked to read by John?

As (1a) illustrates, a phrase can undergo a step of *wh*-movement to the Spec of the embedded CP, but the same phrase cannot then be passivized, as evidenced by the ungrammaticality of (1b). Some further examples considered to involve instances of improper movement are provided in (2). Assuming that locality forces a stop-over in the Spec of embedded CP, these examples again show that a phrase cannot be moved to the subject position from the Spec of CP.

- (2) a. *Who is possible John will see? (Chomsky (1981))
- b. *John is possible Bill will see. (Chomsky (1981))
- c. *John seems that it is certain to fix the car. (Chomsky (1995))

Examples of this type led to the conclusion that operations such as raising or passivization, i.e. operations targeting an A(argument)-position cannot follow movement operations that target a non-argument, i.e. an A'-position. The relevant generalization is stated in (3).

- (3) *Ban on Improper Movement*
A phrase occupying an A'-position cannot be moved to an A-position.

The Ban on Improper Movement rules out examples where an A'-moved phrase undergoes a step of A-movement, but allows the reverse application of these two movement operations. Ever since these restrictions on ordering of movement operations were first noted, the question of how to derive them in a principled way has been one of the major concerns of the linguistic theory. This thesis aims at providing a novel analysis of the phenomenon, relying on theoretical tools and conceptual advances of the now broadly adopted Minimalist framework (Chomsky (1995, 2001)).

In addition to examples like (1) and (2), I will also consider the cases where A- and A'-movement interact in more complex ways and argue, following some recent proposals, that the empirical scope of the original generalization should be extended beyond standard examples featuring consecutive movements of the same phrase. Two additional configurations will be considered and shown to display the same ordering asymmetry between A and A'-operations. The first set of relevant cases are those where only a part of a phrase in an A'-position undergoes A-movement. As (4a) illustrates, such a derivation yields an illicit output. In (4a), first the infinitival clause, *how likely Oscar to win*, undergoes a step of *wh*-movement to the Spec of the embedded C. A part of the moved phrase, namely *Oscar*, is then passivized and the resulting structure is ungrammatical (cf. (1b)). A licit output can be obtained if A-movement applies before the *wh*-movement of the infinitival clause, as shown in (4b) (examples are from Collins (2005)).

- (4) a. * $[_{IP} \text{Oscar}_i \text{ was asked } [_{CP} [\text{how likely } t_i \text{ to win}]_j \text{ it was } t_j]]$.
 b. Max asked $[_{CP} [\text{how likely } t_i \text{ to win}]_j [_{IP} \text{Oscar}_i \text{ was } t_j]]$.

The second set of relevant cases are those involving remnant movement, i.e. configurations where a phrase containing a trace of movement is displaced. Consider (5) (from Abels (2007)). In this example, the phrase, *which king*, extracts out of the containing noun phrase and *wh*-moves to the embedded CP. Subsequently, the remnant phrase, *a picture of*, is passivized. The reverse application of these two movement operations is licit, as already illustrated in (4b).

- (5) * $[_{IP} [\text{A picture of } t_i]_j \text{ is known } [_{CP} [\text{which king}]_i \text{ to have been sold } t_j]]$.

The examples just discussed show that in all three configurations considered, *wh*-movement cannot feed passivization. On the basis of patterns such as these, we will conclude that the same restrictions on the sequencing of A- and A'-operations can be observed not only in cases where both types of operation affect the same phrase, but also in cases involving subextraction out of moved phrases and remnant movement.

I will then argue that the observed ordering restrictions on (movement) operations in all three types of configurations can be derived by relying on the internal featural make-up of the moving phrase, in conjunction with a particular view regarding the timing and manner in which linguistic structure is spelled out. Following recent versions of Minimalism (e.g. Chomsky (2000) and subsequent work), I assume that syntactic derivation proceeds in incremental chunks called *phases*. On the account

pursued here, phase-based locality plays an important part in determining the order in which operations can apply.

I will then turn to a different empirical domain, that of anaphoric binding, and argue that the same factors that I have claimed play a role in constraining movement operations are also crucial in regulating the distribution and interpretation of anaphoric relations, namely the featural composition of the expressions involved and the nature of Spell-Out. I will argue that due to their feature specification, anaphors must enter into a syntactic Agree relation with another c-commanding noun phrase. This syntactic dependency between an anaphor and its antecedent translates into a particular interpretative dependency at the semantic interface. On the other hand, no syntactic dependency is assumed to hold between a bound pronoun and its antecedent. Pursuing this view will allow us to capture the well-known observations that, unlike pronouns, anaphors require the presence of a syntactic antecedent, and that moreover, the antecedent must be within the same local domain as the anaphor. Both of these requirements are met in cases like (6a), for instance. In (6b), though pragmatically salient, *Mary* is not local enough to bind the anaphor. The relation between an antecedent and a pronoun on the other hand is not subject to the same locality restrictions, nor is it the case that a pronoun must find a syntactic antecedent. Thus, in (6c), the pronoun *her* can refer to *Mary*, but it can also pick up another referent from the discourse.

- (6) a. Mary criticized herself.
 b. *As soon as Mary arrived, they asked herself to leave.
 c. As soon as Mary arrived, they asked her to leave.

Given that anaphoric binding is encoded in terms of an Agree-relation, and the application of the Agree-operation itself is constrained by phases, the analysis pursued here predicts that phases constitute the local domain for anaphoric binding. A considerable part of the dissertation will be devoted to showing that treating phases as the relevant locality domain for binding can take us a long way in accounting for empirical facts.

I will also follow the work of Lebeaux (1984), Reinhart and Reuland (1993), Pollard and Sag (1992), among others in making a distinction between local and non-local binding relations and argue that relying on the phase as the relevant locality domain makes the correct split between these two binding environments. I will then further enrich the typology of binding relations, by distinguishing two types of non-local relations: those established at the semantic interface and those established in the discourse component. Though the two nonlocal binding relations pattern alike in being insensitive to syntactic locality restrictions, it will be argued that semantically they differ considerably.

On the view pursued here, phases thus play a crucial role in constraining both movement and binding dependencies. The intuition that locality constraints on movement and binding should be unified is an old one, going back to Chomsky (1973). In Chomsky (1981), the relation between movement and binding phenomena was for-

malized by assuming that traces of movement are also subject to the binding theory. Improper movement effects were then argued to involve violations of binding conditions. In other words, it was the binding theoretic properties of traces that ensured the proper sequencing of movement operations. Chomsky (1993), Chomsky and Lasnik (1993) abandon this line of thinking by arguing that binding relations are purely interface phenomena, leaving thereby the well-known similarities between the locality constraints on movement and binding unaccounted for. The current proposal offers a way of reestablishing the link between movement and binding. The distribution of anaphoric expressions and the proper sequencing of movement operations are here once again attributed to the same factors, though the implementation of this intuition will differ substantially from the one pursued in Chomsky (1981). The primary data in this work will be drawn mostly from English, but other languages, including German, Dutch and Serbian, will be brought in at times where they provide complementary evidence.

1.2 Theoretical background

The analysis of movement and binding phenomena developed in the following chapters is couched in the Minimalist framework and will rely heavily on certain Minimalist assumptions regarding the nature of syntactic derivation. This section therefore provides a brief outline of some key assumptions in the Minimalist framework, focusing particularly on those aspects of the theory that will be relevant in the discussion to follow.

Minimalism explores the hypothesis that the human language faculty (FL) is an optimal solution to minimal design specifications. According to Chomsky (2001), these design specifications can be viewed as ‘legibility conditions’: “for each language L (a state of FL), expressions generated by L must be ‘legible’ to systems that access these objects at the interface between FL and external systems” (Chomsky (2001), p.1), the external systems being the S(ensory)-M(otor) and the C(onceptual)-I(ntentional) system. The role of the computational component of the grammar is to supply the interfaces with legible representations. Syntax does so by manipulating features of lexical items. Features can be viewed as attribute-value pairs. Feature-attributes either enter the derivation already bearing a particular value, or they do not, in which case the feature is said to be unvalued. Unvalued features are not tolerated at the interface, so these must receive a value in the course of the syntactic derivation.¹ Feature-valuation is achieved via the operation Agree. An unvalued feature, or ‘probe’ is the trigger for Agree. It searches within a syntactic domain for the same feature-attribute bearing a value, a ‘goal’. If particular structural conditions are met, i.e. if the probe and the goal are in a c-command relation, then Agree can apply

¹Chomsky (2001) assumes that all and only unvalued features are uninterpretable and therefore must be valued and eliminated before reaching the interface. This assumption will be scrutinized in chapter 8.

copying the value of the goal onto the probe. (7) illustrates this on the example of ϕ -feature agreement. An Agree relation is established between the unvalued ϕ -features of T, acting as a probe, and a matching valued set of ϕ -features borne by the subject, *John*. As a result, T receives the ϕ -features of the subject, yielding the surface effect of noun-verb agreement.

- (7) John misses Mary.
- a. [T' T_[\phi:_] [vP John_[\phi:3,sg] [v-miss [vP [V Mary]]]]] (Agree applies)
 - b. [T' T_[\phi:3,sg] [vP John_[\phi:3,sg] [v-miss [vP [V Mary]]]]]

Not every matching pair induces Agree. In addition to being in the c-command domain of the probe, the potential goal bearing a matching feature must also be the *closest* one available (Chomsky (2000), p. 122). G is the ‘closest’ goal if a probe P c-commands G and there is no G' such that P asymmetrically c-commands G' and G' asymmetrically c-commands G.² Therefore, although the object, *Mary*, bears a valued set of ϕ -features, *Mary* in (7) cannot enter into an Agree relation with T since it is not the closest available goal.

The computational system of human language (C_{HL}) manipulates features (in the manner just described) and at some stage, the generated syntactic object is sent to the interfaces. This point is referred to as Spell-Out. Chomsky (2000, 2001) argues that Spell-Out applies at multiple points in the course of the derivation (cf. Uriagereka (1999)). In other words, the derivation is argued to proceed in incremental chunks, referred to as *phases*. Once a particular chunk of syntactic structure is transferred to the interfaces, it is no longer accessible for further syntactic operations. According to Chomsky, this leads to reduction in computational burden, given that syntactic computation can proceed without checking back to these earlier stages in the derivation. Chomsky argues that phase heads are C and the transitive v (more specifically, v that introduces an external argument, which he marks as v^*).³ Transfer of the entire phasal category, CP or vP, would preclude the derivation to continue beyond the immediate phase. Hence, some notion of the ‘escape hatch’ is generally assumed. Chomsky (2001) argues that the phasal head itself and any of its specifiers remain accessible for operations at the next higher phase level, but nothing in the complement of the phasal head does.

- (8) [α [H β]]

²In Chomsky (1995), this locality requirement was imposed on the operation Attract and referred to as the *Minimal Link Condition (MLC)*. Since in Chomsky (2000) and subsequent work, Agree is the operation that takes over the empirical burden of Attract, the obvious step to take was to incorporate the MLC into the definition of Agree, as Chomsky (2000) does.

³It has been argued that the inventory of phase heads is richer than Chomsky assumes, including possibly intransitive and unaccusative v (Legate (2003)), D (e.g. Svenonius (2004), Bošković (2005)), and even P (Abels (2003)).

(9) *Phase Impenetrability Condition* (Chomsky (2000))

In a phase α with head H, the domain of H [=complement of H] is not accessible to operations outside α [=HP], but only H and its edge [H plus any/all of its specifiers].

The chief consequence of the Phase Impenetrability Condition (PIC) is that operations such as Move cannot access elements embedded inside a phase; only the material in the phase edge is computationally available in the next phase up.⁴ This forces movement to apply successive-cyclically, via phase edges. To make the discussion more concrete, let us look at an example involving long distance *wh*-movement.

(10) What did John say that Mary bought?

The embedded verb in (10) is transitive, which means that the embedded v^*P constitutes a phase. Since the domain of the phase is impenetrable to further syntactic operations, the *wh*-phrase will have to move to the edge of v^*P . This is usually ensured by endowing the phase heads (in this case, the v^*) with an [EPP] (or a P(eripheral)) feature. Once all operations within the phase are carried out, the domain of the phase (i.e. its VP complement) will undergo Spell-Out, making both VP and any constituents of VP invisible to further syntactic computation. The *wh*-phrase will continue to move in this manner through all intermediate phase edges (in this case the intermediate CP and the matrix v^*P) until it reaches the matrix CP, its final landing site. The derivation of (10) thus looks roughly as in (11).

(11) [_{CP} What_{*i*} did John [_{v^*P} *t_i* say [_{CP} *t_i* that Mary [_{v^*P} *t_i* bought *t_i*]]]]?

In Chomsky (2001), the PIC is reformulated so as to allow the domain of the phase to be accessible until the head of the next higher phase is merged. The analysis to be outlined adopts the former more restrictive approach, whereby Spell-Out takes place as soon as the phasal head is introduced. In fact, in chapter 5, I will further limit the accessibility of elements contained within a phase, by somewhat modifying Chomsky's (2000) view regarding what precisely spells out. On the view pursued here, upon the merge of a phase head, not only will its complement become opaque, but under certain conditions, so will the phrase, or part of the phrase, occupying the specifier of the phasal head. This will have important ramifications both for the analysis of movement and for the analysis of binding phenomena.

It should also be pointed out in the context of this discussion that the literature is not always careful in distinguishing the notion of phases from the constituent that

⁴According to Chomsky (2000, 2001), pure Agree is also constrained by phases, i.e. by the PIC. In these works, Chomsky argues that Move is a composite operation, with Agree and Merge as its subcomponents. Agree is a prerequisite for movement as it establishes the required feature dependency, while movement itself is simply the consequence of an EPP-type trigger on the probe. On this view, locality constraints on movement must therefore be reinterpreted as constraints on Agree.

spells out.⁵ Thus, when a phase is described as being sent to Spell-Out, what is often meant is that actually the *domain* of that phase goes to Spell-Out, with the edge remaining accessible. I will use the term *phase* here to refer to the constituent containing the phase head, together with its edge and domain, and assume that whatever is spelled out is opaque for further operations. Thus, on this view, strictly speaking it is not the phase that is sent to Spell-Out, but rather a subpart of it. The parts of the phase that are spelled out cannot then be further manipulated by syntactic operations.

In addition to somewhat modifying Chomsky's assumptions regarding which constituent(s) spells out, another departure from Chomsky's view of phases that will be significant in the context of the present study concerns the identity of the phasal heads. I will argue that the relevant phasal heads are C and Asp, AspP being the projection immediately dominating *v/V*, rather than C and transitive *v*, as Chomsky assumes. The inventory of phasal heads will bear importantly on how the system works. For instance, note that the external argument occupies the phase edge when it is inserted in the structure only if *v* is the phasal head, but not if Asp is. This will have important consequences, particularly for the analysis of binding facts. The approach pursued here also assumes a richer inventory of categories, in the spirit of the cartographic approach to clausal structure (Rizzi (2002), Cinque (1999), among others).

Let me close off this brief introduction with some concerns of a methodological nature. Although achieving empirical coverage is obviously important, Minimalism also places great weight on theoretical standards of parsimony, simplicity, and elegance. Language phenomena should be accounted for by using as little technical machinery as possible and only that which is 'virtually conceptually necessary'. In the chapters to follow, we will try to comply to this methodology as much as possible, in general striving to reduce the number of theoretical tools and assumptions and relying particularly on those that are not relevant solely for the empirical data being discussed, but rather have broader application in the theory.

1.3 Organization of the dissertation

The thesis is organized as follows. The first part (chapters 2, 3, 4, 5) focuses on improper movement phenomena, while the second part (chapters 7 and 8) takes anaphoric binding as its empirical domain.

Chapter 2 lays the empirical groundwork for the investigation of improper movement. Interactions between various movement operations in terms of possible feeding/bleeding relations are examined. I also extend the empirical scope of the original generalization from examples involving consecutive movements of the same phrase, to cases involving remnant movement and subextraction from moved phrases, following some recent proposals to this effect.

⁵See also Svenonius (2004) for relevant discussion.

Chapter 3 reviews some previous analyses of improper movement, discussing their merits and shortcomings. I start off with the classical analysis relying on binding-theoretic properties of traces, but devote most of my attention to more recent accounts of the phenomenon.

Chapters 4 and 5 then outline and discuss two possible ways of deriving the observed restrictions on the ordering of movement operations from the hierarchy of functional projections. Chapter 4 posits a constraint on movement, regulating the choice of possible landing sites. It is shown that in order for this constraint to have any significant empirical value, it must be considered in tandem with a theory of locality, conceived of here in terms of phases. Some cases of what I will argue to be instances of improper movement nevertheless remain beyond the scope of the proposed account, unless additional assumptions are adopted. What is more, an internal theoretical inconsistency will be revealed, prompting us to search for a different solution.

An alternative is developed in chapter 5. The proposal derives the noted ordering restrictions by relying on the internal featural make-up of the moving phrase, again in conjunction with a phase-based locality. I will argue that there is no need to formulate an independent constraint on movement; rather the correct ordering can be derived from the way the derivation proceeds and particular assumptions regarding the nature of the functional hierarchy. As part of the analysis, a modification to the standard view of phases and the nature of Spell-Out will be suggested, with important ramifications not only for improper movement phenomena, but also for binding facts to be discussed in the following chapters.

In the rest of the dissertation, I then turn to the discussion of anaphoric relations, arguing that the featural composition of elements involved and the nature of Spell-Out, as conceived of in chapter 5, play a crucial role in constraining dependencies also in this empirical domain. Chapter 7 investigates in great detail the consequences of assuming that the phase is the relevant locality domain for binding relations. It will be shown that such an assumption has not only theoretical appeal, but empirical bite as well. The proposal will lead us to suggest a three-way distinction in binding relations, the nature of which will be investigated in this and the following chapter.

Chapter 8 explains the relevance of phases for anaphoric binding. Following the intuition behind some recent proposals, I will argue that anaphors enter into an Agree-relation with their antecedents in the course of the syntactic derivation. This featural dependency then translates into a particular interpretative dependency at the interface. A proposal along these lines explains why syntactic factors play a role in governing the distribution of anaphors. Though the main focus will be on binding relations involving anaphors, the proposed analysis, as will be shown, will also have significant ramifications for the distribution of pronouns.

Part I
Improper Movement

Example (2b) can thus be taken to show that A'-movement cannot feed A-movement. On the other hand, movement from an A to an A'-position, or from an A' to another A'-position is allowed, as shown in (2a) and (2c), respectively.

The Ban on Improper Movement also captures the ungrammaticality of examples in (3) under the standard assumption that movement out of a CP proceeds successive cyclically through each intermediate SpecCP. The illicit movement step is again the one which takes the moving phrase from the embedded CP to the matrix TP.

- (3) a. *Who is possible John will see? (Chomsky (1981))
 [_{CP} who [_{IP} <who> is possible [_{CP} <who> [_{IP} John will see <who>]]]]
 b. *Who seems it is likely to leave? (Obata and Epstein (2008))
 [_{CP} who [_{IP} <who> seems [_{CP} <who> [_{IP} it is likely [_{IP} to <who> leave]]]]]]

Examples in (4), cases of so called *superraising*, do not contain a *wh*-phrase and thus omit the final step in the derivation of (3). The derivation of these examples otherwise parallels those in (3) and involves the same illicit movement step.

- (4) a. *John is possible Bill will see. (Chomsky (1981))
 [_{IP} **John** is possible [_{CP} <John> [_{IP} Bill will see <John>]]]
 b. *John seems (that) is likely to win. (Zwart (1996))
 [_{IP} **John** seems [_{CP} <John> (that) [_{IP} <John> is likely to <John> win]]]
 c. *John seems (that) it is likely to win. (Zwart (1996))
 [_{IP} **John** seems [_{CP} <John> (that) [_{IP} it is likely to <John> win]]]

Movement from an A- to an A'-position, on the other hand, does not yield a deviant result. In (2a) for instance, the object moves to SpecCP from its base-generated position, which is an A-position. Alternatively, it might be argued that the object first moves to a case position and then to a position in the C-domain. On either scenario, the launching site of *wh*-movement is an A-position. That a phrase can A'-move from an A-position can also be shown on the basis of examples like (5) and (6), given standard assumptions that subjects extract out of SpecTP and move to the C-domain.

- (5) Who seems to like cheese?
 [_{CP} who [_{IP} <who> seems [_{TP} <who>to like cheese]]]
 A' ← A
 (6) Who seems to be likely to win?
 [_{CP} who [_{IP} <who> seems [_{IP} <who> to be likely [_{IP} <who> to win]]]]
 A' ← A

The Ban on Improper Movement (BOIM) is just a descriptive generalization, which itself calls for an explanation. The question of what the nature of the mechanism responsible for this prohibition is has been a central concern since Chomsky (1973),

though a satisfactory account of the phenomenon remains elusive to this date. Before we begin tackling this issue, we must first reconsider the very generalization we are trying to account for. First of all, the BOIM, as standardly formulated, relies on the binary distinction between A- and A'-movements. However, it has often been pointed out that this binary division is insufficient to adequately describe movement phenomena, given that the diagnostic tools that have been utilized to distinguish A- from A'-movements do not always yield clear-cut results.² If the A/A'-distinction should be refined, then the immediate question we face is how to properly characterize the BOIM. Secondly, the traditional formulation of the BOIM is limited to cases involving consecutive movements of the same phrase. However, it has been argued recently that the same ordering restrictions on the application of movement operations can be observed in cases involving remnant movement and subextraction from a moved phrase (Grewendorf (2003), Williams (2003), Abels (2008b)).

In this chapter, I will explore the empirical basis of the BOIM, thereby setting the stage for subsequent discussion. I will focus in particular on the two issues just raised. Namely, (i) what restrictions on the ordering of movement operations can be observed once a more fine-grained typology of movement operations is assumed, and (ii) whether the observed ordering restrictions extend beyond cases involving consecutive movements of the same phrase.

2.1 Expanding the movement type inventory

2.1.1 On the status of A- vs A'-distinction

The standard formulation of the BOIM rests on the assumption that movement operations can be classified as either A or A'-movements. Though widely assumed, over the years this binary typology of movement operations has proven to be problematic in a number of respects. For one, the notions *A-position* and *A'-position* are not well-defined. When first formulated, A-positions (or Argument-positions) were considered to be those positions in which an element can be base-generated and assigned a θ -role. It was assumed that SpecIP is a position where subjects (of transitive and unergative intransitive verbs) are base-generated in active sentences and receive a θ -role (cf. Chomsky (1981)). On such grounds it was argued that the SpecIP qualifies as an A-position. Note that it was only required that an A-position potentially functions as a θ -position, not that it always does so. Thus, movement to SpecIP, as with passive and raising verbs (see (7)), is commonly treated as an instance of A-movement, i.e. movement to an A-position, though in these cases no θ -role is assigned in SpecIP. An A'-position, such as the SpecCP, on the other hand, can never act as a θ -position.

- (7) a. [IP John [was criticized <John>]]

²See particularly Williams (2003), who argues extensively that the A/A'-typology is too coarse.

- b. [_{IP} John seems [_{IP} <John> to be intelligent]]

That this is not a particularly satisfying way of defining A/A'-positions has been pointed out by Chomsky (1995), among others. Referring to the notions *θ-position* and *A/A'-position*, Chomsky (1995, pp. 63-64) notes the following: "These notions are formally quite different in character. A particular occurrence of a category in a phrase marker is, or not, a theta-position, depending on whether it is theta-marked in that phrase marker. The notion *A-position*, however, depends on 'potential theta-marking', which is to say that it presupposes an equivalent position of some member of the equivalence class. This is not an entirely straightforward notion, and . . . it becomes unspecifiable in any way that will bear the considerable theoretical burden that has been laid on the A- vs. A'-distinction, which enters crucially into large areas of current work." What is more, the VP-internal subject hypothesis calls into question even this potential of IP to function as a *θ-position*. It is now generally assumed that subjects of transitive and unergative verbs are base-generated in Spec of vP and receive a *θ*-role in this position, rather than in SpecIP. If so, then Spec of IP in fact never functions as a *θ-position*. Thus, if *θ*-assignment is taken to be the distinguishing criterion, then SpecIP turns out to be an A'-position and movement to SpecIP in raising and passive construction an instance of A'-movement. This is clearly an unsatisfactory outcome. The goal of grouping together various movement operations under the label of A- or A'-movements was to capture the observation that particular movement operations pattern alike with respect to various properties. For instance, it is argued that elements moved to A-positions can bind anaphoric expressions that are unbound in their base position. On the other hand, A'-movement does not give rise to new binding possibilities. In this respect, raising and passive (movements targeting the SpecIP position) clearly contrast with A'-movement operations, such as *wh*-movement or topicalization (examples are from Grewendorf and Sabel (1999)).

- (8) a. The men_{*i*} seem to each other_{*i*} [<the men> to be nice].
 b. John thinks that the men_{*i*} were [kissed <the men>] by each other_{*i*}'s wives.
- (9) a. *Which actors_{*i*} did pictures of themselves_{*i*} convince the director that he should interview <which actors>?
 b. *The guests_{*i*}, each other_{*i*}'s dance partners criticized <the guests>.

The question then arises as to how A/A'-positions could be characterized. Given that A-movements are said to share a number of properties, we might rely on these properties to identify the set of A-positions. The phenomena that have been argued to distinguish A from A'-movements include the binding possibilities, weak crossover effects, licensing of parasitic gaps, locality restrictions, etc. However, applying these diagnostics to movement operations does not always yield clear-cut results. There are movement operations that seem to display mixed properties, i.e. that seem to behave both as A- and A'-movement, depending on which property is tested. Consider for in-

stance scrambling to the pre-subject position in German. The scrambled direct object cannot bind an anaphor contained in the subject, as illustrated in (10b). (10a) shows that the anaphor contained in the direct object is licensed by the c-commanding subject. In this respect, scrambling behaves as A'-movement (examples from Grewendorf and Sabel (1999)).³

- (10) a. weil der Student_i [die Lehrer von sich_i] zweifellos in
since the student.NOM [the teacher of self].ACC undoubtedly in
guter Erinnerung behalten hat.
good memory kept has
'The student has undoubtedly kept the teachers of himself in good
memory.'
- b. *weil den Studenten_i [die Lehrer von sich_i] zweifellos
since the student.ACC the teachers.NOM of self undoubtedly
<den Studenten> in guter Erinnerung behalten haben.
in good memory kept have
intended: 'The teachers of himself have undoubtedly kept the student
in good memory.'

However, when weak crossover (WCO) effects are considered, scrambling to the pre-subject position in German patterns with A-movement. Weak crossover effects occur in configurations where an element A'-binds both a trace and a pronoun contained in an argument XP that c-commands the trace. *Wh*-movement in English is a paradigm example of a displacement type that exhibits WCO effects, as demonstrated in (11). However, scrambling in German does not yield a WCO violation, as (12b) shows (examples from Grewendorf and Sabel (1999)).⁴

- (11) *Who_i does his_i mother love <who_i>?
- (12) a. *weil seine_i Mutter jeden Studenten_i liebt.
since his mother.NOM every student.ACC loves

³Scrambling in German can also license parasitic gaps, which is standardly taken to be an A'-property. I refer the reader to Grewendorf and Sabel (1999) for relevant data. See also Neeleman (1994) who argues that this test is unreliable since parasitic gaps can also be licensed by A-movement in Dutch (Hinterhölzl (2006) shows the same for German), and Putnam (2007) for criticism of this claim.

⁴A complicating factor here is that short *wh*-movement in German also fails to display WCO effects, as illustrated below. However, Wiltschko (1998) has argued that this unexpected behaviour of *wh*-movement in German can be explained by assuming that the *wh*-movement step is preceded by a scrambling operation. This argument will be presented in more detail shortly.

- (i) Wen_i liebt seine_i Mutter?
who.ACC loves his mother.NOM

- b. weil [jeden Studenten_i] seine_i Mutter <jeden Studenten_i>
 since every student.ACC his motherNOM
 liebt.
 loves

Scrambling in German also patterns with canonical cases of A-movement, such as passive and raising, in being clause-bound. (13) shows that long-distance scrambling of a direct object out of a tensed clause results in ungrammaticality (from Hinterhölzl (2006)).

- (13) a. Gestern hat niemand geglaubt, dass er die Maria einladen wird.
 yesterday has nobody believed that he the Maria invite will
 b. *Gestern hat niemand die Maria geglaubt, dass er <die Maria>
 yesterday has nobody the Maria believed that he
 einladen wird.
 invite will

The patterns discussed show that scrambling (to the pre-subject position) in German cannot be classified as clearly displaying either A or A' characteristics.⁵ This is not just some quirk of German grammar. Saito (1992, 1994) for Japanese and Mahajan (1990) for Hindi have shown that clause-internal scrambling in these languages also does not display either A or A'-properties across-the-board. Cases such as these show that properties attributed to A and A'-movements do not always cluster together, casting doubt on the adequacy of this binary typology as a means of classifying and describing movement phenomena. These considerations call for a reformulation of the BOIM, one that would not appeal to the A/A'-distinction. To this aim, the following sections investigate what generalizations regarding the ordering of movement operations can be formulated if a richer inventory of movement types is assumed.

2.1.2 Exploring feeding/bleeding relations between movement operations

The BOIM formalizes the observation that there is an asymmetric ordering relation between two classes of (movement) operations. In particular, operations that target A-positions can be followed by A'-operations, but not the other way round. If we assume a richer inventory of movement types, many questions immediately arise. Some of these are formulated below (drawing on Abels (2008b)):

⁵The observation that German middle field scrambling exhibits mixed A/A'-properties is often referred to as *Webelhuth's Paradox*, as it was first noted by Webelhuth (1989). One way of resolving the paradox is to make a finer-grained distinctions among movement relations. See for instance Williams (2003) for a recent proposal along these lines.

1. is the ordering of operations total, i.e. is it the case that for any pair of operations $\{\alpha, \beta\}$, either α feeds β , or β feeds α ?
2. are movement operations always *asymmetrically* ordered, i.e. is it always the case that if an operation α feeds a different operation β , then β does not feed α ?
3. is the ordering regarding the application of various operations universal?

A number of authors have argued that, for a subset of movement operations they have investigated, the answer to the second question is positive: various movement operations indeed enter into asymmetric feeding/bleeding relations in a way that allows us to establish a hierarchy regarding the ordering of their application (Sternefeld (1993), Grewendorf (2003), Abels (2008b)).⁶ The proposed hierarchy, as stated in Abels (2008b), is given below.⁷

(14) $\theta \prec \text{A-mvt} \prec (\text{clause internal}) \text{ scrambling} \prec \text{wh-mvt} \prec \text{topicalization}$

What is meant by A-movement here is movement to a case position. As regards scrambling, Abels (2008b) limits the term to denote a local movement operation targeting a pre-subject position, of the type found in German.⁸ The hierarchy thus requires that θ -role assignment applies first, followed by Case related operations, followed by scrambling, etc. Let us now review some of the evidence in favour of the proposed Hierarchy of Operations.

A syntactic argument starts its life in a theta-position. In the Government and Binding (GB) framework, movement from theta to non-theta-positions and from non-theta to non-theta-positions was allowed. However, movement from one theta-position to another or from a non-theta to a theta-position was forbidden. These restrictions followed as a consequence of the assumption that theta-roles are assigned at the level of Deep Structure, ensuring thereby that no movement operation can precede theta-assignment. Although recent claims can be found in the literature that movement from a theta- to a theta-position should be allowed (Hornstein (2001), Ramchand (2008)), there seems to be a general consensus that movement from a non-theta to a theta-position is impossible. No movement operation (other than possibly theta-movement) is able to target a theta position. Consequently, theta-operations occupy, quite uncontroversially, the bottom of the Hierarchy.

⁶The researchers cited also assume that the proposed ordering is universal, though the validity of that assumption is yet to be confirmed.

⁷The hierarchies proposed by Sternefeld (1993) and Grewendorf (2003) are stated slightly differently, though for the data examined all three approaches in general make the same predictions. I will briefly return to these differences after examining the relevant data.

⁸Grewendorf (2003) places scrambling ('Adjunction' movement in his terms) on the same place in the Hierarchy of Operations as Abels does, but he assigns a broader scope to this notion. In particular, both mittelfeld scrambling in German and long-distance scrambling in Japanese would be instances of an Adjunction operation (though not mittelfeld scrambling in Japanese). See Grewendorf (2003) for further details.

The Hierarchy further states that A-movement can feed scrambling, *wh*-movement and topicalization, and that the reverse application of these operations is illicit. Consider first the interaction between A-movement (i.e. case-driven movement) and *wh*-movement. The licit ordering of these two operations is illustrated below.

- (15) a. [CP Who_i [IP t_i was arrested t_i]]?
 b. I know [CP who_i [IP t_i has been invited t_i]]
 c. [CP What_i did you [A_{gr}O t_i buy t_i]] ?
 d. [CP Who_i do you [A_{gr}O t_i believe [to have t_i won]]?

A few comments are in order regarding the examples above. In the first two examples, a phrase undergoes *wh*-movement from the subject position. It is sometimes argued that extraction from the subject position is prohibited (see for instance Rizzi (2006)), which would make the examples (15a) and (15b) irrelevant for the point we are trying to make. This complication however does not arise with respect to the last two examples. In (15c), it is the object that is *wh*-moved. Even if there is no movement from theta to a case position, i.e. if case is assigned in the base position, (15c) still might be taken to show that movement can proceed from a case to a *wh*-position. Anticipating somewhat what is to follow, this is all we need to show since we will later restate the BOIM in terms of possible launching and landing sites. Example (15d) involving *wh*-movement of an ECM subject illustrates the same point. Depending on the analysis, the ECM subject either moves to the embedded TP where it is “exceptionally” assigned case, or it moves to a case position in the matrix clause (as represented here). It then *wh*-moves to the left periphery.

There is a general consensus that a *wh*-moved phrase cannot be A-moved (i.e. moved to a case position). We have already seen some examples illustrating the factual correctness of this claim at the beginning of this chapter. (16) and (17) illustrate the same point.

- (16) a. I know who to invite.
 b. *[IP Who_i is known [CP t_i to invite t_i]]?
- (17) a. They asked which book John had read.
 b. *[IP Which book_i was asked [CP t_i John had read t_i]]? (from Neeleman and van de Koot (2010))

Assuming that objects move to case positions, any example involving scrambling or topicalization of objects might be taken as evidence that a step of case-driven movement feeds both scrambling and topicalization. In any case, it is clear that case-marked phrases can be both scrambled and topicalized. Grewendorf (2003) also provides the following example of a German ECM-construction with the unaccusative verb *misslingen* ‘fail’ to show that A-movement can feed scrambling. The surface subject of the unaccusative verb (italicized in the examples below) undergoes A-movement, followed by scrambling out of the infinitive:

- (18) a. dass der Professor *das Experiment_i* seinem Kollegen *t_i*
 that the professor.NOM the experiment.ACC his colleague.DAT
 misslingen sah.
 fail saw
 ‘that the professor saw the experiment fail on his colleague.’
 (my translation)
- b. dass *das Experiment_i* der Professor *t_i* seinem Kollegen
 that the experiment.ACC the professor.NOM his colleague.DAT
t_i misslingen sah.
 fail saw

The Hierarchy of Operations states that scrambling and A-movement are asymmetrically ordered. However, while evidence can be adduced to show that A-movement can feed scrambling, as we have just seen, it is extremely difficult to either prove or disprove the claim that the opposite ordering of these operations is illicit, i.e. that scrambling cannot feed A-movement. What is more, if scrambling in (14) should be defined as an operation targeting a pre-subject position (in German), then in fact all case positions would be lower than the scrambling position. An appeal to the Extension condition, which requires movement always to target a c-commanding position, would then rule out the possibility of scrambling feeding A-movement quite independently of the Hierarchy of Operations. To circumvent this interfering factor, one would have to focus on cases involving cross-clausal movement, i.e. have an element scramble in the embedded clause and then move to a case position in the matrix clause. However, this test cannot be carried out given that A-movement is strictly clause-bound.

Whether scrambling can feed *wh*-movement and topicalization is a debated issue. It is often noted that *wh*-phrases in German are not allowed to (freely) scramble (Fanselow (1990), Müller and Sternefeld (1996), Pesetsky (2000) among others). (19a) shows the normal order in German multiple questions, with one *wh*-phrase fronted to the left periphery and the other remaining in situ. (19b) shows that the object *wh*-phrase cannot be scrambled over the subject. Assuming then that *wh*-phrases which do not move to SpecCP in overt syntax, do so at LF, examples like (19b) are taken to show that *wh*-movement cannot follow scrambling. This is the view assumed for instance by Müller and Sternefeld (1993). Note that [-*wh*] phrases can scramble in the same context, as illustrated in (19c) (examples from Müller and Sternefeld (1996)).

- (19) a. Wie hat der Fritz was repariert?
 how has the Fritz what fixed
- b. *Wie hat was der Fritz <was> repariert?
 how has what the Fritz fixed
- c. Wie hat das Auto der Fritz <das Auto> repariert?
 how has the car the Fritz repaired

However, it has been noted that *wh*-words can scramble over quantified and other operator-like NPs, as the examples in (20) illustrate.⁹ In fact, scrambling of the second *wh*-phrase seems to be required in these contexts, considering that the failure to scramble yields a deviant result.¹⁰ If there is a general prohibition against *wh*-moving scrambled phrases, then examples like those in (20) become problematic. Fanselow (2001) thus argues that the ungrammaticality observed in cases like (19b) has nothing to do with the feeding relation between scrambling and *wh*-movement, but is rather due to a violation of the ban against placing an indefinite to the left of a definite NP.¹¹

- (20) a. (i) ??Wer hat niemanden wo angetroffen?
 who has nobody where met
 ‘Who didn’t meet anybody where?’
 (ii) Wer hat wo niemanden angetroffen?
 who has where nobody met
 b. (i) ?*Wann hat wem der Mann geholfen?
 when has who the man helped
 ‘When did the man help who?’
 (ii) Wann würde wem nur ein Held helfen?
 when would who only a hero help
 ‘When would only a hero help whom?’

Many other scrambling languages allow *wh*-phrases to scramble more freely than German does, further suggesting that scrambling can in principle feed *wh*-movement. Consider for instance Japanese. It is well-known that Japanese is a *wh*-in-situ language, i.e. that *wh*-phrases typically remain in their base generated positions, as illustrated in (21a). (21b) shows that a *wh*-phrase can be displaced, though it need not end up in the C domain.¹²

⁹(20a) is taken from Beck (1996), while (20b) is from Fanselow (2001).

¹⁰With respect to (20a), Beck notes the following: “[20a(ii)] needs a good context (e.g. a conversation about deliveries in a pizza service). If a good context is provided, the sentence is fine. (20a(i)), on the other hand, is ungrammatical, no matter how good a context is provided.” (Beck (1996), p. 6)

¹¹An interesting alternative way of analysing Beck’s data has been put forth by Grohmann (2000). He argues that both *wh*-phrases always overtly front to the C-domain in German, while everything that intervenes between the two *wh*-phrases has been moved there. The examples in (20a) then do not show that the adverbial *wh*-phrase has scrambled over *niemandem*, but rather that *niemandem* cannot move to a position between the two *wh*-phrases. If this account is on the right track, then neither examples like (19b), nor those in (20) can tell us something about the interaction between scrambling and *wh*-movement.

¹²Examples are provided by Naoyuki Yamato p.c. and modelled on an example from Nemoto (1999). See also Saito (1989) and Grohmann (2000).

- (21) a. Taroo-ga Hanako-ni dare-no sensei-o syookaisita no?
Taro-NOM Hanako-DAT whose teacher-ACC introduced Q
'Whose teacher did Taroo introduce to Hanako?'
- b. Taroo-ga [dare-no sensei-o]_i Hanako-ni t_i syookaisita no?
Taroo.NOM whose teacher.ACC Hanako.DAT introduced Q

Serbian is another language which allows *wh*-phrases to scramble. Unlike Japanese where *wh*-phrases typically remain in situ, in Serbian no *wh*-phrase is allowed to stay in its base position, as witnessed by the ungrammaticality of (22b).

- (22) a. Koga_i je čime_j David udario t_i t_j?
who.ACC aux what.INSTR David hit
'Who did David hit with what?'
- b. *Koga_i je David udario t_i čime?
who.ACC aux David hit what.INSTR

Though all *wh*-phrases must be displaced in Serbian multiple questions, they need not all move to the left periphery. Examples in (23) show that it suffices to scramble the second *wh*-phrase to the left of the verb.

- (23) a. Koga_i je David čime_j udario t_i t_j?
who.ACC aux David what.INSTR hit
- b. Kome_i je David šta_j dao t_i t_j?
whom.DAT aux David what.ACC give

Patterns such as these observed in Serbian and Japanese suggest that the ungrammaticality of German examples like (19b) is not due to the fact that scrambling cannot be followed by *wh*-movement, a conclusion which is further corroborated by the observation that *wh*-phrases in German can in fact scramble under the right conditions. Furthermore, focusing on the distribution of weak crossover effects in German, Wiltschko (1998) provides an additional argument supporting the same conclusion, namely that scrambling can feed *wh*-movement. As already noted, scrambling in German does not give rise to WCO effects. (24) shows that no violation arises when the scrambled phrase binds both a trace and a pronoun in an argument position that c-commands the trace. Another well-known property of scrambling in German is that it is clause-bound, as illustrated in (25) (examples from Abels (2008b)):¹³

- (24) a. Früher haben [jeden Studenten]_k seine_k Eltern t_i unterstützt.
earlier have every student.ACC his parents.NOM supported
'In the past his_k parents supported every student_k'
- b. *Früher haben seine_k Eltern [jeden Studenten]_k unterstützt.
earlier have his parents.NOM every student.ACC supported

¹³By clause-bound I mean that scrambling in German cannot cross the finite CP boundary.

- (25) *Gestern hat [jeden Studenten]_i Hans gesagt, dass seine Eltern _{t_i}
 yesterday has every student.ACC Hans said that his parents
 unterstützen.
 supported

Interestingly, short *wh*-movement does not give rise to weak crossover effects (26a), whereas long *wh*-movement does (26b), but only in the higher clause.¹⁴

- (26) a. Welchen Studenten_k unterstützen seine_k Eltern?
 which student support his parents
 ‘Which student_k do his_k parents support?’
 b. *Welchen Studenten_k glauben seine_k Eltern, dass Maria unterstützt?
 which student believe his parents that Maria supports
 ‘Which student_k believe his_k parents that Maria supports?’

Wiltschko (1998) argues that this pattern receives a straightforward explanation if scrambling can feed *wh*-movement. In (26a), a *wh*-phrase would undergo scrambling, obviating the weak crossover, before undergoing *wh*-movement. Since scrambling is clause-bound, the ungrammaticality of (26b) is expected. According to Abels (2008b), the paradigm above can be replicated for topicalization by replacing the word *welchen* ‘which’ by *jeden* ‘every’, with judgements remaining the same. It can then be argued along the same lines that scrambling can feed topicalization. I conclude therefore that mittelfeld scrambling can in principle feed both *wh*-movement and topicalization.

The reverse feeding relations are illicit. If *wh*-movement could feed scrambling, we wouldn’t expect this type of scrambling to be clause-bound. The fact that a phrase cannot undergo one step of *wh*-movement to the embedded SpecCP, and then scramble to a position in the matrix clause, as illustrated in (27), speaks in favour of the conclusion that *wh*-movement does not feed scrambling (the example is from Abels (2007)).¹⁵

- (27) *Gestern hat [welchen Studenten] nur ein Meisterdetektiv wissen
 yesterday has which student only a master-detective know
 können, Hans beim Schummeln erwischt hat.
 could Hans at.the cheating caught has

¹⁴The data can be replicated in Serbian, with the same results. If the proposed analysis is on the right track, then examples of this sort might be taken as evidence that there is no long-distance scrambling in Serbian.

¹⁵Scrambling which can cross clausal boundaries, found in some languages such as Japanese, would have to constitute an operation distinct from the scrambling operation discussed here, with its placement in the Hierarchy of Operations yet to be determined. Treating long-distance and mittelfeld scrambling as two distinct operations is consistent with the observation that these two movement operations do not pattern alike in many respects (see Saito (1992, 1994), Grewendorf and Sabel (1999), among others). See however Grewendorf (2003) for a different view on this matter.

‘Yesterday only a world-class detective could have known which student Hans caught cheating.’

Finally, let us turn to the interaction between *wh*-movement and topicalization. The Hierarchy states that *wh*-movement should be able to feed topicalization, but that a topicalized phrase cannot be *wh*-moved. However, as Abels (2008b) notes, it seems to be impossible to establish a feeding relation between *wh*-movement and topicalization in either direction. He speculates that the reason for this might be semantic in nature. It is quite plausible to assume that the same argument cannot be topicalized and questioned at the same time. It is constituents which represent discourse-old information that typically undergo topicalization, while a *wh*-question asks for the identity of the relevant discourse referent. *Wh*-elements can thus be viewed as inherently non-topic elements for semantic reasons. This view seems to be reinforced by the fact that a topic marker in languages like Japanese and Korean does not attach to *wh*-elements. The fact then that *wh*-movement and topicalization do not interact would not bear on the Hierarchy of Operations.

The claim that *wh*-movement and topicalization do not enter into a feeding relation might potentially be challenged by examples such as (28), provided by Müller and Sternefeld (1993).

- (28) *Welchen Pudding_i sagt niemand [t_i würde sie t_i mögen]?*
 which pudding says nobody would she like

The *wh*-phrase moves to the matrix C domain from an embedded V2 clause. What is interesting is that *wh*-movement does not typically induce V2 in the embedded clause, while topicalization does, as the following contrast illustrates:

- (29) Ich glaube [[den Fritz]_i hat sie t_i gesehen].
 I believe the.ACC Fritz has she seen
- (30) a. Ich sagte [wen_i (daß) sie t_i gesehen hat.]
 I said who.ACC that she seen has
 ‘I said who she had seen.’
 b. *Ich sagte [wen_i hat sie t_i gesehen.]
 I said who.ACC has she seen

It might be argued then that the moved phrase first topicalizes to the left periphery of the embedded clause in (28), and then undergoes a step of *wh*-movement. Observe that this would yield the exact opposite ordering of these two operations to the one stated in (14). Even if examples such as these are set aside, note that by now there are no empirical facts in support of the proposed ordering. Grewendorf (1994) and Abels (2008b) argue that the pertinent evidence can be adduced once we look beyond cases involving consecutive movements of the same phrase and consider constructions where movements affect two phrases which are in a dominance relation. We will see however that the evidence presented is not as convincing as one would hope.

The relevant examples will be given in the following section, and discussed further in chapter 5.

Let me summarize the conclusions we have reached so far. As we have seen, there is relatively solid evidence supporting the following pairwise orderings of different movement operations:

- (31)
- a. A-mvt \prec *wh*-mvt
**wh*-mvt \prec A-mvt
 - b. A-mvt \prec topicalization
*topicalization \prec A-mvt
 - c. (mittelfeld) scrambling \prec *wh*-mvt
**wh*-mvt \prec (mittelfeld) scrambling
 - d. (mittelfeld) scrambling \prec topicalization
*topicalization \prec (mittelfeld) scrambling

It is also clear that A-movement can precede scrambling. Whether scrambling can feed A-movement is difficult to test. In addition, if we limit our attention to scrambling targeting a pre-subject position, then this feeding relation would be ruled out by the Extension condition and would not require a particular ordering to be imposed on these two operations. It is even less clear whether the ordering between *wh*-movement and topicalization is as stated in the Hierarchy. At least no evidence for the claim that *wh*-movement precedes topicalization seems to be available on the basis of identity cases. In the following section, I explore the claim that further support for the proposed Hierarchy of Operations can be found by looking at cases where the relevant movement steps affect two phrases that are in a dominance relation.

2.2 Non-identity cases

In the previous section, we looked at interactions between different movement operations in cases involving consecutive movements of the same phrase, as schematized below.

- (32) XP ... <XP> ... <XP>
-

Recently it has been claimed that the same feeding/bleeding patterns between different movement operations can also be observed in cases where the relevant movements affect two phrases which are in a dominance relation (Grewendorf (2003), Williams (2003), Abels (2008b)). In this section, I explore this claim by looking at examples involving remnant movement and subextraction from the moved phrase. The two configurations in question are schematized in (33) and (34). Following Abels (2008b), I will often refer to configurations involving consecutive movements of the same phrase as *identity cases*, and use the term *non-identity cases* to refer jointly to the two configurations below.

- (33) Extraction from a moved phrase
 $YP \dots [XP \dots \langle YP \rangle \dots] \dots \langle [XP \dots YP \dots] \rangle$
- (34) Remnant movement
 $[XP \dots \langle YP \rangle \dots] \dots YP \dots \langle [XP \dots \langle YP \rangle \dots] \rangle$

Before turning to the relevant data, two notes regarding the configurations in (33) and (34) are in order. Cases of remnant movement used to be ruled out by the Proper Binding Condition, which requires that every trace has a c-commanding antecedent at surface structure (Fiengo (1977)). The PBC was designed to rule out syntactic lowering, among other things. However, in recent years, remnant movement analyses of various phenomena have been steadily gaining ground and remnant movement has become less exotic and more widely accepted, leading to the conclusion that the PBC cannot be generally correct. The effects of the PBC then must be derived from other principles. Note in this respect that if the Extension Condition is assumed (see Chomsky (1995)), the PBC becomes redundant as a tool for ruling out downward movement. The Extension Condition requires that syntactic operations extend the tree at the root, and is presumably more fundamental as it derives both the prohibition against syntactic lowering and cyclicity. It however allows remnant movement.

Another restriction on movement that used to be widely assumed was the Freezing principle, specifically designed to rule out the configuration schematized in (33) (see Ross (1967), Müller (1998) etc.) However, it has often been pointed out that the Freezing principle is too restrictive, and that movement out of moved constituents should not in principle be disallowed. Some counterexamples to the Freezing principle will be discussed in this section. We will see that many illicit examples of configurations in (33) and (34) can be captured by the generalized version of the Ban on Improper Movement, casting further doubt on the need to postulate either the PBC or the Freezing principle.

Let us start examining the data that bear on the ordering between movement operations by looking at the less controversial cases, namely those involving A-movement on the one hand and *wh*-movement and topicalization on the other. The Hierarchy of Operations states that A-movement (i.e. case-driven movement) can feed both *wh*-movement and topicalization. That this is indeed correct for cases involving subextraction out of moved phrases can be shown on the basis of examples involving ECM and raising verbs.¹⁶ Assuming the now standard raising-to-object analysis of ECM

¹⁶The examples in (35) are taken from Abels (2008b), while those in (36) are from Chomsky (2008). Chomsky's (2008) article is noteworthy in the context of the present discussion since it argues that A-movement does not in fact feed *wh*-movement (A'-movement more generally) in certain constructions where such a feeding relation is standardly assumed. On standard assumptions, arguments of unaccusative and passive verbs, as in (i) below, first undergo movement from the object position to the SpecTP, and then are *wh*-moved to the

verbs (Postal (1974), Lasnik and Saito (1991), among others), the entire phrase *the manager of which football team* in (35a) for instance would move to a case position in the matrix clause.¹⁷ Subsequently, a part of the moved phrase would extract and undergo *wh*-movement to the C domain (cf. (5), (6) and (15) showing feeding of the same type for identity cases).¹⁸

- (35) a. [Which football team]_i do you believe [the manager of t_i] to have paid a large fine?
 b. [Which football team]_i do you want/expect [the manager of t_i] to pay a large fine?
 c. [Which politician]_i do you believe [the rumors about t_i] to be false?
- (36) a. It is the car (not the truck) of which the driver is likely to cause a scandal.
 b. Of which car is the driver likely to cause a scandal?
 c. Of which car did they believe the driver to have caused a scandal?

An argument to the effect that A-movement can feed *wh*-movement in cases of subextraction can also be made on the basis of examples like (37a) (from Starke (2001)), which present exceptions to the subject island condition. According to Starke (2001), (37a) is only slightly worse than the corresponding object extraction case, (37b). The same appears to be true in English.

- (37) a. [de quel film]_i est-ce que tu crois que [la première partie t_i] va
 of which film is-it that you think that the first part goes
 créer un scandale?
 create a scandal
 ‘Which movie do you think that the first part of would create a scandal?’

left periphery if they bear a *wh*-feature. Chomsky suggests a novel analysis of these cases, whereby both A-movement and *wh*-movement proceed in parallel from the base position.

- (i) Who fell/was arrested?

I will discuss Chomsky’s analysis in more detail in the following chapter. What is important to bear in mind now is that though Chomsky (2008) rules out the possibility of A-movement feeding A’-movement in constructions such as (i), he doesn’t in general rule out this feeding relation. Thus, even for him, such a feeding relation would exist in cases like (35) and (36).

¹⁷I discuss ECM-constructions in more detail in chapter 4.

¹⁸Chomsky (1973) marks similar examples as ungrammatical (see also Kayne (1984)). As Abels (2008b) notes, one important difference between the present examples and Chomsky’s original one is that Chomsky’s example contained a non-D-linked *wh*-phrase. More recently, Chomsky (2008) has judged cases like (36c) involving extraction from an ECM subject as acceptable.

- b. [de quel film]_i as tu raté [la première partie t_i]?
 of which film have you missed the first part
 ‘Which film did you miss the first part of?’

The examples discussed above involve *wh*-movement, but the same conclusion can be reached for the interaction between A-movement and topicalization. Abels (2008b) provides the following German example in support of this claim, modelled on Starke’s (37a):

- (38) [Von diesem Film]_i hat [der erste Teil t_i] doch letztes Jahr einen großen
 of this film has the first part PRT last year a big
 Skandal ausgelöst.
 scandal caused
 ‘The first part of this film caused a big scandal last year.’

The same feeding patterns are also licit in cases of remnant movement. The following examples show that A-movement can feed remnant *wh*-movement. In (39b), for instance, *Oscar* is extracted from the infinitival clause and moved to SpecTP of the embedded clause. This movement step is followed by *wh*-movement of the remnant infinitival clause (from Abels (2008b)).

- (39) a. [How likely t_i to win the race]_j is John_i t_j?
 b. Max asked how likely to win Oscar was.
 [TP Max asked [CP [AP how likely t_i to win]_j [TP Oscar_i was t_j]]]

The claim that A-movement feeds remnant topicalization is supported by the following examples from English and German, respectively ((40a) is from Abels (2008b), (40b) is from Grewendorf (2003)). In (40b), the infinitival clause has been topicalized, and the embedded object has undergone long passivization, which is possible in German with certain verbs such as *versuchen* ‘try’.

- (40) a. [Criticized t_i by his boss]_j John_i has never been.
 b. [t_i zu küssen]_j wurde der Student_i von Maria t_j versucht.
 to kiss was the student.NOM by Mary tried

The reverse feeding relations are illicit, i.e. neither *wh*-movement, nor topicalization can feed A-movement in non-identity cases. (41) shows this for the combination of *wh*-movement and A-movement.

- (41) a. *Oscar was asked how likely to win it was. (Collins (2005))
 [TP Oscar_i was asked [CP [AP how likely t_i to win]_j [TP it was t_j]]]
 b. *A picture of is known which king to have been sold. (Abels (2007))
 [TP [A picture of t_i]_j is known [CP [which king]_i to have been sold t_j]].

Example (41a) shows that a phrase cannot A-move out of a *wh*-moved constituent (cf. this example to the well-formed (39b)). On the intended construal the sentence should mean something like *It was asked how likely it was that Oscar would win*. (41b) is a remnant movement configuration. The *wh*-phrase, *which king*, extracts from the dominating DP and moves to the left periphery of the embedded clause. Subsequently, the remnant DP fronts to the matrix SpecTP. On the basis of examples such as these, it can be concluded that clear instances of A'-movement operations, such as *wh*-movement and topicalization, follow A-movement in non-identity cases as well, i.e. non-identity cases pattern in this respect with cases involving consecutive movements of the same phrase.

Let us now turn to the interaction between A-movement and scrambling. Recall that by scrambling in this case we mean clause bound movement to a pre-subject position (in a language like German). Abels (2008b) provides the example (42b) as evidence that A-movement can feed scrambling in cases of subextraction. The argument presupposes that objects in German move to case position. It can be demonstrated that the prepositional adverbial *darüber* has indeed moved out of the NP by observing the specificity effect.

- (42) a. ... weil niemand [{ ✓ein/✓das } Buch darüber] lesen wollte.
 ...because nobody a/the book there-about read wanted
 '...because nobody wanted to read a / the book about that'
- b. ... weil darüber_i niemand [{ ✓ein/*das } Buch t_i] lesen wollte.
 ...because there-about nobody a/the book read wanted
 '...because nobody wanted to read a / *the book about that'

The following examples show that the combination of A-movement and remnant scrambling is also possible (from Abels (2008b) and Grewendorf (2003), respectively). In (43b) for instance, the first step involves long passivization of the embedded object to the matrix clause, followed by scrambling of the remnant infinitival clause.

- (43) a. ?... dass [von einem Studenten t_j angefasst]_i [kein einziges
 that by a student touched no single
 Reagenzglas]_j t_i werden durfte
 test.tube become may
 '... that no student was allowed to touch a test tube.'
- b. ... weil [t_j zu küssen]_i der Student_j von Maria t_i versucht
 since to kiss the student.NOM by Mary tried
 wurde.
 was

The Hierarchy states that the opposite feeding relation, i.e. the one where scrambling precedes A-movement should be impossible. As already noted for the identity cases, this particular prediction is difficult to test and relevant cases might in fact be ruled

out by factors entirely independent of the Hierarchy of Operations. This is because scrambling in the context of this discussion has been defined as movement to a position preceding the subject. Assuming the Extension condition, once a phrase has moved above the subject position, subsequent movement steps cannot target the subject position. Likewise, scrambling followed by A-movement of the object should be impossible as the object case position is lower than the scrambling position.

Consider now the following examples, argued by Grewendorf (2003) to show that scrambling cannot feed remnant A-movement. Though A-movement and mittelfeld scrambling cannot cross the finite clause boundary, it has been observed that these two movement operations can extract constituents out of certain infinitival clauses in German, such as those selected by the verb *versuchen*, ‘try’. This is shown for scrambling in (44a), and for A-movement in (44b). Note that the adverbial *vergeblich* ‘in vain’ can only modify the matrix verb.

- (44) a. weil Maria ihn_i vergeblich [t_i zu küssen] versuchte.
 since Mary him in.vain to kiss tried
 ‘since Mary tried in vain to kiss him.’
 b. weil [ihn zu küssen]_i von Maria vergeblich t_i versucht wurde.
 since him to kiss by Mary in.vain tried was

Grewendorf (2003) argues that in such environments, the combination of scrambling and remnant A-movement is illicit. In (45), first the embedded pronominal object is scrambled, and then the remnant infinitival clause is A-moved.

- (45) a. *weil [t_j zu küssen]_i ihn_j von Maria vergeblich t_i versucht wurde.
 since to kiss him by Mary in.vain tried was
 b. *weil [t_j zu küssen]_i von Maria ihn_j vergeblich t_i versucht wurde.
 since to kiss by Mary him in.vain tried was

There are two objections we might raise against the relevance of examples in (45). First of all, if scrambling is defined as movement to a pre-subject position for the purposes of the Hierarchy, then the ungrammaticality of such examples is irrelevant since the landing site of the scrambled phrase is lower than the subject position.¹⁹

Secondly, German differs from English in that A-movement in passive constructions is not obligatory (see Grewendorf (1989), Müller (1995), etc.). An internal argument can move to the Spec of IP, but it can also remain in its base position.

¹⁹This objection does not extend to Grewendorf’s analysis as his use of the term scrambling is broader in scope, and applies to movement targeting an adjoined position. If both left and right adjunction to AgrSP are allowed (a point which Grewendorf does not discuss), then appealing to the Hierarchy of Operations might provide one way of excluding examples in (45).

There is another clause-internal scrambling operation in German, analyzed typically as adjunction to VP (or AgrOP), which could potentially provide a lower landing site for scrambled phrases. This type of operation is not discussed by Abels (2008b). I will thus leave it aside for now, but will return to it once I have spelled out the proposal I will be suggesting.

- (46) a. daß [IP [VP dem Peter das Buch gegeben] wurde]
 that the.DAT Peter the.NOM book given was
 b. daß [IP das Buch_i [VP dem Peter t_i gegeben] wurde]
 that the.NOM book the.DAT Peter given was

According to Grewendorf (1989), this optionality is due to the availability of an empty expletive *pro* in German, which occupies the SpecIP position and transmits nominative case to a VP-internal NP. Under this view, an empty expletive is not present in (46b); hence, A-movement to SpecIP occurs to satisfy the Case Filter. In (45), there is thus no reason why the infinitival clause would have to move to the matrix IP, given that it does not need to check Case features. Hence, cases such as these might in fact involve two instances of scrambling, namely scrambling of the pronoun out of the infinitive and scrambling of the remnant infinitive, as suggested by Sauerland (1996) for similar examples. It is well-known that scrambling cannot affect a phrase out of which another phrase has scrambled, as illustrated in (47).²⁰ The same restriction might then rule out examples in (45).

- (47) a. *Danny hat [t_i zu putzen]_j vergeblich [das Bad]_i t_j versucht.
 Danny has to clean unsuccessfully the bathroom tried
 (from Sauerland (1996))
 b. *dass [t_i zu lesen]_j [das Buch]_i keiner versucht hat
 that to read the book noone tried has
 (from Müller (1998))

Thus, although there seem to be no counterexamples to the claim that scrambling cannot feed A-movement, it is important to bear in mind that constraints independent of the Hierarchy of Operations might rule out such derivations.

Turning to the interaction between scrambling on the one hand, and *wh*-movement and topicalization on the other, Abels (2008b) provides the following examples to show that scrambling can feed *wh*-movement ((48a) and (49a)), and topicalization ((48b) and (49b)) in both subextraction and remnant movement cases:²¹

- (48) Extraction out of a moved phrase
 a. Worüber_i kann [einen Südkurier-Artikel t_i] selbst Peter nicht am
 what.about can a Südkurier-article even Peter not at.the
 Stand verfassen?
 beach write
 'For which topic is it the case that even Peter cannot write an article
 about it for the Südkurier when he is at the beach?'

²⁰See Müller (1998) for a list of relevant references. Müller (1998) argues that the generalization is in fact broader in scope, namely that remnant creating movement and remnant movement can never be of the same type.

²¹Abels (2008b) also notes that these examples are not accepted by all speakers.

- b. [Über Karl den Großen]_i hatte [ein fertiges Manuskript t_i]_j
 about Charlemagne had a done manuscript
 leider keiner der anwesenden Historiker t_j anzubieten.
 unfortunately none the present historians to.offer
 ‘About Charlemagne, none of the historians present could offer a finished manuscript.’

(49) Remnant Movement

- a. [Was für ein Buch t_i]_j hat [über die Liebe]_i niemand t_j gelesen?
 what for a book has about the love nobody read
 ‘What type of book about love has nobody read?’
- b. [[Ein Buch t_i] gelesen]_j hat darüber_i keiner.
 a book read has there.about noone
 ‘A book about that, nobody has read.’

The opposite feeding relations are not possible, as illustrated here for cases of subextraction (from Abels (2008b)):

- (50) a. *Gestern hat [über Karl den Großen]_i keiner gefragt, was für
 yesterday has about Charlemagne nobody asked what for
 [eine Arbeit t_i] er schreiben soll.
 a work he write should
 ‘Nobody asked yesterday what kind of paper about Charlemagne he should write.’
- b. *Auf dem Kongress hat [über Karl den Großen]_i keiner gesagt [ein
 at the congress has about Charlemagne nobody said a
 fertiges Manuskript t_i] habe er anzubieten.
 done Manuscript have he to.offer
 ‘At the congress nobody said that he had a completed manuscript about Charlemagne to offer.’

Of course, we might argue that examples in (50) are ruled out because scrambling is clause-bound. However, the clause-boundedness itself is plausibly just another property that an articulated theory of improper movement should capture. We might ask why the scrambled PP can’t be smuggled out of the clause (to use Collins’ (2005) term) by the *wh*-phrase that contains it. We will address this issue in more detail in the following section.

Finally, let us see whether the interaction between *wh*-movement and topicalization conforms to the proposed Hierarchy of Operations. The Hierarchy given in (14) states that *wh*-movement asymmetrically feeds topicalization. Recall that no evidence for this ordering could be adduced from the identity cases, given that the two operations in question cannot affect one and the same constituent. Abels (2008b) argues that the evidence for the proposed ordering can be found once we turn to non-identity cases. Consider the pair of examples in (51). (51a) shows that topicalization

out of a *wh*-moved constituent, though degraded, is possible in German, and contrasts with cases of *wh*-movement out of a topicalized constituent in (51b).

- (51) a. ??[Über Karl den Großen]_i weiß ich nicht, [was für ein Buch t_i]_k er t_k
 about Charlemagne know I not what for a book he
 schreiben will.
 write wants
 ‘About Charlemagne I don’t know what kind of book he wants to write’
- b. *[Über welchen deutschen Kaiser]_i sagt er [ein fertiges
 about which German emperor says he a completed
 Manuskript t_i]_j hat keiner t_j anzubieten.
 manuscript has nobody to.offer
 ‘About which German emperor does he say that a completed manuscript
 nobody can offer?’

The same conclusion can be drawn on the basis of examples involving remnant movement. The relevant cases with *wh*-movement feeding remnant topicalization are given in (52) (from Abels (2008b) and Grewendorf (2003), respectively). The sentences are degraded as they involve extraction out of a *wh*-island. However, both Grewendorf (2003) and Abels (2008b) note that they are no more degraded than (53), where topicalization takes a complete, rather than a remnant category, out of a *wh*-island. The opposite feeding relation is not possible.

- (52) a. ??[Hans t_i zu geben]_j weiß ich nicht was_i Fritz t_j abgelehnt hat.
 Hans to give know I not what Fritz refused has
 ‘I don’t know what Hans refused to give Fritz.’
- b. ??[t_j zu überreden]_i weiss ich nicht wen_j sie t_i versucht hat.
 to persuade know I not who.ACC she tried has
- (53) a. ??[Hans das Buch zu geben]_i weiß ich nicht warum Fritz t_i abgelehnt
 Hans the book to give know I not why Fritz refused
 hat.
 has
 ‘I don’t know why Fritz refused to give Hans the book.’
- b. ??[Den Studenten zu überreden]_i weiss ich nicht warum sie t_i
 the student.ACC to persuade know I not why she
 versucht hat.
 tried has
 ‘I don’t know why she tried to persuade the student.’

The German patterns just discussed thus seem to support the claim that *wh*-movement asymmetrically feeds topicalization. However, this claim is challenged by the English data. According to Lasnik and Saito (1992), in English, it is at least marginally possible to *wh*-move out of topicalized constituents, as illustrated below.

- (54) a. ??Who_i do you think that [pictures of t_i]_j, Mary believes t_j are on sale?
 b. ??Who_i do you think that [pictures of t_i]_j, John wanted t_j?

It might even be possible to construct better-sounding examples than those of Lasnik and Saito. (55) is an attempt to do so.

- (55) (?)? [Which emperor]_i did you say that [even pictures of t_i]_j John worships t_j?

If any conclusions can be reached on the basis of these grammaticality judgements, then these examples suggest that topicalization can feed *wh*-movement in English.²² Therefore, if these two operations in English are ordered, then it seems that the ordering should be as stated in (56). Note that this is the exact opposite of the ordering proposed on the basis of German data, challenging the universality of the established Hierarchy of Operations.

- (56) topicalization \prec *wh*-movement

Both Grewendorf (2003) and Abels (2008b) note this problem and suggest that it can be dealt with if either *wh*-movement or topicalization are not the same kind of operation in German and English. If the relevant operations in the two languages are distinct, then discrepancies regarding the Hierarchy of Operations are not unexpected. Considerations such as these reveal how important it is to clearly define criteria that determine what counts as an instance of a particular operation if the proposal is to be falsifiable. Abels (2008b) however remains quite vague on this point, noting only that the feature-types that enter into computation of relativized minimality/attract closest define the operational types.²³ Grewendorf (2003) is a bit more explicit on this issue. The hierarchy he proposes makes reference to the following operations:

- (57) A-movement \prec Adjunction movement ('scrambling') \prec A'-movement as non-operator movement \prec A'-movement as operator movement

At first site, Grewendorf's hierarchy seems to state exactly the opposite ordering for *wh*-movement and topicalization than the one proposed by Abels (2008b), given that *wh*-movement is standardly treated as an operator movement, and topicalization as an instance of a non-operator movement.²⁴ In this respect, English facts such as (54) and (55) are expected and conform to the predictions of the Hierarchy. What is prima

²²The examples are not fully acceptable, but neither are the German examples discussed earlier.

²³This requires a substantially refined view of classes relevant for computing relativized minimality effects, as in the recent work of Starke (2001), Rizzi (2006).

²⁴Sternefeld (1993) states the hierarchy in terms of landing sites, distinguishing between movement to an object position (=SpecV), a subject position (=SpecI), an adjunction position (=scrambling position), an operator position (=SpecC), and 'topicalization' position (=Vorfeld position in a V2 clause). His view thus resembles more closely Abels's proposal. It should be however noted that Sternefeld only considers identity cases.

facie unexpected is the German pattern of feeding/bleeding relations. Grewendorf argues that particular properties of *wh*-movement and topicalization in German make this pattern less puzzling. In his view, the key diagnostic for determining the status of a particular operation as an operator or a non-operator movement is whether or not it shows WCO effects. In German, short *wh*-movement thus counts as a non-operator movement since it does not display WCO effects (see (26a)). On the other hand, topicalization (as in the pertinent examples in (51a) and (52)) would in fact be an instance of operator movement since topicalization out of *wh*-islands triggers WCO effects, as illustrated below.

- (58) *Den Studenten_i weiss sein_i Vater nicht wer t_i gesehen hat.
 the student.ACC knows his father.NOM not who seen has
 ‘His_i father doesn’t know who saw the student_i.’

Thus, both approaches in principle have a way of capturing the discrepancy between German and English data by relying on factors that determine what counts as an instance of a particular operation.

Consider next the following example:

- (59) ?[That medicine]_i I don’t know [how much of t_i]_j I should take t_j.

In this case, *wh*-movement feeds topicalization in a configuration involving subextraction out of a moved phrase. Given that both (59) and (54)/(55) are (at least marginally) acceptable, we might conclude that the feeding relation between *wh*-movement and topicalization can in fact go in both directions. If only asymmetric feeding relations are allowed, as Abels (2008b) hypothesizes, then further distinctions between different types of *wh*-movement and/or topicalization in English seem to be required to capture these patterns. The upshot of this discussion is that it is not entirely clear that *wh*-movement and topicalization do indeed enter into an asymmetric feeding relation. We might interpret these facts as indicating that no ordering is imposed on these two operations, and attribute the ungrammaticality of German configurations where topicalization feeds *wh*-movement to an independent factor. A proposal along these lines will be developed in chapters 4 and 5.

2.3 Clause-boundedness

Consider again cases of superraising, such as those below.

- (60) a. *John seems (that) likes cheese.
 [TP John seems [CP <John> (that) [TP <John> [VP <John> likes cheese]]]]
 b. *John seems (that) it likes cheese.
 [TP John seems [CP <John> (that) [TP it [VP <John> likes cheese]]]]

- c. *Who seems (that) likes cheese?
 [TP who seems [CP <who> (that) [TP <who> likes cheese]]]

It is standardly assumed that locality forces any phrase moving out of the CP to stop off in the Spec of CP. Given this assumption, examples in (60) can be ruled out as instances of improper movement. The DP first undergoes a step of A'-movement to the embedded C domain, and then A-moves to the matrix TP. Thus, the clause-boundedness of A-movement can be captured by appealing to the BOIM. It remains to be shown how, if at all, the ungrammaticality of such examples could be accounted for by a theory of improper movement that relies on finer-grained distinctions in movement types.

Whether clause-boundedness of certain operations can be captured in a theory that appeals to a Hierarchy of Operations will ultimately depend on the way operations relevant for the Hierarchy are defined. If *wh*-movement is defined either as a movement operation which checks a *wh*-feature or a movement that targets a position in which the *wh*-expression is interpreted, then fronting of the *wh*-phrase to the embedded CP in (60c) would not count as an instance of *wh*-movement. It is even clearer that A'-movement steps in (60a) and (60b) would not be instances of *wh*-movement on this view, since no *wh*-expressions are involved. If the operations relevant for the Hierarchy are defined in terms of the landing sites, as on Sternefeld's view (1993), then examples in (60) would fall under the purview of the Hierarchy of Operations.²⁵ In that case, these examples would be treated in the same way as those in (61), where a *wh*-phrase moves to the embedded C-domain, where it is interpreted, and then subsequently fronts to the Spec of matrix TP.

- (61) *[IP what_i was asked [CP t_i [IP to read t_i]]] ?

I take it that an updated theory of improper movement that refines the taxonomy of movement relations and redefines the BOIM accordingly should at least capture all the cases that the standard theory of improper movement did. Our goal in the following chapters will thus be to develop a theory of improper movement that would capture the observed feeding/bleeding relations between different movement operations, while keeping examples such as those in (60) under the purview of such a theory.²⁶

²⁵See Müller and Sternefeld (1993) for a theory of improper movement that relies on a similar definition of operations.

²⁶Some of the illicit ordering relations might be ruled out by Rizzi's (2006) Criterial Freezing, a principle which prohibits a phrase moved to what Rizzi calls a Criterial position, from moving further. Criterial positions are positions associated to certain scope/discourse properties, such as *wh*, topic, focus, etc. Thus, operations such as *wh*-movement or topicalization would not be expected to feed any other movement operation. However, Criterial Freezing does not capture a large set of illicit feeding relations and is thus insufficient as a theory of improper movement. Examples such as those in (60) cannot be captured by Cri-

2.4 Summary and conclusions

In this chapter, I have focused on interactions between various movement operations, relying on a finer-grained taxonomy of movement than the traditional binary division into A- and A'-movement types. In particular, I have been concerned with the issue of whether there are restrictions on ordering of movement operations such that a hierarchy of operations, as given in (54), might be proposed:

$$(62) \quad \theta \prec \text{A-mvt} \prec (\text{clause internal}) \text{ scrambling} \prec \text{wh-mvt} \prec \text{topicalization}$$

Though the predictions made by the Hierarchy of Operations are largely supported by the data reviewed, some pairwise orderings between different movement operations are more difficult to test than others. One case in point is the combination of A-movement and middlefield scrambling, as we have seen. What is more, the possibility of scrambling feeding A-movement might be excluded by factors entirely independent of the Hierarchy of Operations. As regards the interaction between *wh*-movement and topicalization, it is far from clear how these movement operations should be ordered. In Abels's (2008b) view, *wh*-movement precedes topicalization, while for Grewendorf (2003), the canonical cases of *wh*-movement in fact follow topicalization. As we have seen, there is evidence suggesting that finer-grained distinctions need to be made between different types of topicalization and/or *wh*-movement, if we are to assume that movement operations are always asymmetrically ordered. Alternatively, we might interpret the relevant facts as suggesting that *wh*-movement and topicalization are in fact not ordered. The analysis to be proposed in the following chapters will lead us to such a conclusion.

Setting these two problematic cases aside for now, the key observation stated as the standard BOIM, namely that (at least some) operations are asymmetrically ordered with respect to each other, is corroborated by the data investigated here. Even assuming a finer-grained typology of movement, asymmetric feeding relations between different movement operations can be observed. What is more, the same ordering patterns seem to hold both in cases involving consecutive movements of the same phrase, as well as in configurations involving remnant movement and subextraction from a moved phrase. These observations call for an explanation. The following chapters thus aim at developing an articulated theory of improper movement, capable of capturing the illicit orderings between various movement operations both in identity and in non-identity cases.

terial Freezing since the moved phrases are not interpreted in the embedded C domain. In addition, the approach does not extend to non-identity contexts as Rizzi explicitly states that subextraction from a phrase occupying a Criterial position is allowed.

Chapter 3

Previous analyses of improper movement

Why is movement from an A'- to an A-position prohibited? The question of what underlies this prohibition has been one of the central concerns ever since it was first formulated (Chomsky (1973)). In this chapter, I will discuss some of the previous accounts of improper movement. I'll start off with the classic GB analysis, before turning to more recent proposals.

3.1 The binding-theoretic approach

In the Government and Binding framework, traces of movement were assumed to be subject to the principles of the binding theory. Traces left by A-movement were treated as anaphoric and thus came under the purview of the Principle A of the binding theory. On the other hand, traces of A'-movement were argued to be variables. Variables counted as R-expressions for the purposes of the binding theory and were hence subject to the Principle C. Principle A requires that an anaphor be bound within its binding category, while Principle C states that an R-expression must be free, i.e. unbound. In addition, an element is thought to be bound if it is coindexed with a c-commanding antecedent *in an A-position*. Improper movement violations were then reduced to violations of binding conditions (Chomsky (1981), following May (1979)). Consider a standard case of superraising in English.

- (1) a. *John seems that it is certain to like ice cream. (Chomsky (1981))
b. [_{IP} John_i seems [_{CP} t'_i that [_{IP} it is certain [_{IP} t_i to like ice cream]]]]]

The trace in the most embedded IP (t_i) is a trace of an A'-movement and hence a variable subject to Condition C. However, it does not remain free as Condition C would require it to be, since it is bound by *John* in the matrix SpecIP. The ungrammaticality of (1) is thus attributed to a Condition C violation.

The reverse application of A- and A'-movements is correctly predicted to be legitimate on this approach. Consider (2). t_i is a locally bound anaphoric trace. On the other hand, t''_i and t'_i are variables which are not bound (since the antecedent is not in an A-position), conforming thus to principle C. No violation of the binding theory is incurred.

- (2) a. Who do you think was kissed?
 b. $[_{CP} \text{ who}_i \text{ do you think } [_{CP} t''_i [_{IP} t'_i \text{ was kissed } t_1]]]$

It is important to note that the binding theory can be invoked to rule out (1) only if there is a trace in the intermediate CP. In other words, in order to ensure the illformedness of (1), *John* must not be allowed to move from embedded IP to the matrix IP in one fell swoop, as illustrated in (3).

- (3) $[_{IP} \text{ John}_i \text{ seems } [_{CP} \text{ that } [_{IP} \text{ it is certain } [_{IP} t_i \text{ to like ice cream}]]]]]$

This type of derivation was ruled out by the Subjacency Condition in the GB, which prohibited movement from crossing more than one bounding node (where bounding nodes were IP and NP).¹ Hence, improper movement violations were captured by the binding theory only in tandem with the theory of locality.

In the following chapters, I will argue that improper movement effects can be captured by a theory of locality, conceived of here in terms of phases, without appealing to the principles of the binding theory, providing thereby a more parsimonious account of the phenomenon.² This is a welcome result given that various aspects of the binding theory have become problematic with the advent of Minimalism.³ As we will see, pursuing this line of thinking does not imply that we must abandon the GB intuition that common syntactic factors constrain both movement and binding dependencies. The final chapters of the dissertation will argue that phases play a crucial role in regulating possible binding relations, establishing thereby the link between movement and binding.

3.2 Müller and Sternefeld (1993)

Müller and Sternefeld (1993) point out that in addition to A'-movement not being able to feed A-movement, various types of A'-operations do not seem to feed each other either. Thus, both *wh*-movement and topicalization cannot affect the same phrase. In addition, they argue that the same is true for combination of scrambling and *wh*-movement/topicalization. These restrictions cannot be explained by Principle C of Chomsky's (1981) Binding theory. M&S thus set out to develop a theory of improper

¹In Chomsky (1986a), Subjacency was restated in terms of barriers. This difference is not relevant for the current discussion. See Chomsky (1986a) for details.

²In addition, the analysis to be outlined will have a broader scope since it will also regulate the sequencing of movement operations in non-identity cases.

³We will return to this issue in greater detail in chapters 7 and 8.

movement that would capture these cases, as well as those covered by the binding-theoretic analysis.

The key ingredient of their proposal is the following principle:

- (4) *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 that some of the different types of positions relevant for the PUB are SpecIP (the landing site of A-movement), SpecCP (the landing site of *wh*-movement), SpecTopP (the landing site of topicalization), and various left-adjunction sites (the landing sites of scrambling). The PUB then predicts that A'-movement to a certain kind of position cannot be followed by another movement step that ends up in a different kind of position. Observe how this captures the lack of interaction between topicalization and *wh*-movement for instance. A phrase undergoing *wh*-movement lands in a SpecC position. If the same phrase is subsequently topicalized, it will end up in SpecTopP. This leads to a configuration where the trace of the first movement step is bound ambiguously, both by its antecedent in the Spec of TopP and by the intermediate trace in Spec of CP, in violation of the PUB. Reversing the two operations produces the same result.

The analysis straightforwardly extends to the canonical cases of improper movement. Thus, in (5), *wh*-movement to SpecCP is followed by A-movement to SpecIP.⁴ Such a mixing of movement types is disallowed by the PUB. Locality, seen here in terms of Barriers, ensures that there is a stop-over in intermediate SpecCP.⁵

- (5) a. *John seems that it is certain to like ice cream.
 b. [IP John_i seems [CP t'_i that [IP it is certain [IP t_i to like ice cream]]]]]

The analysis still crucially relies on the A/A'-distinction. Only traces of A'-movement, but crucially not those of A-movement, are variables and therefore subject to the PUB. Consequently, a configuration where A-movement feeds A'-movement is predicted to be licit. In (6) for instance, no violation of the PUB is incurred since

⁴Note that what counts as *wh*-movement on this view is not movement that checks a [wh]-feature, but any movement that targets SpecCP.

⁵Müller and Sternefeld (1993) rely on the following definition of a *Barrier*:

- (i) *Barriers Condition*
 Movement must not cross a barrier.
- (ii) *Barrier*
 XP is a barrier for A iff
 (1) Xⁿ includes A
 (2) Xⁿ is not directly selected
 (3) X⁰ is distinct from Y⁰, where Y⁰ directly selects XP.

t_i is not a variable and t'_i is bound unambiguously.

(6) $[_{CP} \text{ who}_i [_{IP} t'_i \text{ was kissed } t_i]]$?

M&S argue that the theory of improper movement based on the PUB is superior to Chomsky's Principle C account because it rules out not only the illicit combination of an A'- and an A-movement, but also illicit combinations of different types of A'-movements.⁶ It is thus worth investigating how solid the evidence really is that different A'-movements do not enter into feeding relations. The A' movement operations that M&S discuss are scrambling, *wh*-movement, and topicalization. Consider the combination of scrambling and *wh*-movement/topicalization. Scrambling in M&S's view is any instance of adjunction operation to either VP, IP or CP. They further assume that scrambling is uniformly A'-movement across languages. This is a highly controversial assumption, as is clear from the discussion of scrambling in the previous chapter, but the one that is crucial for M&S. If scrambling were A-movement, then traces of scrambling would not count as variables, making the PUB irrelevant in regulating the application of this movement operation. Setting this concern aside, the evidence that M&S provide in favour of the claim that scrambling does not feed *wh*-movement/topicalization is based on the theory of barriers. They argue that if *wh*-movement were allowed to proceed via intermediate adjunction (i.e. scrambling) to either VP or IP, then Chomsky's (1986a) *Barriers* framework would not be able to derive the violation in extraction from clausal adjuncts. Consider (7).

(7) ??Who_i did you [_{VP} t'_i [_{VP} go home [without [_{IP} Mary [_{VP} t_i [_{VP} talking to t_i]]]]]]]?

The adjunct is a barrier for movement. However, if adjunction to VP is allowed, as illustrated above, then it is in fact the only barrier that intervenes between t_i and t'_i , thus the sentence cannot be ruled out by Subjacency. If intermediate adjunction is prohibited, as dictated by the PUB, then the ungrammaticality of (7) is captured since the moved *wh*-phrase would cross two bounding nodes, namely, the adjunct, and, by inheritance, the VP dominating the adjunct. On the basis of theory-internal considerations such as these, M&S conclude that the combination of scrambling and *wh*-movement is illicit.

On the other hand, in the previous chapter I have reviewed some evidence leading

⁶Rizzi's (2006) Criterial Freezing principle is very reminiscent of the PUB as it also rules out feeding relations between A'-movement operations. There are some significant differences though. Criterial Freezing cannot account for cases of superraising since the Spec of the embedded CP in this case would not be associated to scope/discourse properties and thus would not count as a freezing position. On the other hand, the PUB, unlike the Criterial Freezing, would not rule out cases such as (i), as these would involve movement from SpecCP to SpecCP.

(i) *Which book does Bill wonder [_{CP} t_i she read t_i]?

to the opposite conclusion, namely that scrambling can feed *wh*-movement. Recall that in languages like Serbian and Japanese, *wh*-phrases can be displaced though they need not move to the left periphery of the clause. Assuming that *wh*-phrases that do not overtly move to SpecCP do so at LF, such cases have been taken to suggest that scrambling can be followed by a step of LF *wh*-movement. M&S are aware of such data and they propose to handle facts of this type by parametrizing the level at which the PUB applies. Thus, in languages like Serbian and Japanese, the PUB is assumed to apply only at S-structure, but not at LF. Consequently, *wh*-phrases can be scrambled, and followed by an LF-movement of another type. In German, on the other hand, the PUB applies, they argue, both at S-structure and at LF, which is why the corresponding examples are ungrammatical in German. However, we have seen evidence suggesting that even in German scrambling can feed *wh*-movement. Recall that although overt scrambling of *wh*-phrases in German is not allowed as freely as in Japanese or Serbian, it is nevertheless possible under certain conditions (see chapter 2, section 2.1.2). The same conclusion is further corroborated by Wiltschko's argument based on the distribution of weak crossover effects in German, also reviewed in the previous chapter. Thus, if scrambling can in principle be followed by *wh*-movement and topicalization, as these facts seem to suggest, then the PUB in fact incorrectly rules out this option.

Given these considerations, empirical support for the PUB now reduces to the interaction between topicalization and *wh*-movement. As we have seen, it is indeed the case that topicalization and *wh*-movement cannot affect the same constituent. However, it could plausibly be argued that this restriction is due to a semantic incompatibility, rather than to an illicit feeding relation. The PUB is then not needed to account for these cases, and if scrambling can feed *wh*-movement/topicalization, it is also empirically inadequate.⁷ In any case, it is not obvious that the PUB has clear empirical advantages over the Principle C account reviewed in the previous section.⁸

3.3 Williams (2003)

Williams (2003) develops an alternative to the standard model of syntactic theory, with several interesting consequences, one of them relating to the ban on improper movement. In Williams's Representation theory, syntax is divided into multiple levels of representation, related to one another by shape-conserving mapping principles. Global shape conservation requires that two levels within a representation relation maximally correspond, modulo the introduction of functional material. Thus, grammar privileges mappings that conserve the linear and hierarchical relations across the

⁷See also Culicover (1996) for a critical review of this proposal.

⁸Note also that if ordering restrictions between movement operations in non-identity cases should also be covered by a theory of improper movement, as I have argued in the previous chapter, then a whole set of improper movements remains unaccounted for under this proposal (as well as under the Principle C account).

levels. The relevant levels include, but are not limited to, theta structure, case structure, predicate structure, surface structure, quantifier structure, focus structure and accent structure.⁹

An interesting component of Representation theory is the way embedding takes place. All embeddings are done in order of complement type, rather than in a bottom-to-top fashion, as on the standard approach. This is ensured by the Level Embedding Conjecture (LEC), which says that an item is embeddable exactly at the level at which it is defined, and no other. Consider how *that* clause embedding, as in (8), takes place on this view.

- (8) Bill believes that Mary is sick.
- a. Level of theta structure: [Bill believes that]_{VP} [Mary is sick]_{VP}
 - b. Level of predicate structure: [Bill believes that]_{IP} [Mary is sick]_{IP}
 - c. Level of surface structure: [Bill believes [that Mary is sick]]_{CP}_{CP}

The two clauses (matrix and embedded) are built in parallel. Embedding of the second clause occurs only when the level of surface structure is reached, and no earlier. Thus, while (9) is a licit embedding in the standard model, it is not a possible structure in Representation theory.

- (9) [_v V [_{CP} ...]]

This kind of architecture has a number of interesting consequences, one of them relating to the ban on improper movement. The RT levels determine different kinds of embedding, and are in addition associated with different kinds of movement operations. Thus, *that*-clause embedding and *wh*-movement take place at the level of surface structure, whereas NP-movement operations, such as raising and passivization, happen at an earlier level. Consider in this light an example of superraising:

- (10) *John seems [_{CP} that it is certain to like ice cream] (Chomsky (1981))

LEC dictates that both the embedded and the matrix clause must be at the same level of development for embedding to be possible. Since *that*-clause embedding takes place at the level of surface structure (SS), both clauses in (10) must have reached this level. At this point however, it is too late to perform raising, since this operation takes place at an earlier level. Thus, while cases of superraising can in principle be derived in bottom-up theories of embedding and must be ruled out by independent mechanisms, in the RT-model examples corresponding to (10) cannot be derived in

⁹Williams's (2003) proposal is a further development and elaboration of the model outlined in Riemsdijk and Williams (1981). Riemsdijk and Williams (1981) argue for the existence of a separate level of representation called NP-structure, in addition to the D-structure and S-structure that were already part of the GB model of grammar. They assume that all NP-movements (and some other operations) first apply to derive NP-structure. This structure then serves as input to the transformational operation of *wh*-movement that derives S-structure.

the first place. The ban on improper movement thus falls out from the very architecture of the model.

As Williams himself notes, RT crucially needs some notion of Extension to prevent trivial defeat of the most interesting predictions of the LEC. He thus assumes that Extension, essentially as it is used in Chomsky (1995), must also be part of the RT-model: any operation has to affect material that could not have been affected in a previous level. Without Extension there is no good reason why movement in SS would have to be to the periphery of the structure available at that level (i.e. the CP-structure), rather than to a position already made available at an earlier level (for example the Spec of IP). In general, Extension ensures that the periphery be affected by an operation.

As Hornstein and Nevins (2005) observe, an additional assumption is required in order to ensure ungrammaticality of improper movement configurations. Suppose a clause is embedded at an early level (somewhere before SS). Raising, ECM, and other operations/relations that depend on the absence of CP structure could take place at this point. At one of the subsequent levels, the embedded clause “grows” a CP through correspondence with a surface structure representation. The SpecCP could then be the target of movement. The question is what blocks this kind of derivation, which would rule in cases like (10) along with other improper movement violations and in general undermine the restrictiveness of the model. Hornstein and Nevins (2005) note that the answer can be found in discussion of an unrelated phenomena in one of the later chapters where Williams suggests that correspondence across levels must respect categories, i.e. “the only ‘growth’ that is allowed is growth that preserves the category, essentially adjunction” (Williams (2003), p. 185).

It is easy to see that Williams’s theory is restrictive enough to capture all the cases we have argued should fall under the rubric of improper movement, including the interactions between movement operations in non-identity contexts. As an illustration, consider the illicit feeding relation between A-movement and *wh*-movement in a remnant movement configuration. (11b) gives the derivation of (11a) in the standard model.

- (11) a. *A picture of t_i is known which king to have been sold. (Abels (2008b))
 b. $[_{IP} [A \text{ picture of } t_i]_j \text{ is known } [_{CP} [\text{ which king }]_i [_{IP} \text{ to have been sold } t_j]]]$

Recall that on Williams’s view SS is the level at which CP structure is introduced. It is also the level at which *wh*-movement takes place. The problem with (11) is that the matrix and the embedded clause are at different stages of development, a state of affairs prohibited by the LEC. The embedded clause is fully developed to the level of CP, which means that the matrix clause that embeds it must also be a CP. Again, at this point, a phrase cannot be raised from the embedded to the matrix clause since raising happens at a level prior to the level of SS. The ungrammaticality of (11) thus falls out in exactly the same way as in the identity cases. Since movement operations are assigned to particular levels, and the levels themselves are ordered via the LEC,

asymmetric ordering between different movement operations is also derived.

Note that in principle any type of grammatical relation, not just movement operations, can be assigned to a particular level within the RT-model. This allows Williams to link various properties of syntactic relations. For instance, both the locality and binding-theoretic properties of movement operations can be attributed to the same factors (cf. the GB-approach reviewed in section 3.1). As an illustration, consider the interaction between scrambling and anaphoric binding in German. As already noted in the previous chapter, scrambling to the pre-subject position does not give rise to new binding possibilities, i.e. an anaphor in need of a local binder cannot be bound by the scrambled phrase, as (12a) illustrates. In addition, this type of scrambling permits reconstruction. In (12b), the anaphor cannot be bound in its surface position since in this configuration it is not c-commanded by the antecedent. The grammaticality of (12b) thus shows that for the purposes of the binding theory, the scrambled phrase can be interpreted in the base position. On the other hand, scrambling to a position below the subject displays exactly the opposite properties: (i) it creates new binding possibilities (see (13a)), and (ii) does not permit reconstruction (see (13b)).¹⁰ For the sake of convenience, let me follow standard practice and refer to these two operations as IP-scrambling and VP-scrambling, respectively.¹¹

- (12) a. *weil den Studenten_i [die Lehrer von sich_i] zweifellos t_i in guter
since the student.ACC the teachers of self undoubtedly in good
Erinnerung behalten haben.
memory kept have
intended: ‘The teachers of himself have undoubtedly kept the student
in good memory.’
- b. dass sich_i der Fritz t_i schlau vorkommt.
that self the.NOM Fritz intelligent appears

¹⁰Examples (12b) and (13a) are from Müller (1995). (12a) is from Grewendorf and Sabel (1999), already discussed in the previous chapter. (13b) is from Putnam (2007), citing Moltmann (1991). The binding possibilities in German scrambling constructions are further discussed in chapter 7, section 7.7.

¹¹Instead of (12a), Williams provides the following example taken from Müller (1995):

- (i) *dass [den Frank]_i sich_i manchmal t_i nicht gemocht hat.
that the.ACC Frank self sometimes not liked has

(i) features a nominative anaphor. It is a well-known fact that many languages ban anaphors in nominative positions. Various analyses of this restriction have been offered, some of which do not attribute it to the binding theory (see Everaert (1991b) for an overview). Therefore, it might be argued that the ungrammaticality of (i) has nothing to do with the possibility of a scrambled phrase acting as an antecedent. Example (12a) avoids this interfering factor.

Example (13b) is also not the one that Williams provides to illustrate the relevant property. The problem with the example that Williams uses is that it does not clearly show that displacement has occurred since the two objects appear in the uninverted order.

- (13) a. dass der Arzt_j den Patienten_i sich_{i/j} t_i im Spiegel
 that the.NOM doctor the.ACC patient self in.the mirror
 zeigte.
 showed
- b. *Ich habe [die Bilder von einander_j]_i den Männern_j t_i
 I have the.ACC pictures of each.other the.DAT men
 gezeigt.
 shown
 ‘I have shown the pictures of the men to each other.’

In Williams’s system, these observations are explained in the following way.¹² Suppose that nominative subjects are defined at the level of predicate structure (PS). Scrambling beneath the surface subject position (identified with nominative case here) must therefore occur before PS, while scrambling to a pre-subject position must occur after PS. Assume further that reflexive binding also applies at PS. It then follows that only scrambling which applies before the PS (i.e. VP-scrambling) can affect binding relations, scrambling at subsequent levels cannot. Thus, as with locality, reconstruction relations which a particular syntactic operation enters into are determined entirely by where in the model the relevant operation occurs: an operation will “reconstruct for” any relation defined at previous levels, and for no relation defined at the same or at later levels.¹³

In the standard theory, the possibility of creating new binding relations and the lack of reconstruction, as well as tighter locality restrictions, are seen as properties of A-movements, while A’-movement is argued to display the opposite behaviour.¹⁴ On Williams’s view, the A/A’-distinction is relativized: positions at level R_i are A positions with respect to positions at level R_{i+1}.¹⁵

As we have seen, Williams’s system is restrictive enough to capture all ordering restrictions on movement operations noted in the previous chapter. We might ask

¹²In Williams’s model, scrambling is not considered to be an instance of real movement, but rather involves mismapping, i.e. displacement that arises from the mismatch of two levels. Though there are significant differences between mismapping and movement operations, these are not relevant for the point I am illustrating here.

¹³The term *reconstruction* is used here in a broader sense than usually assumed. As Williams notes, theta-relations could be thought to be assigned “under reconstruction”, in which case even VP-scrambling would show reconstruction effects of a certain kind. Examples like (11) can also be viewed in terms of reconstruction. In order to have remnant movement, the remnant moving rule must “reconstruct for” the remnant creating rule. In the RT model, this will happen only when the remnant creating rule is strictly earlier than the remnant moving rule. NP-movement occurs at case structure (or possibly predicate structure), while *wh*-movement occurs at a later level. Thus, NP-movement does not reconstruct for *wh*-movement, and cases like (11) are correctly predicted to be ungrammatical.

¹⁴The generalization that A-movement does not reconstruct is highly controversial. See the following chapter for further discussion of this issue.

¹⁵For further discussion of anaphoric binding in the context of the RT-model, see chapter 7.

however whether it is too restrictive, in that it also rules out certain well-formed constructions. In fact, Williams (2003) himself points out a couple of potential counter-examples. One case concerns quantifier movement in French, illustrated below.

- (14) a. elle rien_i voulu [que j' achète t_i]
 she nothing wanted that I buy
 'There is nothing that she wanted me to buy.' (Starke (2001))
- b. Il a tous_i fallu [qu' ils t_i parlent]
 it has all needed that they speak
 'It was necessary that they all speak.' (Williams (2003))

The quantifier moves from the embedded clause, crossing the complementizer *que*, but it clearly does not land in the matrix CP domain as it follows the subject and in (14b) also the finite auxiliary. This is precisely the type of derivation that is ruled out by the LEC, in tandem with Extension. If the embedded clause is a CP, then so is the matrix, and Extension would prohibit any movement that does not target the edge of that CP. Observing that the clause out of which the quantifier moves is necessarily subjunctive (or infinitival), Williams (2003) speculates that the answer might lie in treating the subjunctive clauses as smaller than full CPs, despite the presence of *que*. Note however that the embedded clause would have to be considerably smaller than CP (possibly smaller than TP), so that movement (which is always to a peripheral position) could target a position below both the subject and the auxiliary verb. As Abels (2008b) notes, there are more examples of this type. In general, the problem seems to be that LEC rules out any displacement operation which targets a position lower on the functional hierarchy (or, on Williams's view, earlier in terms of the level where it is defined) than one or more of the heads crossed on the path of movement. In some, but arguably not all cases, this problem can be circumvented by reanalyzing the relevant operation as involving mismapping between levels, rather than actual movement. This is suggested by Williams for ECM constructions.

Setting these concerns aside, Williams's account is the only proposal discussed so far that captures all the restrictions that we have argued should fall under the rubric of improper movement. An additional advantage of this approach is that no constraints or filters specific to the phenomenon in question are required.¹⁶ Once operations are assigned to particular levels, the restrictions on their ordering fall out from the very architecture of the system. In chapter 5, I will show that there is no need to adopt Williams's model of syntactic theory in order to achieve these results. Even within the confines of the standard model, the generalized version of the ban on improper movement can be derived from the basic principles governing derivation. In chapters 7 and 8, I will then go on to show that, as is the case in Williams's system, the proposal to be put forth also allows us to establish a link between movement and binding, by

¹⁶Arguably, Extension and the restriction that correspondence under Shape Conservation must respect categories, which as we have seen are crucial in deriving the ordering restrictions, would be needed independently, as they have consequences beyond the improper movement phenomena.

arguing that the same assumptions about the nature of the syntactic derivation not only play a role in fixing the ordering of movement operations, but also determine the binding possibilities.

3.4 Obata and Epstein (2008)

An attempt to account for improper movement within the Minimalist framework has been made by Obata and Epstein (2008). The analysis they put forth relies on the phase-based model outlined in Chomsky (2008), which differs in some crucial respects from Chomsky's earlier Derivation-by-Phase model. Let us therefore first briefly review the key features of the system Chomsky (2008) proposes.

Chomsky (2008) argues that both Edge- and Agree-features are borne by phase heads. This means that T has no inherent ϕ -features, but *inherits* these from C, C being a phase head. One consequence of this view is that movement to SpecCP and SpecTP takes place simultaneously and independently, which, according to Chomsky, allows one to capture certain asymmetries in extraction from subjects. Consider the following set of examples.

- (15) a. Of which car did they find the driver?
 b. Of which car was the driver awarded a prize?
 c. *Of which car did the driver cause a scandal?

Example (15b) shows that extraction from a derived subject (internal argument) of a passive verb is as acceptable as extraction from an object, (15a). On the other hand, extraction from an underlying subject is prohibited, as illustrated in (15c).¹⁷ In both (15b) and (15c), the surface position of the subject DP is the same, namely SpecTP. If extraction applies in SpecTP, then these two cases cannot be distinguished. This suggests that the availability of extraction is determined by the base structure, and therefore that extraction should take place from the base position of the subject. Chomsky argues that this can be achieved by adopting the feature-inheritance system and thus forcing *wh*-extraction and A-movement to happen in parallel. Consider first (15b). All the action takes place once the C head is merged. T inherits ϕ -features from C and starts acting as a probe. *Wh*-extraction and A-movement of the dominating DP then proceed in parallel from the base position of the subject.¹⁸ No intermediate stop-over at the edge of *v* is needed, given that on Chomsky's assumptions, unaccusative/passive *v* is not a phasal head. On the other hand, transitive

¹⁷These empirical claims have not gone unchallenged. Starke (2001) for instance argues that extraction from the subject position is in principle possible (cf. (15) to (37a) from the previous chapter), and Broekhuis (2005) argues that extraction is equally (im)possible from both derived and underlying subjects.

¹⁸If all operations apply at the phase level, then intra-phasal violations of cyclicity constraints, such as the Extension condition, must be allowed. In the current system, the only cyclicity violation involves going back to previously transferred phases.

v^* is a phasal head, therefore extraction in (15a) must proceed via the v^*P edge (as dictated by the Phase Impenetrability Condition). From Spec of v^*P , the *wh*-phrase can then be attracted by the next higher phase head C. According to Chomsky, the PP complement in the illicit (15c) cannot be extracted in the same way, because it is not in the search domain of the probe v^* , the search domain of the probe being its c-command domain. The PP thus remains embedded within the external argument and can no longer be attracted when C gets merged. It is however not entirely clear why the PP should be inaccessible, given that it is contained within a phrase that occupies the phase edge. According to Chomsky, the external argument can be attracted from the Spec of v^*P (since it can be probed by T), but nothing embedded inside it can, though he doesn't make it explicit what this restriction should stem from. He only notes that in cases such as these, the PP is inaccessible to the higher probe in C since, at the point when C is merged, the PP would be embedded in a phase "which has already been passed in the derivation." (Chomsky (2008), p. 147)

Note that although there is no movement from SpecTP to SpecCP in (15), Chomsky (2008) does not categorically rule out the possibility of A-movement feeding A'-operations. In fact, his analysis of ECM and raising constructions relies on such a feeding relation. The derivation of (16), for instance, would go as follows. Assuming that A-movement proceeds successive-cyclically through every intermediate TP, the Agree feature of C-T would raise the subject step-by-step to the matrix TP, while the edge feature of C would extract the PP from the intermediate position, SpecTP. We know that extraction cannot take place from the base position of the subject, due to the ungrammaticality of examples like (15c). Thus it must take place from SpecTP.¹⁹

(16) Of which car is the driver likely to cause a scandal?

Let us now see how Obata and Epstein (2008) employ this system to derive the ban on improper movement. In addition to feature-inheritance which they adopt from Chomsky's work, O&E introduce another mechanism they refer to as *feature-splitting* to account for the relevant facts. Consider an example of superraising:

(17) *Who seems will leave?

a. $[_{CP} \langle \text{who}_{[Q]} \rangle C_{[EF]} [_{TP} \langle \text{who}_{[Q][Case]} \rangle T_{[u\phi]} [_{vP} \langle \text{who}_{[Q]} [\phi][Case] \rangle \dots]]]$

b. $[_{CP} C_{[EF]} [_{TP} T_{[u\phi]} \text{seems} [_{CP} \langle \text{who}_3[Q] \rangle [_{TP} \dots]]]]$

¹⁹We might now wonder why this kind of derivation, involving extraction from TP, is impossible in (15c). According to Chomsky, the reason is "the generalized inactivity condition". He argues that the head of the A-chain is invisible to Agree. In other words, A-chain becomes invisible to further computation when its uninterpretable features are valued. It is thus important that the SpecTP out of which extraction takes place in ECM and raising constructions is not a position where all the uninterpretable features are valued.

(17a) illustrates the point in the derivation when the embedded C is merged. The edge feature on C and ϕ -features on T (inherited from C) each independently attract the single element *who* occupying the edge of vP. As a consequence of this simultaneous attraction of the single element by two distinct heads, O&E propose that *feature-splitting* takes place. Namely, the features on *who* get separated: [*uCase/i* ϕ] move to SpecTP, while [Q] moves to the edge of CP, as illustrated in (17a). Then, the embedded TP is spelled out and the derivation goes on to the matrix clause. In (17b), [*u* ϕ] on the matrix T seeks a matching goal with [*i* ϕ]. However, *who* at the edge of CP bears only [Q], not [ϕ], by virtue of feature-splitting and thus it is not an appropriate goal. On the other hand, *who* in the embedded SpecTP is inaccessible due to the Phase Impenetrability Condition. As a result, [*u* ϕ] on matrix T is never valued, which causes the derivation to crash.

Regarding examples like (18), O&E only note that the ungrammaticality of (18) can be explained by saying that [*uCase*] on *who* is transferred unvalued along with the lowest TP, which causes the derivation to crash.

(18) *Who seems it is likely to leave?

Presumably, the Q-feature of *who* can move to the edge of CP, while the insertion of the expletive would preclude the movement of [*uCase/i* ϕ]. Clearly, feature-splitting must take place in this case as well, even though [*uCase/i* ϕ] are not attracted. If movement of [Q] could pied-pipe these features, then (18) would not be ruled out. The details regarding when precisely feature-splitting can and must take place remain to be worked out.

Many issues regarding the mechanism of feature-inheritance also remain unclear. This becomes particularly obvious if one tries to extend the proposed analysis to other cases we argued should be considered as instances of improper movement. For example, given that the system accounts for clause-boundedness of A-movement as in (17) and (18), it would be desirable and more parsimonious to derive clause-boundedness of other movement operations by employing the same mechanisms, rather than assuming additional ones. Thus, one would hope that the mechanisms proposed would suffice to capture the clause-boundedness of middle field scrambling, in languages like German. Following the logic above, the C head would presumably come equipped with a scrambling feature, which would be inherited by a clausal head in the middle field. A number of non-trivial questions regarding the feature-inheritance system arise at this point. Are all features of the middle field present on the C head? If so, which features are inherited and which not? Also, can features be inherited by multiple heads in the IP domain? For German, we would have to assume that the scrambling feature, if also inherited, can be spread to multiple heads since scrambling can be iterated, as illustrated in (19) (example from Müller (1995)). The same however could not be true for [*uCase/i* ϕ], as we might then expect German to allow more than one element within the same clause to surface bearing nominative

case.²⁰

- (19) daß Ellen_i [die Gerüchte über Ina]_j keiner t_i t_j geglaubt hat
that Ellen.DAT the rumours.ACC about Ina noone.NOM believed has

There are also concerns of a more general nature. Minimalism dictates that language phenomena should be accounted for by using as little technical machinery as possible and only that which is ‘virtually conceptually necessary’. From this perspective, feature-inheritance and feature-splitting look like extra technical devices, marking a departure from optimal design, unless it can be shown that they are necessary or provide a more parsimonious way of accounting for particular properties of language. In chapter 5, I will show that improper movement effects can be captured within a phase-based theory of locality without appealing to these additional mechanisms.

3.5 Abels (2008b)

Abels (2008b) captures improper movement effects by positing an extrinsic ordering on the application of movement operations. As already noted in chapter 2, Abels argues that there is a universal constraint on the order in which operations apply to a given constituent, which he calls the Universal Constraint on Operational Ordering in Language (UCOOL).²¹ A proposed subpart of UCOOL is given in (20).

- (20) $\theta \prec A\text{-mvt} \prec (\text{clause internal}) \text{ scrambling} \prec wh\text{-mvt} \prec \text{topicalization}$

UCOOL requires θ -related operations to apply first, followed by case-related operations, followed by scrambling and so on. In the previous chapter, I have discussed how strong the empirical evidence supporting this particular arrangement of operations is. Let us now see how UCOOL is implemented.

Abels proposes the following principle in order to regulate how UCOOL is deployed:

- (21) *Generalized Prohibition against Improper Movement (GenPIM)*
No constituent may undergo movement of the type σ if it has been affected by movement of the type π , where $\sigma < \pi$ under UCOOL.

UCOOL is relevant not only for cases involving consecutive movements of the same phrase, but also for cases involving subextraction and remnant movement, as

²⁰We might also ask why T is needed at all in a system where C carries all the features; or, more generally, why do nonphase heads exist at all, if phase heads do all the work? See Richards (2007) for discussion of this issue and a possible solution.

²¹Recall that similar hierarchies of syntactic operations have also been proposed by Sternefeld (1993) and Grewendorf (2003). Since these authors do not discuss how the hierarchies in question should be implemented, I will focus here on Abels’s proposal.

Abels argues.²² To ensure that the restrictions stated in UCOOL extend to these latter configurations as well, Abels introduces the notion of *affectedness* in the definition above. *Affectedness* is defined in the following way:

- (22) A constituent α is affected by a movement operation iff
- i. α is reflexively contained in the constituent created by movement, and
 - ii. α is in a (reflexive) domination relation with the moved constituent.

(22) ensures that when a constituent α is moved, this movement affects not only the moving constituent itself (it is in a reflexive domination relation with itself), but also constituents that make up α (the elements that α dominates), and the nodes along the path of movement, since those dominate α in the pre-movement configuration. Note that specifiers and heads along the path of movement are unaffected. Assuming that the hierarchy stated in (21) is valid, UCOOL, in conjunction with the GenPIM, captures the ordering restrictions between different movement operations for both identity and non-identity cases.

A host of questions spring to mind, relating both to the hierarchy itself and to the GenPIM.²³ A particularly pertinent one is how operational types relevant to the UCOOL should be defined. Though this seems to be a highly significant issue, Abels does not discuss it in any detail.²⁴ He merely suggests, as a first approximation, that features that enter into the calculation of relativized minimality/attract closest define operational types. As regards the GenPIM, though the notion of *affectedness* that Abels (2008b) defines seems to be descriptively adequate in that it encompasses the relevant improper movement configurations, one might wonder why exactly this notion/definition should hold and not some other plausible formulation.²⁵

The arbitrary nature of the notion of *affectedness* has also been pointed out by Neeleman and van de Koot (2010) (henceforth N&K). They propose an alternative way of implementing UCOOL which explains why the nodes affected by movement are precisely those stated in Abels's definition of *affectedness*. The main goal of their paper is to argue for a new conception of syntactic dependencies, which, according to the authors, provides a way of capturing patterns of interaction between different movement types in both identity and non-identity cases, in addition to having some other beneficial consequences that I will not discuss in this summary. They start by investigating the implications of the Inclusiveness Condition, which leads them to a

²²For relevant empirical evidence, see chapter 2, section 2.2.

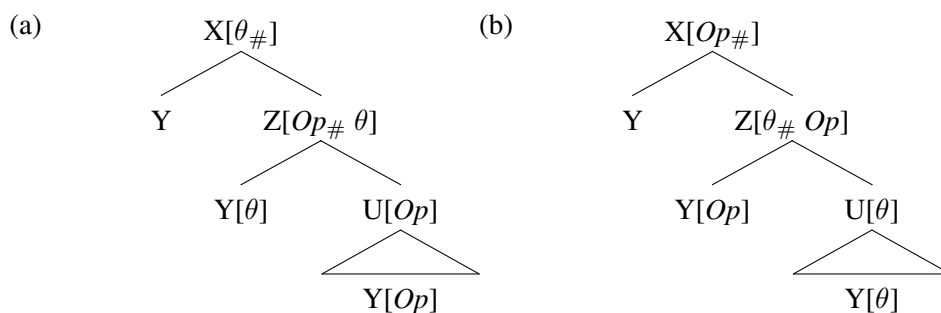
²³For issues related to the empirical evidence for the proposed hierarchy, I refer the reader to the previous chapter.

²⁴It seems to me that the question of how operations relevant for the hierarchy are to be defined must be clarified if the theory is to be falsifiable. In order to establish whether there are potential counterexamples to the proposed hierarchy, we must be able to identify instances of the same kind of operation in other languages. See chapter 2, section 2.2 for relevant discussion.

²⁵See Abels (2008b) for a suggestion as to how the rather complex definition of affectedness might be simplified.

revision of movement theory. In particular, movement is on their view mediated by a selectional requirement (SR), which must be introduced by a trace, copied up the tree, and satisfied by the trace's antecedent. On their view, traces are lexical items. Traces of A'-movement are lexical items that bear the SR *Op* (for operator); traces of A-movement bear the SR θ . The syntactic properties of a terminal node (i.e. features and selectional requirements) are freely generated, on the condition that they are licensed either by matching the terminal to a lexical entry (what they refer to as *external matching*), or by matching it to another node in the same representation (*internal matching*). In case of A'-movement, the SR *Op* is licensed by an external matching relation with a particular lexical entry, i.e. that of A'-trace. *Op* is then copied up the tree and satisfied by the *wh*-feature of the antecedent. The function of *Op* is to identify the source for internal matching of the remaining properties of the terminal. In N&K's view, it is this internal matching that is responsible for syntactic reconstruction effects. Syntactic reconstruction is obligatory with A'-movement, they assume, and to ensure its obligatoriness they stipulate the Copy Condition, which bans the presence of a copy of an attribute (a feature or a SR) in the node containing a SR that licenses the reconstruction. On their assumptions, θ never licenses internal matching and therefore syntactic reconstruction should be unavailable in A-chains. It is this difference in the reconstructive behaviour of θ and *Op* that ensures the correct ordering of A- and A'-movements. The derivation in (23a) violates the Copy Condition, and is therefore ruled out. *Op* is copied to U from the trace that introduces it, and then onward to Z where it is satisfied (indicated by #). Since Z contains *Op* and *Op* is a SR which licenses the reconstruction of θ , the Copy Condition dictates that θ cannot be copied onto Z. Therefore, A'-movement cannot be followed by A-movement. The reverse ordering of these movement operations, illustrated in (23b) is licit, because θ does not give rise to syntactic reconstruction of *Op*. Non-identity cases are captured in the same manner. I refer the reader to Neeleman and van de Koot (2010) for illustration and discussion.²⁶

(23)



²⁶Compare this account to that of Riemsdijk and Williams (1981) and Williams (2003). Like these proposals, Neeleman and van de Koot (2010) link the reconstructive properties of A- and A'-movement to the patterns of interaction between these movements.

The account can be generalized beyond the A/A'-movement typology, N&K argue, if Abels's UCOOL is understood as a hierarchy of selectional requirements. They dub this variant UCOOLR, for the Universal Constraint on Operational Ordering in Language Reinterpreted:

- (24) UCOOLR
 $\dots \theta \prec \text{Scr} \prec \text{Op} \prec \text{Top} \prec \dots$

A principle regulating when syntactic reconstruction is possible replaces Abels's GenPIM:

- (25) *Principle of Syntactic Reconstruction (PSR)*
 Satisfaction of a selectional requirement SR1 gives rise to syntactic reconstruction of SR2 if and only if $\text{SR2} \ll \text{SR1}$ under UCOOLR.

Although N&K's proposal approaches the problem from a different angle, in certain crucial respects it highly resembles Abels's account. As the authors themselves note, the predictions of their analysis and Abels's account are empirically indistinguishable. Furthermore, like Abels (2008b), they also rely on an extrinsic ordering hierarchy (in their case the hierarchy of selectional requirements), in addition to a constraint regulating how the effects of the hierarchy come about, both of which can be dispensed with, as I will argue in the following chapters.

Before concluding the review of this type of proposal, consider another notable aspect of the proposed hierarchy of operations. It is immediately obvious that UCOOL bears a clear resemblance to the functional hierarchy of the clause. Within the clausal structure, theta-domain is below the case-related projections in the IP domain, and in turn the IP domain is below operator positions in the CP domain. Given this clear resemblance, having two hierarchies seems redundant and one should ask whether one of these hierarchies could be derived from the other. Pursuing this line of inquiry, the following chapters will argue that the hierarchy of operations is not needed and that improper movement effects can be derived by relating movement to the functional hierarchy in a particular way.

3.6 Summary

In this chapter, I have reviewed some of the previous accounts of the ban on improper movement, and discussed their benefits and shortcomings. Along the way, I have already hinted at the direction that the alternative account to be proposed in the following chapters will take. Namely, I will argue that the observed ordering restrictions between movement operations in both identity and non-identity cases can be derived within the confines of the Minimalist approach to syntactic theory and in strictly derivational terms, without appealing to representational filters. In addition, it will be shown that this can be achieved by employing only those principles/properties

of the grammar that are independently needed, namely the hierarchy of functional projections and a phase-based theory of locality.

Chapter 4

Relating improper movement to the functional sequence

It has been established in chapter 2 that there are restrictions regarding possible sequencing of movement operations that go beyond those imposed by the standard formulation of the ban on improper movement. Assuming a richer inventory of movement operations than the traditional A/A'-typology, there appears to be evidence that movement operations apply in a particular order. As we have seen, a number of authors have argued that the relevant ordering can be stated in terms of a hierarchy, an example of which is repeated below (from Abels (2008b)).

- (1) *Hierarchy of Operations*
 $\theta \prec \text{A-mvt} \prec (\text{mittelfeld}) \text{ scrambling} \prec wh\text{-mvt} \prec \text{topicalization}$

An obvious question to ask is whether these restrictions on sequencing of operations follow from deeper principles of grammar, or the hierarchy needs to be stated as such. It is easy to notice that the hierarchy given above highly resembles the hierarchy of clausal functional projections. In the clause, theta-domain is below case-related projections in the IP domain, and in turn the IP domain is below operator positions in the CP domain. Since the hierarchy of functional projections (or the functional sequence) is needed independently of improper movement effects, positing a hierarchy of operations seems redundant. Thus, a desirable goal would be to try to capture the restrictions on ordering of movement operations by relating movement/operations to functional structure. In fact, within a single clausal domain, the sequencing of operations follows from the functional hierarchy, coupled with the standard assumption that movement is always to a c-commanding position. Consider as an illustration the interaction between *wh*-movement and A-movement.

- (2) $[_{CP} \text{Who}_i [_{TP} t_i [_{VP} t_i \text{likes cheese}]]]$?

The landing site of *wh*-movement (the SpecCP) is higher in terms of functional hierarchy than the landing site of A-movement, SpecIP. Since downward movement is

prohibited, it follows that A-movement can feed *wh*-movement, but not the other way round.

However, as soon as we turn to examples featuring cross-clausal movement, this simple account breaks down. Consider (3):

- (3) * $[_{CP} \text{Who}_i [_{TP} t_i [_{VP} \text{seems} [_{CP} t_i \text{that} [_{TP} t_i [_{VP} t_i \text{likes cheese}]]]]]]]$?

Example (3) does not violate the functional hierarchy, and movement consistently targets a c-commanding position, yet the example is ungrammatical. The problem is that relatively “big” structures can be embedded under “small” structures, so that (4) is a stage in the derivation of (3). At this point, nothing said so far prohibits movement from embedded SpecCP to matrix TP.

- (4) * $[_{TP} \text{Who}_i [_{VP} \text{seems} [_{CP} t_i \text{that} [_{TP} t_i [_{VP} t_i \text{likes cheese}]]]]]$?

The challenge therefore lies in cases that involve embedding. Something more is needed to regulate possible feeding relations between movement operations across clausal domains. As a way of approaching this challenge, in this chapter I will explore one possible way of relating the ban on improper movement to the structural hierarchy in the clause, namely by stating the conditions on movement over possible landing sites.

The idea has its origins in the work of Sternefeld (1993), and Williams (2003). Although Williams (2003) proposes a novel syntactic theory, he suggests that in the standard model, the Generalized Ban on Improper Movement (i.e. the one that relies on a richer inventory of movement operations) might be stated in the following way:

- (5) *Generalized Ban on Improper Movement* (Williams (2003), p. 72)
Given a Pollock/Cinque-style clausal structure $X_1 > \dots > X_n$ (where X_i takes $X_{i+1}P$ as its complement), a movement operation that spans a matrix and an embedded clause cannot move an element from X_j in the embedded clause to X_i in the matrix, where $i < j$.

Slightly modifying and simplifying Williams’s formulation, let me restate the Generalized Ban on Improper Movement in the following way:

- (6) *Generalized Ban on Improper Movement (GBOIM)*
A phrase cannot be moved to a position lower in terms of the functional hierarchy than the launching site of movement.

Note that no appeal to an independent hierarchy of operations is made on this view. Rather, the ordering restrictions are predicted to follow from the functional architecture of the clause itself. The rest of this chapter is devoted to exploring just how successful a theory that incorporates the GBOIM is in capturing improper movement effects. The first conclusion that will be reached is that the GBOIM alone is too weak, leaving many relevant cases unaccounted for. It will then be shown that the restrictive

power of GBOIM considerably increases once it is combined with a theory of locality. The proposal calls for a reanalysis of ECM constructions, which we will consider in some detail in section 4.3. Finally, it will be shown that even this modified version of the GBOIM-based theory falls short of being fully satisfactory, which will lead us to considerably revise the proposal in the chapter to come.

4.1 The GBOIM is too weak

A theory which relies only on the GBOIM is not restrictive enough to capture all the cases we argued should fall under the purview of improper movement. While it accounts for the observed feeding/bleeding relations in identity cases, it makes no predictions regarding examples that involve remnant movement and subextraction from moved phrases. Let us first look at a couple of cases that do fall out from the GBOIM.

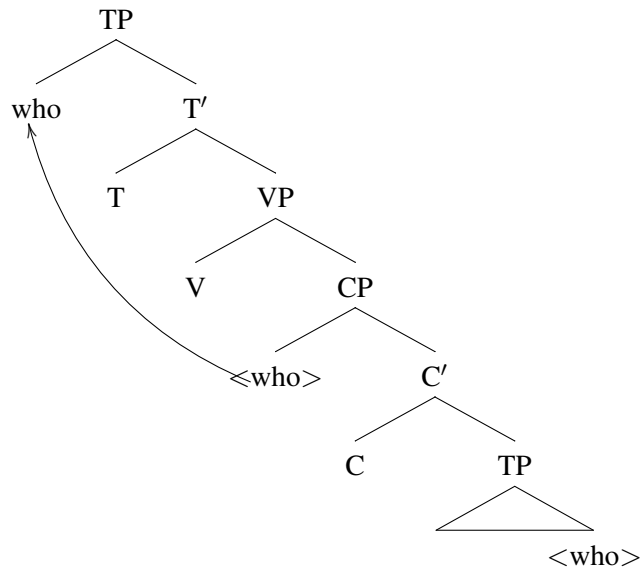
Consider the pair of examples in (7), illustrating the proper sequencing of *wh*-movement and A-movement (i.e. case movement). Following standard practice, I assume that *wh*-movement targets SpecCP, while subjects undergo A-movement to TP:¹

- (7) a. It is known who was invited.
 [_{TP} It is known [_{CP} who [_{TP} <who> was invited <who>]]]
 b. *Who is known (that) it was invited?
 [_{TP} Who is known [_{CP} <who> [_{TP} it was invited <who>]]]

In the clausal structure, CP dominates TP. Consequently, the GBOIM correctly rules out (7b). In this case, the *wh*-phrase moves from the Spec of CP to a projection which is lower on the functional sequence than CP, precisely the kind of configuration prohibited by GBOIM. This is schematized in (8). Considering that case positions are in general below the CP domain, the GBOIM predicts that a phrase which has undergone *wh*-movement to CP cannot move further for case purposes.

¹As the structures involving remnant movement and subextraction from moved phrases can quickly become extremely complex and cumbersome, for the sake of clarity, the clausal structure is maximally simplified in the representations to follow, and only projections and movements that are relevant for the discussion at hand are represented.

(8)



It is easy to see how this logic extends to other feeding/bleeding relations between movement operations in identity cases. Consider for instance the position which mittelfeld scrambling occupies in the hierarchy of operations, repeated below.

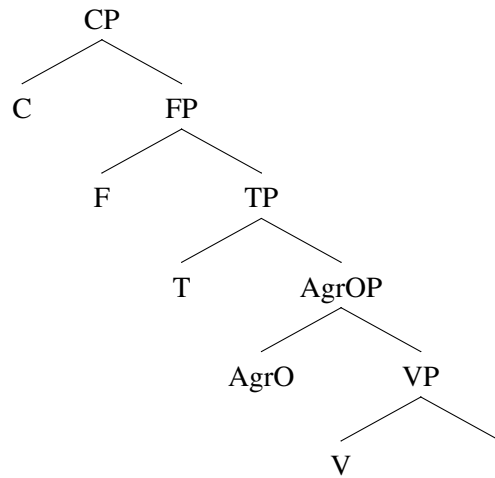
(9) *Hierarchy of Operations*

$\theta \prec \text{A-mvt} \prec (\text{mittelfeld}) \text{scrambling} \prec \text{wh-mvt} \prec \text{topicalization}$

The Hierarchy states that A-movement can feed scrambling, and scrambling can in principle feed *wh*-movement and topicalization. Recall also that what was meant by *scrambling* in this case was ‘movement to a pre-subject position in a language like German’. Let us assume therefore that scrambling in German targets the specifier of a projection on top of TP, which I will for the sake of convenience label as FP. Let us also assume that objects in German move to case positions, which I have marked here as AgrOP.²

²The predictions still hold even if object case is assigned VP internally, since VP is below FP in the functional hierarchy. Nothing crucial thus hinges on this, and I resort to AgrOP at this point just for the sake of explicitness.

(10)

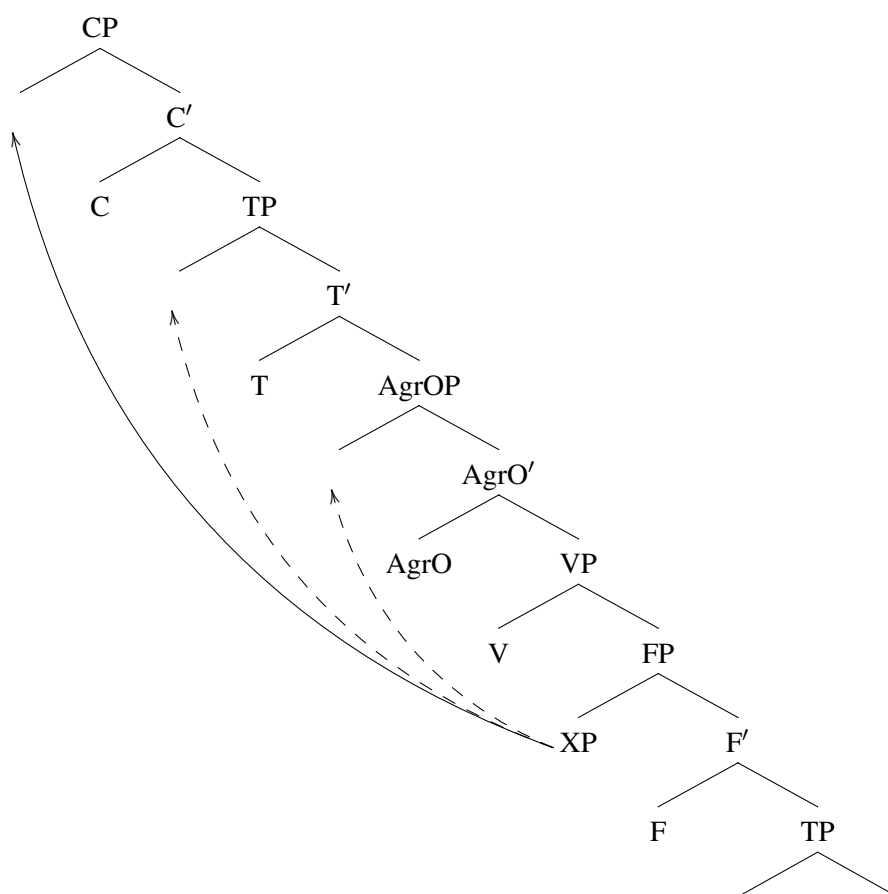


Given the functional sequence in (10) and the GBOIM, we predict that a phrase which has landed in FP can be further *wh*-moved to SpecCP, but cannot be moved for case purposes as all case positions are below FP in the functional hierarchy. Crucially, this prohibits scrambling from feeding A-movement even in cases of embedding, as sketched in (11).³ Movement steps marked with the dashed line are predicted to be illicit, while the one marked with the full line is allowed. It is therefore expected that mittelfeld scrambling should be sandwiched between A-movement and *wh*-movement in the Hierarchy of Operations.⁴

³I present here the case where the embedded clause is smaller than CP, to circumvent the restriction that A-movement in general cannot cross the finite CP boundary.

⁴German has another clause-internal scrambling operation which inverts the order of the direct and indirect object, and shows properties distinct from scrambling to a pre-subject position. Proponents of the Hierarchy of Operations do not discuss this type of scrambling or its placement in the Hierarchy. I will follow this practice for now and set this type of scrambling aside, but will return to it in the following chapter.

(11)

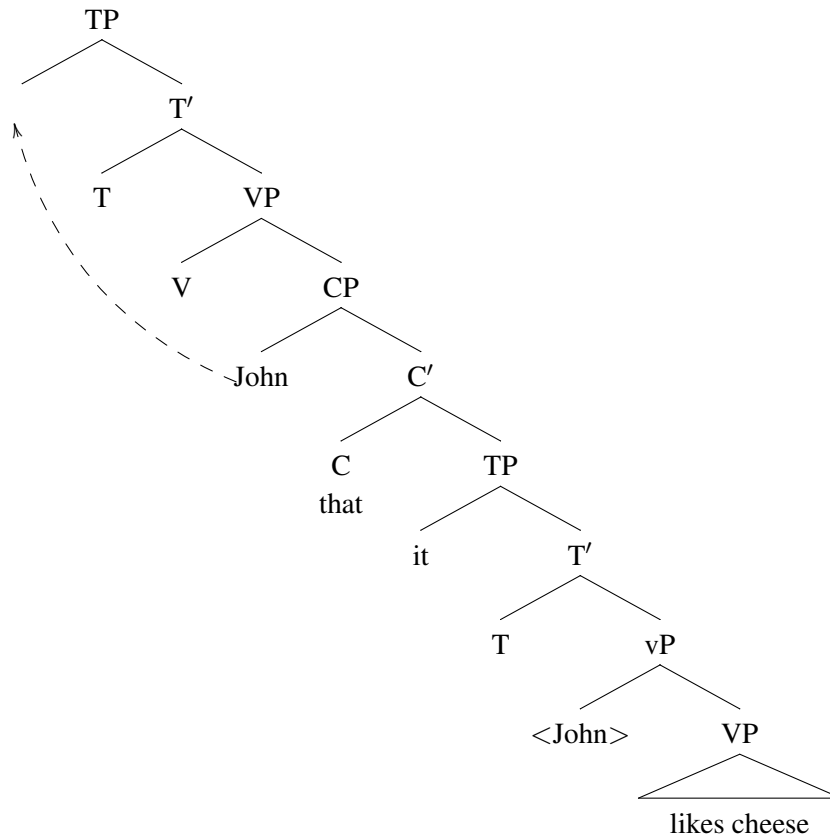


Therefore, assuming that the Hierarchy of Operations is identical to the hierarchy of functional projections (i.e. that θ -positions are dominated by case positions, which are in turn dominated by the scrambling position(s), etc.), the GBOIM seems to be delivering the desired results as far as identity cases are concerned.⁵

However, problems start to emerge once we look at cases of superraising. Consider (12). The GBOIM rules out (12), on the derivation illustrated below. Observe that the launching site of A-movement is SpecCP, which is higher in terms of the functional hierarchy than the target.

⁵Rizzi (1997) argues that there are multiple Top positions in the left periphery. If so, then on the current view, this leads to the conclusion that topicalization and *wh*-movement should not be asymmetrically ordered, contrasting with the predictions based on the Hierarchy of Operations. Note that this is not necessarily a problem since it seems that the proponents of the Hierarchy of Operations must distinguish different types of topicalization and/or *wh*-movement if they are to maintain the assumption that these two operations are asymmetrically ordered, as discussed in chapter 2, section 2.2.

(12) *John seems that it likes cheese.



However, note that the stop-over in the intermediate CP is crucial for GBOIM to have an effect. The derivation where *John* moves directly from the base position to the matrix TP is not blocked by the GBOIM. This type of derivation is however prohibited by locality, which ensures that movement proceeds successive cyclically. Thus, the GBOIM captures the ungrammaticality of superraising examples only in tandem with a theory of locality. This is in fact true of all the other accounts of improper movement discussed in the previous chapter,⁶ i.e. they all must resort to some version of locality to make sure that all possible derivations of (12) are ruled out.

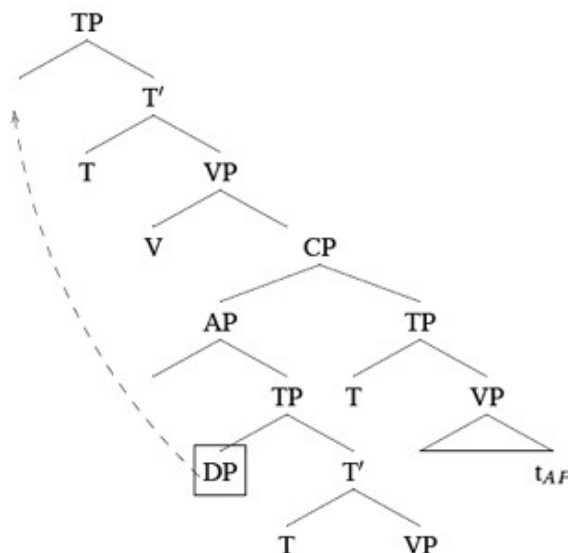
Once we turn to non-identity cases, it becomes even more evident that the GBOIM is too weak on its own to rule out illicit combinations of movement operations. To see this, let us look at the interaction between A-movement and *wh*-movement, starting with the configuration involving subextraction from a moved phrase, as in (13) (from Collins (2005)).

⁶With the possible exception of Williams (2003). In Williams's system, both locality and improper movement effects fall out from the LEC.

- (13) *Oscar was asked how likely to win it was.
 [_{TP} Oscar_i was asked [_{CP} [_{AP} how likely t_i to win]_j [_{TP} it was t_j]]]

Here, the AP undergoes *wh*-movement to the embedded CP. The derivation presumably goes wrong when we try to extract the DP *Oscar* from the moved phrase and A-move it to the matrix TP, as schematized in (14).

(14)



The crucial question to ask is what the launching site of A-movement is. The DP is embedded in a phrase that occupies the SpecCP. However, the DP itself occupies a lower position in the embedded infinitival clause: either the base position (SpecvP), or the Spec of TP if TPs always have an EPP feature. Clearly we want the position of the embedding phrase to count, but this does not seem to be ensured by the GBOIM.

The inadequacy of GBOIM is further revealed by examples involving remnant movement. Recall that *wh*-movement cannot feed remnant A-movement.

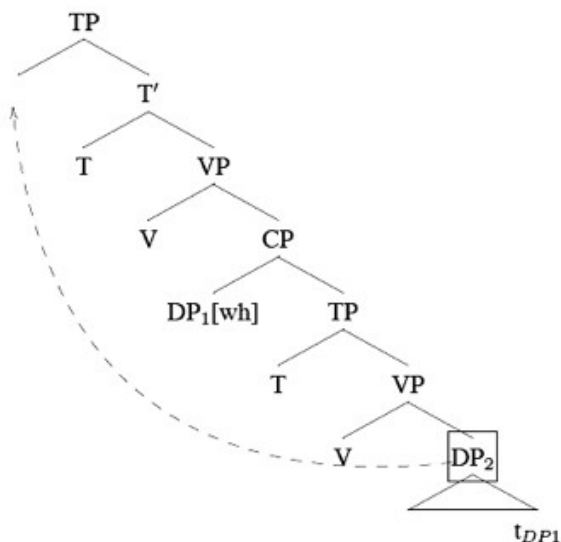
- (15) *[A picture of t_i]_j is known [which king]_i to have been sold t_j. (Abels (2007))

The remnant DP cannot undergo A-movement after a subpart of it has been *wh*-moved. It is even clearer in this case that the landing site targeted by the remnant DP (namely the matrix TP) is not lower in terms of the functional sequence than the launching site.⁷ Therefore, (15) does not fall under the purview of the GBOIM.⁸

⁷Again there might be a trace in the embedded TP if non-finite TP also has an EPP feature (see section 4.3 for discussion of this issue). This does not affect the argument developed above.

⁸One might object that (15) does not bear on the ordering of operations since we know independently that A-movement is clause-bound, i.e. it cannot cross the CP. However, recall that on usual assumptions, the clause-boundedness of A-movement is also described in terms

(16)



To summarize, we have seen that the GBOIM as defined in (6) does derive asymmetric feeding relations between movement operations in identity cases, without appealing to an independently stated hierarchy of operations. However, it does not capture the restrictions on sequencing of operations in non-identity cases, nor is it sufficient to rule out cases of superraising. The following sections discuss a way of strengthening the GBOIM which significantly increases its empirical scope.

4.2 Strengthening the GBOIM

4.2.1 Enter locality

Can the restrictive power of the GBOIM be increased in a non-stipulatory way? One possible venue to explore presents itself when looking at cases of superraising. Recall that ungrammaticality of superraising constructions, i.e. the clause-boundedness of A-movement, follows from the GBOIM only in tandem with a theory of locality. Locality forces the moved phrase to stop off in the intermediate SpecCP, as illustrated in (17). Since A-movement then targets a position in the matrix clause which is lower in terms of the functional sequence than CP, superraising examples are ruled out.

(17) *_{[TP John_i seems [_{CP} t_i that it t_i likes cheese]]}.

of improper movement (as in the cases of superraising). (15) also shows that it is not finiteness of the CP which matters in this respect. The embedded CP here is non-finite, but A-movement out of it is still blocked. The same point is illustrated by the following example, which is equally ungrammatical but does not feature remnant movement.

(i) *A picture of the king is known where to have been sold.

This account carries over straightforwardly to cases of scrambling. Recall that scrambling in German targets a position lower than CP, which we have labelled FP. If an obligatory stop-over in intermediate CP is forced by locality, then the GBOIM correctly predicts that scrambling in German cannot cross the CP boundary (example from Müller (1995)).

- (18) *daß [_{FP} Hygrometer_i [_{TP} keiner sagt [_{CP} t_i daß [_{TP} Antje t_i mag]
 that hygrometers noone says that Antje likes

This does not mean however that the domain of scrambling is confined to a single clause. Cross-clausal scrambling is possible as long as there is no CP intervening. Wurmbrand (2001) has convincingly shown that a certain class of infinitives in German (often referred to as *coherent* or *restructuring* infinitives) are smaller than CPs. Scrambling out of this type of infinitive is licit (example from Müller (1998)).

- (19) dass [das Buch]_i keiner [t_i zu lesen] versucht hat.
 that the book noone to read tried has

Since some version of locality is needed to fully account for cases of superraising, or more generally, to capture clause-boundedness of certain operations, it seems natural to ask whether locality might play a role in capturing ordering restrictions in non-identity cases as well.

That movement proceeds successive cyclically is in the current model of syntactic theory ensured by appealing to *phases*. In chapter 1, I have briefly reviewed Chomsky's (2000; 2001) phase-based conception of locality. Recall that the general idea behind phases is that once these domains have been built, the complements of the phase heads are sent to the interfaces, i.e. they are spelled out. The internal domain of the phase is thus no longer accessible to syntactic computation. This is stated in the Phase Impenetrability Condition, repeated below. In order for an element embedded in the phase domain to be visible to an outside probe, that element must move to the edge of the phase. The PIC therefore ensures that movement proceeds phase-edge by phase-edge, i.e. successive-cyclically.⁹

- (20) *Phase Impenetrability Condition* (Chomsky (2001))
 In a phase α with head H, the domain of H [=complement of H] is not accessible to operations outside α [=HP], but only H and its edge [H plus any/all of its specifiers].

Forcing movement to proceed successive-cyclically, which is ensured here by appealing to the PIC, eliminates the problem that superraising constructions posed for the GBOIM. A phrase moving out of a CP is then forced to stop off in the Spec of CP, C being a phase head, and any subsequent movement to a position below C would violate the GBOIM. Many other feeding/bleeding relations in identity cases can be captured by relying on the GBOIM, even without a theory of locality, as illustrated in

⁹For discussion and illustration, refer back to chapter 1.

section 4.1. What remains to be determined is whether the GBOIM, coupled with a phase-based theory of locality, suffices to account for ordering restrictions on movement operations in non-identity cases. I turn to this issue in the following section.

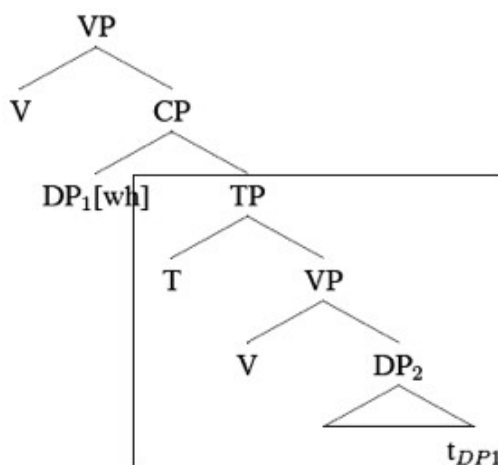
4.2.2 Non-identity cases again

Let us consider again an example involving illicit sequencing of *wh*-movement and A-movement in a remnant movement configuration:

(21) *[A picture of t_i] $_j$ is known [which king] $_i$ to have been sold t_j . (Abels (2007))

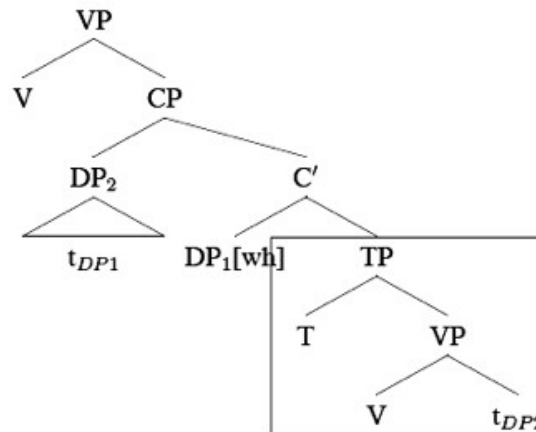
The ungrammaticality of (21) can now be ruled out in the following way. The crucial point in the derivation is the merge of the embedded CP. The *wh*-phrase then extracts out of the dominating DP and moves to SpecCP, where its [wh]-feature is checked. C, being a phase head, then triggers Spell-Out of its complement TP, as illustrated below (the spelled-out domain is boxed). Consequently, the remnant DP becomes inaccessible to further syntactic computation, i.e. it is not visible and thus cannot be attracted by the matrix T. The derivation crashes because the features on the remnant DP and/or matrix T remain unchecked.

(22)



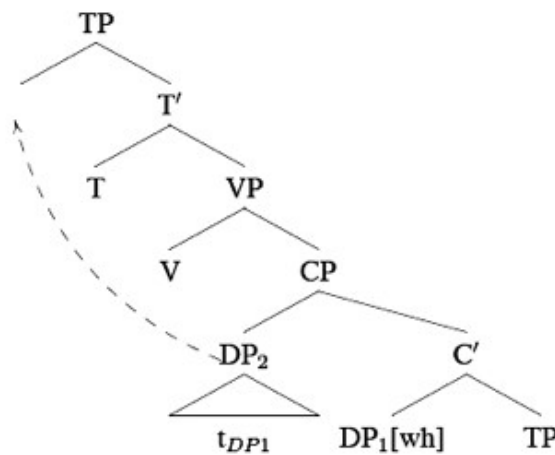
Recall however that there is a possibility of bypassing the PIC by moving to the phase edge. The remnant DP could then in principle move to an (outside) specifier of the CP and be accessible to probes in the matrix clause, as illustrated below.

(23)



The remnant DP sitting in the SpecCP would then be available for attraction by the matrix T. However, notice that in such a case the remnant DP would be moving to a projection lower in terms of the functional hierarchy than the launching site, precisely the kind of scenario ruled out by the GBOIM.

(24)



Both plausible derivations of (21) are thus successfully excluded. If the remnant CP does not move to the embedded CP, it will be caught in the Spell-Out domain and will hence be inaccessible to probes in the matrix clause. If on the other hand, the remnant DP moves to SpecCP, the subsequent movement steps can only target the matrix CP, but not any projections lower than CP, due to the GBOIM.

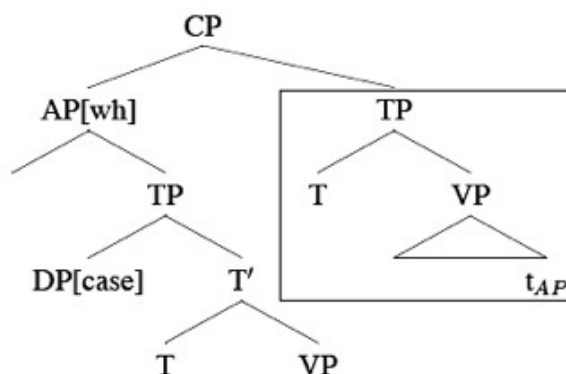
Let us now see whether this kind of approach ensures the correct sequencing of A-movement and *wh*-movement in cases involving subextraction from a moved phrase. A relevant example is repeated below (from Collins (2005)):

(25) *Oscar was asked how likely to win it was.

[_{TP} Oscar_{*i*} was asked [_{CP} [_{AP} how likely *t_i* to win]_{*j*} [_{TP} it was *t_j*]]]

Consider the point in the derivation where the embedded CP is merged. The *wh*-phrase moves to the embedded CP, pied-piping the DP with an unchecked case feature. The complement TP then spells out, making whatever is embedded inside inaccessible to further syntactic operations, as sketched in (26). Note however that the DP in need of case checking is not in the Spell-Out domain, but is embedded in the phrase occupying the specifier of CP, the escape hatch position.

(26)



The question is whether goals that are buried inside a phrase sitting in the specifier of a phase head are visible to outside probes or not. Nothing said thus far seems to guarantee their invisibility. It should thus be possible in principle to extract a DP out of the specifier of a CP and move it to a case position in the matrix clause. In other words, ‘smuggling’-type derivations, such as (26), do not seem to be ruled out unless additional stipulations are made.¹⁰

To summarize, we have seen that a theory which employs the GBOIM in combination with the phase-based locality derives a broad range of improper movement effects. In particular, it captures the fact that certain movement operations are clause-bound, i.e. cannot cross the CP boundary. It also predicts asymmetric feeding relations between different movement operations in identity cases.¹¹ Finally, ordering restrictions on movement operations in configurations involving remnant movement also straightforwardly fall under its scope. In addition, all this is achieved by relating movement to functional structure and without appealing to an extrinsic hierarchy of movement operations. The GBOIM does rely on a theory of locality to achieve its full

¹⁰Recall that Chomsky (2008) encounters the same problem when trying to account for asymmetries in extraction from subjects. He then assumes that extraction from within a phrase sitting in a phase edge is prohibited, though it is unclear what this prohibition follows from.

¹¹To be precise, the account predicts asymmetric feeding relations to hold between many, though not all movement operations in cases of cross-clausal movement. This particular prediction will be discussed in greater detail in the following chapter. For now, let me just point out that the asymmetric feeding relations that were shown in chapter 2 to be empirically well-supported are captured by the proposed account.

effects, but this is not a setback since locality is needed independently of improper movement effects.

The cases of subextraction however are not captured and require an additional assumption, prohibiting extraction from a phrase occupying a phase edge.¹² In the following sections, I will discuss some further challenges and problems that the proposal faces. I will first focus on ECM constructions and show that, with some modifications, the standard analysis of these constructions in terms of raising-to-object can be maintained. I will then close off the chapter with a discussion of a more serious challenge, revealing an incompatibility of the GBOIM with the phase-based locality the account relies on.

4.3 ECM constructions

A potential challenge to the validity of GBOIM is raised by ECM constructions, assuming that the now standard analysis of these constructions in terms of raising to object is correct (Postal (1974), Lasnik and Saito (1991), among others). On this approach, the subject of the embedded clause moves to a case-licensing position in the matrix clause. In (27) for instance, *Mary* first undergoes movement to the embedded TP, to satisfy the EPP feature of the T head. It then moves further to the object position of *believe*, which is taken to be some functional specifier position (AgrOP here). It is also necessary to assume that the verb *believe* moves to a position to the left of the landing site of the raised NP.¹³

- (27) a. John believes Mary to be a genius.
 b. [_{TP} John [_{XP} believes [_{AgrOP} Mary_i [_{TP} t_i to [_{VP} t_i be a genius]

If the GBOIM is adopted, then (27b) cannot be the correct analysis of (27a). Observe that in (27b) the ECM subject moves from Spec of embedded TP to Spec of AgrOP in the matrix clause. This movement step violates the GBOIM since TP is higher in the functional hierarchy than AgrOP. The GBOIM is compatible with the alternative analysis of (27a) in terms of exceptional case-marking, which used to be the standard account of ECM constructions in the Government and Binding era. On that approach, the ECM subject stays in SpecTP of the embedded clause, where it is case-marked by the matrix verb.¹⁴ However, there are good reasons to think that the ECM subject does in fact move to the matrix clause. The key piece of evidence concerns the surface

¹²As Abels (2008b) points out, it is more pressing to rule out cases involving remnant movement since all illicit examples of subextraction can in fact be derived if remnant movement is freely available.

¹³There is no consensus in the literature on what position the raised verb occupies. I won't have much to say about this issue, labeling the relevant projection simply as XP. I will focus rather on the positions that the raised NP can occupy.

¹⁴In the GB, it was assumed that a verb like *believe* can govern into an IP and assign case to its subject NP. The account of ECM constructions in English in terms of exceptional case-marking is generally disfavoured nowadays, though an updated version has been reintroduced

ordering of the raised NP and various kinds of main clause material. For instance, the raised NP can precede adverbs and particles that are part of the matrix clause, as originally observed by Postal (1974).

- (28) a. I believed Nixon incorrectly to be interested in ending the war. (Postal (1974))
 b. I have found Bob recently to be morose. (Postal (1974))
 c. Mikey made George out to be a liar. (Johnson (1991))

If the ECM subject remains in the Spec of embedded TP, then the surface strings in (28) are completely unexpected. Examples such as those in (28) clearly show that the final landing site of the ECM subject is in the matrix clause.

The overt raising analysis also easily captures the interpretational effects discussed by Lasnik and Saito (1991). Examples in (29) show that the ECM subject can bind into a matrix-clause adverbial.

- (29) a. the DA proved [no one to have been at the scene of the crime] during any of the trials.
 b. the DA proved [two men to have been at the scene of the crime] during each other's trials.

The negative polarity item *any* in (29a) requires a c-commanding negated phrase in order to be licensed. The fact that the negated ECM subject can license the NPI in the matrix adverbial suggests that it does appear in the (higher) c-commanding position in the matrix clause at some relevant level of representation. The same conclusion can be reached on the basis of the Condition A satisfaction in (29b).¹⁵ When the binder is unambiguously in the embedded clause, as in the corresponding examples with finite complements, binding into matrix adverbial is not possible.

- (30) a. ?*the DA proved [that no one was guilty] during any of the trials.
 b. ?*the DA proved [that two men were at the scene of the crime] during each other's trials.

Faced with such compelling evidence, we are forced to conclude that the ECM subject does indeed move to a case-licensing position in the matrix clause. We might however ask if there is equally compelling evidence that the launching site of this movement step is Spec of embedded TP. Consider an alternative derivation of (27a), sketched in (31).

- (31) [TP John [XP believes [AgrOP Mary_i [TP to [VP t_i be a genius]

to account for similar constructions in some other languages (see Bruening (2001)).

¹⁵When combined with Lasnik and Saito's (1991) arguments that covert raising does not give rise to new binding configurations, the examples in (29) strongly argue for overt raising in ECM constructions.

In (31), the ECM subject moves to AgrOP from the base-generated position in the lower VP. The raised NP thus ends up occupying a position in the matrix clause, as the evidence just reviewed suggests, but there is no stop-over in the intermediate SpecTP. This type of derivation does not violate the GBOIM. The rest of this section reviews evidence in favour of intermediate touchdown in SpecTP. The conclusion that will be reached is that this kind of evidence is rather slim, paving thus the way for an analysis of ECM constructions which conforms to the GBOIM.

4.3.1 On the status of EPP in non-finite TP

What forces the raised NP in ECM constructions to stop off in intermediate SpecTP? The answer is EPP. The EPP (or the Extended Projection Principle) was originally stated as a requirement that all clauses have a subject (Chomsky (1981, 1982)). The principle was intended to capture the ungrammaticality of constructions like (32), where the subject position of the matrix clause is not filled.

(32) *Is likely that Peter likes Mary.

The EPP was then generalized to non-finite clauses, making it a general requirement of T (finite or non-finite) to have a filled specifier. The stipulatory nature of EPP has been noted often enough and it seems that thirty years of research have not brought us any closer to understanding the true nature of EPP. As a result, there have been a number of attempts to abandon the EPP completely and deduce its effects from other independent principles (Grohmann et al. (2000), Bošković (2002), etc.). My goal here will not be to argue for total elimination of EPP from the grammar. Rather, I will focus my attention on non-finite clauses, leaving it as an open question whether the EPP should in general be dispensed with.

Postulation of the EPP forces A-movement to proceed in successive cyclic fashion through each intermediate SpecTP, as illustrated below for a raising (33a), and an ECM construction (33b), respectively.

- (33) a. John_i seems [_{TP} t_i to be likely [_{TP} t_i to be [t_i ill]]]]
 b. John believes Mary_i [_{TP} t_i to be [t_i ill]]]

We have already seen that the derivation in (33b) involving an ECM verb conflicts with the proposed GBOIM. However, the intermediate movement steps in subject raising constructions do not at first sight seem to be problematic in the same way. At no point in the derivation of (33a) does the raised subject move to a position lower in terms of the functional hierarchy than the launching site. We might therefore be inclined to limit our attention strictly to cases of ECM constructions, leaving the status of the raising TP aside. However, apart from the fact that without a specific proposal it seems rather ad hoc to assume that only the TP in ECM constructions does not have the EPP requirement/feature, a moment of reflection reveals that positing intermediate traces in SpecTPs in raising constructions would still be problematic

on an approach adopting GBOIM. The reason is that a raising construction can be embedded under an ECM verb. In this case, even if we were to assume that only the raising T, but not the ECM one, has an EPP feature, there would still be a movement step from SpecTP to SpecAgrOP, in violation of the GBOIM.

(34) John believes [_{AgrOP} Mary_i [_{TP} (t_i) to be likely [_{TP} t_i to cheat]]]

I conclude therefore that there cannot be a stop-over in intermediate TP in either subject raising or ECM constructions. If this conclusion is correct, then it calls for reevaluation of the evidence supporting the presence of intermediate traces in these constructions, a task to which I now turn.¹⁶

4.3.1.1 Binding facts

Theoretical motivation aside, one empirical argument for the presence of EPP in non-finite TP concerns the following paradigm, taken from Grohmann et al. (2000), who attribute the data to Danny Fox.

- (35) a. (i) John seems to Mary to appear to himself to be ill.
 (ii) John_i seems to Mary [_{TP} t_i to appear to himself t_i [_{TP} t_i to be t_i ill]]
 b. (i) *Mary seems to John to appear to himself to be ill.
 (ii) *Mary_i seems to John [_{TP} t_i to appear to himself [_{TP} t_i to be t_i ill]]

The argument goes as follows. *John* in (35a) raises from its base-generated position over the reflexive to Spec of intermediate TP, before landing in its surface position. As a result, the reflexive is locally bound in virtue of the trace in the intermediate SpecTP position (the pertinent trace is boldfaced). (35b) instantiates a ‘blocking effect’. The experiencer is attempting to bind the anaphor across a closer binder, namely the trace of *Mary* in SpecTP. The contrast in grammaticality between (35a) and (35b) is thus attributed to the presence of an intermediate trace in SpecTP.

Note that the argument outlined assumes that the experiencer embedded in a *to*-PP can bind a reflexive. If this were not the case, then the ungrammaticality of (35b) would simply be due to the fact that the experiencer is not a potential antecedent

¹⁶Interestingly, Bošković (2002) who argues for complete removal of the EPP from the grammar, nevertheless retains the intermediate movement steps in (33). According to him, there is strong evidence that A-movement proceeds successive cyclically through each intermediate TP, although the driving force behind these movements, he claims, is Takahashi’s (1993) Minimize Chain Links Principle, rather than the EPP. Clearly, this conclusion is again incompatible with the GBOIM. If the GBOIM is correct, then there simply cannot be a trace in intermediate SpecTP, regardless of the possible trigger. As we will see shortly, the evidence seems to be less compelling than Bošković claims.

and no appeal to a ‘blocking effect’ would be needed. Grohmann et al. (2000) provide the following examples to show that experiencers can indeed bind out of their PPs. In (36a), *John* is the only available antecedent for the reflexive. The fact that the sentence is grammatical thus suggests that the experiencer can act as a potential binder.¹⁷ (36b) shows that variable binding by a quantified NP embedded in a *to*-phrase is equally possible. Finally, (36c) shows that the experiencer can induce a Condition C violation.

- (36) a. It seems to John_i to appear to himself_i that Mars is flat.
 b. It seems to [every boy]_i to appear to his_i mom that Mars is flat.
 c. *It seems to him_i to appear to John_i that Mars is flat.

Even acknowledging that the experiencer is a potential binder, a careful consideration of the pattern in (35) nevertheless reveals that the desired blocking effect can be achieved even without successive cyclic movement through SpecTP. Consider the following alternative derivation of (35a):

- (37) [_{TP} John_i [_{VP} t_i seems to Mary [_{TP} to [_{VP} t_i V-appear [_{VP} to himself [_{TP} to be [_{VP} t_i ill]]]]]]]]]

In (37), the raised DP moves through each intermediate VP on its way to the matrix clause. *John* can therefore bind the reflexive from the Spec of the intermediate VP, where the binding configuration is sufficiently local. The raised DP in (35b), *Mary* in this case, would move in the same way to the matrix clause, by stopping off at each intermediate VP. This is illustrated in (38). The ‘blocking effect’ in (35b) can equally well be attributed to the trace of *Mary* in Spec of VP, rather than in the Spec of TP.

- (38) *[[_{TP} Mary_i [_{VP} t_i seems to John [_{TP} to [_{VP} t_i V-appear [_{VP} to himself [_{TP} to be [_{VP} t_i ill]]]]]]]]]

There is therefore no need to assume that a trace of the raised NP is present in the Spec of non-finite TP.

The following argument from Bošković (2002) is deficient in the same way. Consider the following pair of examples (originally from Lebeaux (1991)):

- (39) a. *[His_i mother’s_j bread] seems to her_j _ to be known by every man_i _ to be _ the best there is.

¹⁷Epstein and Seely (2006, p. 146, ft.31) note that the grammaticality status of (36a) is not entirely clear. Assuming that A-movement reconstructs, Bošković (2002) and Sportiche (2006) provide the following examples to show that experiencers c-command out of the PP for purposes of anaphoric binding:

- (i) a. Pictures of himself seem to John to be cheap. (Bošković (2002))
 b. Pictures of each other seemed to the boys to be fuzzy. (Sportiche (2006))

- b. [His_i mother's_j bread] seems to every man_i _ to be known by her_j _ to be _ the best there is.

According to Bošković (2002), the data in question can be easily accounted for if the matrix-clause subject passes through embedded SpecTPs, which can serve as reconstruction sites. (39a) is ruled out because the subject has to be reconstructed into the most embedded clause to license the bound variable reading, but it then induces a Condition C violation. On the other hand, in (39b) we can reconstruct the matrix-clause subject to the higher infinitival SpecTP position, where a bound-variable reading can be licensed without giving rise to Condition C violation. However, we can just as easily conclude from (39b) that the reconstruction site is on the edge of VP, not TP, and keep the rest of the argument intact. In other words, the paradigm in (39) might be taken as evidence for successive cyclic A-movement, though not necessarily for successive cyclic movement through SpecTP.

Is there a reason to think that the raised DP stops off in Specs of intermediate VPs, rather than TPs? In fact there is. If we adopt Legate's (2003) conclusion that all types of Vs, not just those that introduce an external argument, are phase heads, then the PIC would force successive cyclic A-movement through each intermediate VP.¹⁸ There is also an empirical argument to be made that successive cyclic A-movement targets intermediate VPs, rather than TPs. Consider first an example like (40) (taken from Sauerland (2003)).

- (40) Every child isn't smart.

(40) allows an interpretation where the universal quantifier is in the scope of the negation. This interpretation can be paraphrased as 'Not every child is smart', and is most natural if the sentence is followed by a clarifying continuation like 'In fact, half of them aren't smart'. We can derive this reading by reconstructing the subject to a position lower than negation.¹⁹

¹⁸As a result, there would be a single driving force behind all cases of successive-cyclic movement- namely, the presence of phase boundaries. It seems to me that this makes Chomsky's argument for phases regarding the reduction of computational complexity more compelling. If Vs headed by raising verbs would not constitute phases, then clauses containing raising predicates would have no phase-inducing heads and would be able to grow to unboundedly large structures before they are spelled out. Postulating the presence of a phasal head in these constructions partitions the derivation into smaller chunks, reducing thereby the computational load. Let me also note that the investigation of binding phenomena in chapter 7 will lead us to the conclusion that C and Asp (dominating v/V) are phase heads, forcing successive cyclic movement via AspP, rather than v/VP. However, since nothing crucial hinges on this assumption in the present context, the discussion in this section will be phrased in terms of C and v/V as the relevant phasal heads.

¹⁹The exact nature of the mechanism underlying reconstruction effects is not relevant for the argument to go through. I use the term *reconstruction* rather informally to refer to interpretation of non-final positions in non-trivial chains.

Consider next the example (41), discussed by Sauerland (2003).

- (41) Every child_i doesn't seem to his_i father to be smart.

In (41) the quantifier *every child* can bind the pronoun *his*, while at the same time taking scope under the negation. This interpretation can be paraphrased as 'It's not the case for every child that it seems to his father to be smart'. Sauerland (2003) provides several additional examples of the same type, with the subject scoping below negation and binding a pronoun in the experiencer argument of *seem*.

- (42) a. Every participant_i didn't seem to his_i coach to be in bad shape.
b. All linguists_i didn't seem to their_i employer to work hard.

To derive the desired scope effect, the subject must be reconstructed to a position below negation. However, the position in question cannot be the embedded TP since the possibility of binding into the experiencer would remain unexplained. The correct interpretation can however be derived if there is a reconstruction site at the VP edge, as sketched below. In this position, the subject is below the scope of the negation, but still high enough to bind into the experiencer.

- (43) every child_i doesn't [_{VP} t_i seem to his father [_{TP} (t_i) to be t_i smart]]

Sauerland (2003) also shows that the desired interpretation cannot be derived by raising negation over the subject. One piece of argument concerns the scope of negation relative to material occurring between the subject and negation. Both examples in (44) have only one interpretation where negation takes scope below the quantificational adverb, and the modal, respectively, suggesting that negation cannot raise over quantificational material.

- (44) a. Tom usually doesn't follow.
b. Jan mustn't get an A.

Compare examples in (44) with those in (45).

- (45) a. Every student usually doesn't follow. In fact, half of them usually don't follow.
b. Every student mustn't get an A. At most a third of them can get one.

In these cases, the subject is able to scope under negation. Sauerland (2003) provides the following scenario that brings about the relevant reading in (45b): A junior teacher gave every student in his class an A. However, the school has a rule that only a third of all students may get an A. A senior teacher could then use (45b) to reprimand his junior colleague. Note that even when the subject takes scope below negation, the modal *must* has to take scope above negation. This would be unexpected if movement of negation was the relevant mechanism to achieve wide scope of negation because in that case negation would outscope *must* as well. The right interpretation can there-

fore be achieved only by reconstructing the subject to a position below the negation, which does not affect the relative scope of the negation and the modal. The same point can be made on the basis of (45a).

To conclude, we have seen that whereas the paradigm in (35) does not decide between VP and TP as possible intermediate landing sites, examples like (43) do. We know on the basis of such examples that there must be a reconstruction site on the edge of VP since reconstructing the raised DP to intermediate TP cannot explain the relevant facts. However, what hasn't been established yet is that there *cannot* be a trace in the intermediate TP. What we have shown is that the data discussed by now can be accounted for by assuming that A-movement proceeds through each intermediate VP. Positing in addition traces in intermediate TPs would be redundant, but nothing said so far speaks against this option. However, as we will see in the rest of this section, there are reasons to believe that there is no stop-over in the intermediate TP. Before we turn to examples in support of this conclusion, let me first discuss another apparent piece of evidence in favour of successive cyclic A-movement through intermediate TPs that will play a role in a later argument. Again, the purported evidence can be reinterpreted as showing that long-distance A-movement proceeds through Specs of intermediate VPs, rather than TPs.

4.3.1.2 Quantifier Float

Since Sportiche (1986), floating quantifiers (FQs) have been taken as a standard diagnostic for A-movement. According to this analysis, FQs are generated together with the element they modify and subsequently stranded under movement of the element in question. If the stranding analysis is correct, then FQs mark positions from or through which a DP has moved.²⁰ Consider (46a) in this light:

- (46) a. I believe the students all to know French.
 b. I [_{XP} X-believe_i [_{AgrOP} the students_j [_{VP} t_i [_{TP} [all t_j]_k to [_{VP} t_k know French]]]]]]

²⁰There are two leading proposals regarding FQs: the adverbial analysis (Kayne (1975), Doetjes (1997), etc.) and the stranding analysis (Sportiche (1986), Shlonsky (1991), McCloskey (2000), Bošković (2004), among others). Although there are still many open questions on both approaches, proponents of the stranding analysis have summoned some very compelling evidence in their favour. One particularly telling observation is that in many languages, FQs show agreement (typically in case, number and gender) with the DP, strongly suggesting that they are modifiers of the DP with which they are initially merged, rather than adverbs (see for instance Shlonsky (1991) for Hebrew, Merchant (1996) for German). Discussing other evidence in favour of the stranding approach would take us too much afield. I will simply assume, as is standardly done, that FQs can be used as a tool to identify positions which the associate DP has occupied in the course of the derivation, and refer the reader to the relevant literature for justification of this view. Needless to say that if the stranding analysis eventually proves to be incorrect, then the data to be discussed cannot be used as evidence either in favour or against the proposal developed here.

On the standard derivation of ECM constructions, (46a) is simply taken to show that the quantifier is stranded in the Spec of the embedded TP, through which the DP passes on its way to the case-licensing position in the main clause, as illustrated in (46b). However, this type of analysis must now be reassessed if our claim that there is no intermediate stop-over in TP is correct. Now we have already reached the conclusion that successive cyclic A-movement targets each intermediate VP. This opens up another possibility of analysing (46a), sketched below:

- (47) I [_{XP} X-believe_i [_{AgrOP} the students_j [_{VP} all t_j]_k [_{VP} t_i [_{TP} [to [_{VP} t_k know French]]]]]]]

Thus, we are again faced with the conclusion that the purported evidence for the presence of an intermediate trace in SpecTP is far from convincing.²¹

An interesting set of data regarding quantifier float is discussed by McCloskey (2000). McCloskey (2000) investigates a variant of English (West Ulster English), which unlike Standard English, allows quantifier float under *wh*-movement. He provides the following pair of examples that are relevant for the discussion at hand:

- (48) a. Who did you expect your mother all to meet at the party?
b. *Who did you expect all your mother to meet at the party?

Assuming that the VP headed by the ECM verb is phasal, the possibility of stranding the quantifier to the right of the ECM subject occupying AgrOP is unsurprising, as all movement out of the domain of the phase is forced to proceed through the phase edge.²²

- (49) [_{CP} who_i did you [_{XP} expect [_{AgrOP} [your mother]_j [_{VP} [all t_i]_k [_{VP} t_j [_{TP} to t_j t_k meet at the party]]]]]]]

Now the example does not argue against the possibility of intermediate movement through SpecTP per se, but it does show that even on the standard account of ECM constructions the specifier of the matrix VP must be made available for intermediate stop-over (at least in the case of A'-movement). The contrast between (48a) and (48b) also shows what McCloskey (2000) intended it to show, which is that the ECM sub-

²¹The ungrammaticality of the examples in (i) does not bear on the possibility of an intermediate stop-off in SpecTP. On either view, the ill-formedness of these examples can be attributed to the ban on stranding a quantifier in a θ position (for more details see Bošković (2004)).

- (i) a. *The students believed John all to be smart.
b. *The students believed all John to be smart

²²I find this a more plausible analysis than McCloskey's suggestion that the embedded clause is a CP and the quantifier is stranded in the Spec of CP. This would lead to a problematic conclusion that A-movement out of a CP is possible in English (given that the ECM subject moves to a case position in the matrix clause).

ject must be moved overtly to the matrix clause. If it could remain in the embedded SpecTP, the ungrammatical (48b) would be derived. On our assumptions, the ECM subject never lands in SpecTP in the first place so the grammaticality judgement regarding (48b) is expected.²³

The considerations discussed above force on us the following analysis of quantifier float in subject raising constructions:

- (50) a. The students seem all to know French.
 b. The students_i [XP seem_k [VP [all t_i]_j [VP t_k [TP to [VP t_j know French]]]]]

The quantifier is again stranded at the edge of VP, which as we have already seen from Sauerland's data, must be available as an intermediate landing site.

Having shown that the quantifier float data can be handled without assuming an intermediate stop-over in SpecTP, let us now turn to another argument for the presence of the EPP feature in non-finite T, discussed by Lasnik (1999). Some further facts involving quantifier float will become relevant in evaluating the force of Lasnik's argument.

4.3.1.3 Scope interactions

While in his earlier work (Lasnik (1995)), Lasnik argued that overt raising of the ECM subject to the matrix clause is obligatory, in Lasnik (1999) he argues that it is optional. In other words, both derivations in (52) are available to (51).

- (51) John believes Mary to be a genius.
 (52) a. John believes [_{AgrOP} Mary_i [VP [TP t_i to be [VP t_i a genius]]]]
 b. John believes [TP Mary_i to be [VP t_i a genius]]

The argument rests on certain observations regarding scope in ECM constructions. According to Lasnik, (53) has two available interpretations, with the universal quantifier taking either narrow or wide scope with respect to negation.²⁴

- (53) I believe everyone not to have arrived yet. ($\forall > \text{neg}$, $\text{neg} > \forall$)

Lasnik (1999) analyzes these facts in the following way. When the ECM subject remains in embedded TP, the narrow scope reading with respect to the negation is available. This reading however is lost when the ECM subject raises overtly to the matrix clause. In that case, only wide scope reading for the universal is possible.

²³The contrast between (48a) and (48b) also suggests that AgrOP, the surface position of the ECM subject is above the phasal vP (with the concomitant need to raise the verb relatively high in English), rather than sandwiched between the VP shells, as Lasnik assumes in a series of works (Lasnik (1995), Lasnik (1999)). See also McCloskey (2000), Bošković (1997) for the same conclusion.

²⁴This observation goes back to Chomsky (1995).

This proposal leads to the prediction that in cases where it is clear that the subject must have moved to the matrix clause, the negation should not be able to outscope it. According to Lasnik, this prediction is borne out. Consider the following pair of examples:

- (54) a. The mathematician made every even number out not to be the sum of two primes. ($\forall > \text{neg}$, $*\text{neg} > \forall$)
 b. The mathematician made out every even number not to be the sum of two primes. ($\forall > \text{neg}$, $\text{neg} > \forall$)

In the rather unusual ECM particle construction above, many speakers allow the ECM subject to surface both to the right and to the left of the particle *out*. According to Lasnik, for speakers who do allow both word orders, there is a difference in interpretation though. In (54a), where it is clear from the word order that raising to matrix clause has taken place, narrow scope for the universal ECM subject is impossible. Apparently, the only available reading in (54a) is the implausible one where the mathematician was engaged in a futile activity of trying to convince everyone that no even number is the sum of two primes.

This analysis of the pattern in (54) makes a further prediction that the ECM subject should display the ‘high’ binding behaviour only when it precedes the particle since only in this case it occupies a position in the matrix clause. According to Lasnik, the results obtained by testing this prediction point in the right direction, though he concedes that more data collection is necessary. For example, a negative ECM subject to the left of *out* ‘seems much more comfortable with a negative polarity item in the matrix clause, than does one to the right of *out*’ (Lasnik (1999), p. 201).

- (55) a. the lawyer made no witnesses out to be idiots during any of the trials.
 b. $?*$ the lawyer made out no witnesses to be idiots during any of the trials.

The same subtle judgements arise with other height tests, such as anaphor binding.

- (56) a. the DA made the defendants out to be guilty during each other’s trials.
 b. $?*$ the DA made out the defendants to be guilty during each other’s trials.

There are several aspects regarding technical implementation of Lasnik’s proposal which are unclear or worrying. First of all, the analysis considerably complicates case assignment, with two different mechanisms available for checking structural case in ECM constructions. When overt raising takes place, the nominal checks its case features in AgrOP. On the other hand, when the ECM subject remains in the embedded clause, case is checked by covert raising of formal features to V. In addition, since Case features need not be checked overtly, movement to the matrix clause cannot be driven by Case. Lasnik assumes that the driving force for the overt raising to AgrOP is a strong EPP feature on AgrO. Case checking is then just a side-effect of EPP satisfaction. Since he argues that overt raising is optional, he then concludes that AgrO only optionally has the EPP feature, making it crucially different from AgrS

which obligatorily contains the EPP feature.

Let us now see what implications the scope interactions discussed by Lasnik have for an analysis of ECM constructions within a theory that endorses the GBOIM. The derivation in (52a) is not compatible with the GBOIM since it involves a movement step from SpecTP to AgrOP, as already discussed. However, Lasnik's account of scope interactions in terms of two surface positions for the ECM subject could in principle be made compatible with the GBOIM. All we need to do is to remove a trace in intermediate SpecTP in (52a), for which Lasnik provides no empirical argument. The scope facts can then be captured in the way Lasnik suggests, by allowing two derivations: one where the ECM subject remains in the embedded SpecTP, and the other where it moves to the matrix AgrOP, possibly via intermediate SpecVPs. The drawback of this kind of approach is that it would lead to the conclusion that the EPP feature in embedded TP is optional in a rather arbitrary way: whenever the ECM subject would have some business to do in the matrix clause, there would be no EPP feature in the intermediate TP. A more plausible alternative is to assume that intermediate TP is simply not available as a potential landing site. Thus, Lasnik's argument merits some more discussion.

An underlying assumption that Lasnik's argument is based on is that there is no reconstruction with A-movement. Consequently, two available scopes in ECM constructions such as (53), repeated below, suggest that the ECM subject can occupy two different positions at surface structure. On the other hand, if this assumption were incorrect, then the observed scope effects could be derived from a single derivation. For instance, the wide scope reading would be derived by interpreting the ECM subject in the surface position, in AgrOP. The narrow scope reading would be derived via reconstruction, by interpreting the ECM subject in the position of the trace in (57b). As a result, no movement to embedded TP would be required to capture the two interpretations.²⁵

- (57) a. I believe everyone not to have arrived yet. ($\forall > \text{neg}$, $\text{neg} > \forall$)
 b. I believed [_{AgrOP} everyone_i [_{VP} [_{NegP} not [_{TP} to [have [_{VP} t_i arrived yet]]]]]]]

In light of this, the claim that A movement does not reconstruct deserves careful consideration. Lasnik (1999) bases himself on an observation noted in Chomsky (1995). Chomsky (1995) provides the following paradigm:

- (58) a. (It seems that) everyone isn't there yet.
 b. Everyone seems not to be there yet.

In (58a), the negation can take wide scope over the quantifier. On the other hand, in (58b) where the subject undergoes A-movement, only the narrow scope for negation is available. Though these facts are suggestive, basing such a strong

²⁵For the time being, I am setting aside the high binding behaviour and the particle order with the *make out* examples.

claim solely on (58) seems premature. In fact, there is considerable evidence that A-reconstruction is possible. The most famous argument is given in May (1977), based on (59).

- (59) Some politician is likely to address John's constituency. (some > likely, likely > some)

The intuition about (59) is that it is ambiguous. According to May, "[59] may be taken as asserting either (i) that there is a politician, for example, Rockefeller, who is likely to address John's constituency, or (ii) that it is likely that there is some politician (or other) who will address John's constituency" (May (1977), p. 189).

Consider another example taken from Sportiche (2006), with three scope inducing elements involved: the existential quantifier introduced by the singular indefinite subject *a southerner*, the universal quantifier *every*, and the verb *predict*.

- (60) A southerner is predicted to win every senate race.
- a. It is predicted that for every senate race, a (different) southerner will win it.
 - b. For every senate race, there is a (different) southerner who is predicted to win it.
 - c. For every senate race, it is predicted that there is a (different) southerner who will win it.

In principle, there should be six scopal readings available, but the three in which *a southerner* outscopes *every senate race* are pragmatically odd (since they require the same southerner to win every senate race). The two readings in (60b) and (60c) express a summary of individual predictions, one for each senate race, and are according to Sportiche difficult to test. The most natural reading is the one in (60a), which reports a unique global prediction. This suggests that at least one input for scope computation is (61), where both the universal and the existential are under the scope of the verb. Then, within a single clause, *every senate race* can outscope *a southerner*.

- (61) is predicted [a southerner to win every senate race]

The issue regarding the availability of A-reconstruction is complex and far from settled, and I cannot do justice here to all the literature written on the topic. However, examples such as these just discussed suggest that reconstruction with A-movement is in principle possible, a conclusion I will adopt here (see also Sportiche (2006) for a longish list of authors supporting this position). The fact that (58b) is unambiguous presents then an interesting challenge to any theory of reconstruction, but it does not necessarily show that A-reconstruction does not exist. As Bobaljik and Wurmbrand (1999) point out, the lack of *some* expected non-overt readings simply reveals that we do not yet fully understand all the factors influencing the availability of non-overt

scope interpretations.

The possibility of A-reconstruction then takes care of the ambiguity in (57a), without the need to posit two derivations of ECM constructions. However, the true force of Lasnik's proposal lies in the prediction that in cases where the ECM subject clearly occupies a position in the matrix clause, only the narrow scope for negation is available. The relevant examples are those showing 'high' binding behaviour of the ECM subject, such as (62).²⁶

- (62) John proved [every defendant]_i not to be guilty during his_i trial. ($\forall > \text{neg}$, * $\text{neg} > \forall$)

To get the bound variable reading of the pronoun which is contained in the matrix clause adjunct, the ECM subject must move to a c-commanding position in the matrix clause. In this case, the universal must take wide scope with respect to negation. How could these facts be captured assuming that A-reconstruction exists? In fact, observations similar to these have been used to argue in favour of A-reconstruction. Consider (63) (from Lebeaux (2009)). The sentence is ambiguous, allowing the two interpretations given in (63a) and (63b).

- (63) Two women seem to be expected to dance with every senator.
 a. There are two women who dance with every senator ($2 > \forall$)
 b. Every senator has two women who dance with them - not necessarily the same two women ($\forall > 2$)

Interestingly, placing an anaphor in the top clause freezes the scope ordering of the quantifiers. Thus, (64) is unambiguous, and has only the reading given in (63a).

- (64) Two women_i seem to each other_i to be expected to dance with every senator.
 ($2 > \forall$, * $\forall > 2$)

Lebeaux (2009) calls this a 'trapping' effect. Assuming that there is A-reconstruction, opens a way of approaching these facts. To get the lower scope reading of *two women* in (63), the raised subject must be reconstructed into the most embedded clause, producing an LF representation as in (65a). Being in the same clause, the two quantifiers can now scope over each other and the ambiguity arises. The sentence in (64) 'traps' the quantifier upstairs. If the quantifier were reconstructed, the putative LF representation would be as in (65b). The reconstructed position is too low to bind the anaphor in the matrix clause. As a result, it is not possible to get the inverse scope without sacrificing binding at the same time.

- (65) a. e seem e to be expected two women to dance with every senator.

²⁶Lasnik (1999) doesn't in fact discuss these examples. The example given here is taken from Boeckx (2001). I assume that the data displaying other kinds of high binding behaviour yield the same results.

- b. e seem to each other e to be expected two women to dance with every senator

(64) also shows that the proper way to derive the wide scope reading of the universal in (63) is not by scoping the downstairs quantifier upstairs via QR. If the wide scope reading could be achieved by raising *every senator* above *two women*, the addition of the anaphor in the matrix clause should not make a difference, and we would still expect both scopes to be possible.

I conclude therefore that the limited scope possibilities in ECM constructions that require binding within the matrix clause should also be attributed to ‘scope trapping’, rather than to two possible Spell-Out positions of the ECM subject. In fact, Lasnik’s account becomes dubious since it cannot be extended to the similar facts just discussed, involving subject raising constructions.

The strength of Lasnik’s argument now rests on examples featuring the ECM verb *make out*. The relevant patterns are repeated below.

- (66) a. The mathematician made every even number out not to be the sum of two primes. ($\forall > \text{neg}$, $*\text{neg} > \forall$)
 b. The mathematician made out every even number not to be the sum of two primes. ($\forall > \text{neg}$, $\text{neg} > \forall$)
- (67) a. the lawyer made no witnesses out to be idiots during any of the trials.
 b. ?*the lawyer made out no witnesses to be idiots during any of the trials.
 a. the DA made the defendants out to be guilty during each other’s trials.
 b. ?*the DA made out the defendants to be guilty during each other’s trials.

If the judgements indicated prove to be valid, then some additional assumptions may be required.²⁷ I have nothing interesting to say about these examples at this point.

Let me now turn to another set of facts which seem to suggest that Lasnik’s analysis of scope effects in ECM constructions might not be on the right track. Recall that floating quantifiers are standardly taken as a tool to identify positions which the raised DP has occupied during the course of the derivation. One observation regarding quantifier float that becomes relevant now is that floated quantifiers are frozen in scope in their surface position (Dowty and Brodie (1984)). Consider the scope interaction between negation and the floated quantifier in (68) (examples taken from Bošković (2004)).

- (68) a. The students did not seem all to know French. ($*\forall > \text{neg}$, $\text{neg} > \forall$)
 b. The students seemed not to all know French. ($*\forall > \text{neg}$, $\text{neg} > \forall$)
 c. The students all seemed not to know French. ($\forall > \text{neg}$, $*\text{neg} > \forall$)

²⁷There is some controversy regarding the given judgements. For instance, Bošković (2002) notes that Norbert Hornstein disagrees with the judgements, and Lasnik himself concedes that the judgements are subtle and need to be checked more carefully.

The same observation can be reproduced with other quantificational material.

- (69) (from Bošković (2004))
- a. The students rarely all go to the Zoo. ($*\forall > \text{rarely}$, $\text{rarely} > \forall$)
 - b. The students all rarely go to the Zoo. ($\forall > \text{rarely}$, $*\text{rarely} > \forall$)
- (70) (from Bobaljik (1993))
- a. The contestants could have all won. ($*\forall > \diamond$, $\diamond > \forall$)
 - b. All the contestants could have won. ($\forall > \diamond$, $\diamond > \forall$)

Consider (70). A sentence such as (70b) where the quantifier is not stranded is ambiguous with respect to the relative scope of the universal quantifier and the modal. On the wide scope for the universal, the sentence asserts that the predicate [*can win*] is true of all the contestants, i.e. that any of them can win. On the reading where the universal takes scope under the modal, (70b) asserts that a universal tie is possible, e.g., every one of the contestants will receive a prize. The example (70a) has only the second reading, i.e. the one where the floated quantifier takes scope in its surface position, under the modal. We might interpret these facts as showing that a quantified DP can undergo reconstruction, but a quantifier separated from the DP cannot. Regardless of how this restriction is derived, the fact that FQs are restricted to taking scope in their surface position makes examples involving floating quantifiers particularly relevant in evaluating potential analyses of scope interactions in ECM constructions.

Recall that on Lasnik's view, the scope is read off the surface structure, without resorting to reconstruction. Thus, the ambiguity of (71) is due to the availability of two derivations, sketched in (71a) and (71b). When the quantified DP raises to the matrix clause, only the wide scope with respect to negation is available. On the other hand, when the quantified DP remains in the embedded SpecTP, the narrow scope reading becomes possible as well.²⁸

- (71) I believe everyone not to have arrived yet. ($\forall > \text{neg}$, $\text{neg} > \forall$)
- a. I believe_j [_{AgroP} everyone_i [_{VP} t_j [_{TP} t_i [_{TP} not [_{T'} to [have t_i arrived yet]]]]]]]
 - b. I believe_j [_{VP} t_j [_{TP} everyone_i [_{TP} not [_{T'} to [have t_i arrived yet]]]]]]]

Consider now the following example involving quantifier float. Given Lasnik's assumptions, (72) would plausibly be derived as illustrated in (73).

- (72) I believe the students all not to have arrived yet.

²⁸It is not entirely clear what status negation has on Lasnik's view. I assume here that it adjoins to TP. If Neg projected on top of TP, then the DP in (71b) would have to occupy Spec-NegP, rather than SpecTP at Spell-Out. Note also that some scope reconstruction might be needed on Lasnik's view after all, since negation does not seem to c-command the universal.

- (73) I believe [_{AgrOP} [the students]_i [_{VP} [_{TP} [all t_i]_j [_{TP} not [_{T'} to have t_j arrived yet]]]]]]

The DP would first move to SpecTP to check the EPP feature of T. From this position, it would continue to move to the matrix clause, leaving the quantifier behind. As already discussed, this is the standard analysis of quantifier float facts. If Lasnik is correct that the narrow scope reading with respect to negation becomes available when the quantifier occupies the embedded SpecTP, then we expect (72) to allow this reading. This prediction does not seem to be borne out however. According to my informants, (72) allows only the wide scope reading of the universal.

I see no simple way of getting around this problem on Lasnik's view. Note that the floated quantifier occupies the same position as the quantified DP in (71b). We might hypothesize that negation in (72) and similar constructions can only scope over full quantified DPs sitting in SpecTP, but not over floated quantifiers. Beside the fact that it is difficult to see what this might follow from, it would make the account of scope facts in ECM constructions completely unrelated to observations regarding (68), (69), and (70). Alternatively, we might argue that stranding in intermediate SpecTP is for some reason prohibited. If such a reason were found, then we are led back to the conclusion (argued for in this section) that quantifier float facts do not argue for the presence of an intermediate stop-off in SpecTP. Rather there must be another landing site for the ECM subject in the matrix clause where the quantifier could be stranded.

The alternative account of scope interactions in ECM constructions that I have been assuming leads to a more coherent understanding of the non-ambiguity in case of (72). On my assumptions, the quantified DP in (71) always moves all the way to AgrOP, as illustrated in (74a_{ii}). The ambiguity is due to the possibility of reconstructing the DP to the trace position below the negation.²⁹ The unavailability of the narrow scope for the universal in quantifier float cases is due to the noted general impossibility of reconstructing floated quantifiers.

- (74) a. (i) I believe everyone not to have arrived yet. ($\forall > \text{neg}$, $\text{neg} > \forall$)
 (ii) I believe_j [_{AgrOP} everyone_i [_{VP} t_i [_{V'} t_j [_{NegP} not [_{TP} to [have t_i arrived yet]]]]]]]
 b. (i) I believe the students all not to have arrived yet. ($\forall > \text{neg}$, $*\text{neg} > \forall$)
 (ii) I believe_j [_{AgrOP} [the students]_i [_{VP} [all t_i]_k [_{V'} t_j [_{NegP} not [_{TP} to [have t_k arrived yet]]]]]]]

To summarize, we have seen that there is an alternative way of analysing Lasnik's scope facts in ECM constructions, without assuming that the SpecTP of the ECM infinitive is either an intermediate or the final landing site for the raised DP. The crucial difference between the two proposals is the assumption regarding the availability of

²⁹There is no problem here with allowing the negation to head its own projection on top of TP.

reconstruction with A-movement, A-reconstruction is assumed to be possible on the approach argued for here, but not on Lasnik's view. As already noted, the jury is still out on the issue regarding the availability of A-reconstruction, but there seems to be suggestive evidence that A movement can reconstruct. If so, then the scope facts in ECM constructions can be derived from a single derivation. What is more, the limited scope possibilities in quantifier float cases make Lasnik's account dubious, but are expected under an A-reconstruction analysis. Finally, the fact that the quantified ECM subject cannot scope below negation when it displays the high binding behaviour is just one instance of a more general phenomenon, referred to by Lebeaux (2009) as 'scope trapping'.

4.3.1.4 Against the EPP in non-finite T

So far, I have been investigating evidence given in the literature as support for the presence of the EPP requirement/feature in non-finite TPs, forcing in turn A-movement to proceed successive cyclically through each intermediate TP. I believe to have shown that the purported evidence is far from conclusive, and that most, if not all of the data can be reanalyzed by having A-movement proceed through Specs of intermediate VPs, and without a stop over in SpecTP. Before concluding this section, let me briefly mention some arguments against the presence of the EPP feature in non-finite TP. The arguments noted here are discussed in more detail in Grohmann et al. (2000) and Bošković (2002).

• Merge-over-Move preference

Consider the pair of examples in (75) (from Chomsky (1995)).

- (75) a. There_{*i*} seems t_{*i*} to be someone in the room.
 b. *There seems someone_{*i*} to be t_{*i*} in the room.

Under the assumption that the non-finite T hosts an EPP-feature, two options become available at the point when the embedded T is merged: (i) merge of *there*, or (ii) movement of *someone*. Choosing either option suffices to satisfy the EPP. The question then arises why (75b) is ungrammatical. To rule out this example, Chomsky (1995) suggests the Merge-over-Move preference. Movement is treated as a complex operation, involving Merge as one its suboperations, and is therefore argued to be dispreferred. The possibility of merging the expletive thus blocks movement of the indefinite. There are a number of problems with the Merge-over-Move account however. To name one, consider the following example, where the indefinite has moved to SpecTP, although an expletive was available for lexical insertion (from Grohmann et al. (2000)):³⁰

- (76) There was a rumor that a man_{*i*} was t_{*i*} in the room.

³⁰To handle these facts, Chomsky (2000) introduces the concept of subnumeration.

The contrast in (75) can be straightforwardly dealt with if there was no EPP in the non-finite TP. In that case, the non-finite T hosts no features in need of checking, therefore nothing needs to either move or merge there. On this view, one way of analyzing (75a) is to assume that the expletive merges directly in its surface position (see for instance Bošković (2002) who extensively argues for this kind of analysis).³¹

(77) There seems [_{TP} to be a man in the garden]

• **Double *there***

Another problem for the EPP is raised by the double-*there* constructions, such as (78):

(78) *There seems there to be someone in the garden.

According to Bošković (2002), it is difficult to see how (78) should be ruled out, especially if Chomsky (1995) is correct in assuming that *there* does not have case. On the other hand, if there is no EPP in non-finite TP, there is no reason to merge the expletive.

• **‘Wanna’-contraction**

Elimination of the EPP from non-finite T has interesting consequences for the analysis of ‘wanna’-contraction. The core observation regarding ‘wanna’-contraction is that intervening traces of A’-movement block the contraction (79a), but PRO and a trace of A-movement do not ((79b) and (79c)).

- (79) a. *Who do you wanna be there?
 b. He doesn’t wanna scare them.
 c. Little heed is gonna be paid to my proposal.

The standard analysis of these effects relies on the assumption that only case marked empty categories block the contraction (Jaeggli (1980)). Though the distinction seems to yield the right results, it is unclear why case should make an expression a phonetic “intervenor”. A simpler alternative would be to assume that any intervening empty category blocks contraction. This alternative however is not available as long as non-finite T has an EPP feature, forcing the presence of either PRO or trace in SpecTP. On the other hand, if the EPP feature is removed from non-finite T, then the simpler analysis becomes viable. In raising and control constructions there would be no intervener in SpecTP, making *wanna*-contraction possible. The surface position of PRO in control constructions then must be SpecVP, rather than SpecTP, as illustrated below. This is essentially what Baltin (1995) proposes.

³¹Bošković (2002) also presents an argument from ellipsis constructions against the presence of EPP in non-finite T. I won’t discuss this argument here.

- (80) He doesn't [_{VP} want [_{CP} [_{TP} to [_{VP} PRO [_{V'} scare them]]]]]

4.3.2 Interim summary

It has been shown that the standard analysis of ECM constructions conflicts with the GBOIM because it involves a movement step targeting a position lower in terms of the functional hierarchy than the launching site. The overt raising approach to ECM constructions can however be reconciled with the GBOIM if the trace in intermediate TP is eliminated. This in turn leads to the conclusion that there can be no EPP feature in non-finite T. That such a conclusion is not necessarily problematic has been shown by reanalyzing the relevant empirical data without a stop-over in non-finite TP. However, unlike some other approaches that argue for elimination of the EPP in non-finite T (or in any T for that matter), the proposal here does not try to eliminate A-chains altogether. In fact, we have seen that there are good reasons to think that A-movement proceeds successive cyclically, though intermediate movement steps do not target SpecTPs, but rather SpecVPs. I conclude therefore that ECM constructions do not present an insurmountable problem to the GBOIM.³²

4.4 Long movement out of a CP

Chomsky (2000) states that in addition to the CP, the transitive vP is also a phase. If v is a phasal head, then the PIC should force any phrase moving out of the vP to stop off at the edge of vP. That successive cyclic A'-movement indeed uses Specs of vPs as intermediate landing sites has been argued by Fox (1999) on the basis of reconstruction facts. Consider the following pair of examples:

- (81) a. [Which of the papers that he_i gave Mary_j] did every student_i ✓ ask her_j to read * carefully?
 b. *[Which of the papers that he_i gave Mary_j] did she_j * ask every student_i to revise *?

The interesting thing about (81a) is that it allows a reading where the first pronoun is bound by the quantified subject, while the second pronoun is coreferential with the R-expression, *Mary*, in the relative clause. The *wh*-expression however does not fulfill the c-command requirement on binding of pronouns in its surface position, leading to the conclusion that the correct interpretation must be derived via reconstruction. However, reconstructing the *wh*-phrase to the most embedded clause would give rise to a violation of the Binding Principle C. If (81a) is to be derived via reconstruction, then the *wh*-expression must be reconstructed to a position in between *every student*

³²ECM constructions present a more serious problem for Williams's LEC approach, so much that the raising of the ECM subject can no longer be treated as movement proper. Williams suggests that these constructions should be analysed as 'mismappings', a strategy he uses to derive certain marked word orders.

and *her* - a vP adjoined position. In (81b) on the other hand, the *wh*-phrase must reconstruct to its merged position in order for *he* to be bound by *every student*, but this then leads to a violation of Condition C, since *she* c-commands *Mary*. Hence, the sentence is ungrammatical.

Legate (2003) points out that passive and unaccusative verbs show the same reconstruction effects. This is illustrated in (82) for a passive verb. In (82), *Mary* keeps being introduced to her own date at parties.

- (82) a. [At which of the parties that he_i invited Mary_j to] was every man_i ✓
introduced to her_j *?
b. *[At which of the parties that he_i invited Mary_j to] was she_j *
introduced to every man_i *?

Like the sentences in (81), those in (82) contain a *wh*-word which must reconstruct below *every man* in order for *he* to be bound, but above *Mary* in order to avoid Principle C violation. This leads to the conclusion that any type of V, not just the transitive one, constitutes a phase.³³

In the previous sections, we have seen that there is evidence that successive-cyclic A-movement also proceeds via the v/VP edge. There is therefore relatively compelling evidence that v/V is a phase head, and this assumption enabled us to reanalyze ECM constructions in a manner compatible with the GBOIM-based approach.

Nevertheless, it is precisely the phasehood of v/VP which poses a serious problem for the analysis of improper movement developed here. Given that both C and v/V are phase heads, a phrase undergoing long movement must stop off at every intermediate C and v/V head lying on the path of movement. This is sketched below.

(83)

[_{CP}XP [C ... [_{vP} <XP> [v ... [_{CP} <XP> [C ... [_{vP} <XP> [_v [_{VP} V <XP>]]]]]]]]]]]

└-----┘

The problem with the derivation in (83) is that the dashed movement step violates the GBOIM, since the target of movement is lower in terms of the functional hierarchy than the launching site. Thus, locality forces an intermediate stop-over in SpecvP, but this intermediate stop-over violates the GBOIM.

As far as I can see, this problem persists even if we assume that the Spell-Out of phases is delayed until the next higher phase head is merged, as Chomsky (2001) suggests. To see this, consider (84) (phase heads are boldfaced). At the point when the embedded CP is merged, the material inside the lower vP phase is still accessible to it (see (84a)). The XP with an unchecked feature then moves to the SpecCP, without a stop off in vP. The vP then spells out. Note that the XP has to move at this

³³Legate (2003) also goes through other evidence for vP phases, showing that they equally well support the phasehood of unaccusative and passive VPs. I won't discuss these diagnostics here and refer the interested reader to Legate's work.

point. If it remains inside the vP, it will be spelled out and become inaccessible to any further syntactic operations. At the point where the matrix vP is merged the problem again arises (see (84c)). XP cannot move from CP to vP because that would violate the GBOIM. However, if it stays in CP, it will be spelled out as soon as the matrix vP has checked all its features and thus will not be available for attraction by a higher probe in the matrix clause.

- (84) a. $[_{CP} \uparrow [C [_{TP} T [_{vP} v [_{VP} V XP]]]]]$
 b. $[_{VP} [V [_{CP} XP [C [_{TP} [T' T vP]]]]]]$
 c. $[_{vP} \uparrow [v [_{VP} V [_{CP} XP [C [_{TP} [T' T vP]]]]]]$

Modifying slightly this view by allowing the edge of the phase to remain accessible after the phase itself has spelled out would allow long distance movement, but would also rule in constructions displaying improper movement effects. Two potential derivations need to be considered. If the XP moves, as in the previous case, by skipping the most embedded vP and targeting first the embedded CP, the right result seems to be derived. At the point where the matrix CP is merged, the XP occupying the edge of the embedded CP will still be accessible and could move directly to the matrix CP, without a stopover in vP.

- (85) $[_{CP} \uparrow [C [_{TP} T [_{vP} v [_{VP} V [_{CP} XP [C TP]]]]]]$

However, the problem is that the following derivation is also allowed on this view. The XP could first move to the edge of the most embedded vP. When the next higher phase head is merged, namely the embedded C, the complement of vP will be spelled out, but the edge of vP will not. As a result, the XP sitting in the edge of vP could still be attracted and move directly to the matrix vP, without a stop-over in the embedded CP, as in the (a) example below. From the edge of the matrix vP, it could then be attracted to the matrix TP, without violating the GBOIM. This way we end up deriving the ungrammatical superraising constructions, the very constructions we set out to exclude.

- (86) a. $[_{vP} \uparrow [v [_{VP} V [_{CP} C [_{TP} T [_{vP} XP [v vP]]]]]]$
 b. $[_{TP} \uparrow [T [_{vP} XP [v [_{VP} V [_{CP} C TP]]]]]]$

As far as I can see, the only way out of this conundrum would be to assume that Vs headed by bridge verbs, i.e. verbs that allow long extraction out of their complement clauses, are not phasal. However, such a conclusion would conflict with other arguments supporting the phasehood of vP, such as those based on reconstruction effects reviewed above. In addition, it is hard to see why this class of verbs should differ from all other verb types in not constituting a phasal boundary. If even rais-

ing Vs are phasal, as we have argued in the previous section, then non-phasehood of the bridge V would be extremely difficult to substantiate. The problem regarding the phasehood of v/V is thus a serious one, and reveals an internal inconsistency of the GBOIM-based approach. On the one hand, without the phase-based locality, the GBOIM captures only a small subset of improper movement operations. On the other hand, once the phase-based locality is adopted, the GBOIM rules out a number of licit movement operations.³⁴ In the following chapter, I will propose an alternative account of improper movement violations that will still rely on the phase-based locality, but will dispense with the GBOIM. The problem noted in this section will therefore not arise.

4.5 Conclusions

In this chapter, I have showed that the following constraint, in tandem with a phase-based theory of locality, rules out a broad range of illicit movement operations that we have argued should be viewed as falling under the rubric of improper movement:

- (87) *Generalized Ban on Improper Movement*
 A phrase cannot be moved to a position lower in terms of the functional hierarchy than the launching site of movement.

In particular, we have seen that the GBOIM is capable of accounting for asymmetric feeding relations between different movement operations in cases involving consecutive movements of the same phrase, as well as in remnant movement configurations. It also captures clause-boundedness of certain movement operations. Finally, these results are achieved without postulating and appealing to an extrinsic hierarchy of operations. What is needed is the hierarchy of functional projections and a phase-based theory of locality, both of which are required independently of improper movement effects.

Nevertheless, several problems for the GBOIM-based approach have been noted. First of all, ordering restrictions on movement operations in cases involving subextraction out of moved phrases are not captured. The reason is that nothing in the account developed so far prohibits extraction from within the phrase occupying the phase edge. Therefore, illicit movement operations in these configurations can only be ruled out by appealing to additional stipulations. In addition, we have seen that there is an internal inconsistency in combining the GBOIM with the phase-based locality. Locality requires long movement out of the CP to proceed via Specs of intermediate CPs and vPs, as these are phasal heads. However, the movement step from the Spec of embedded CP to the Spec of matrix vP is prohibited by the GBOIM,

³⁴Note that the GBOIM, in conjunction with the phase-based locality, rules out any movement out of the CP that targets a position in the matrix middle field. I won't discuss here if and how problematic this prediction is, given that the problem related to the phasehood of v is already serious enough to call for a revision of the proposal.

since v is lower than C on the functional hierarchy. Consequently, when combined with phase-based locality, the GBOIM in fact rules out long movement out of the CP. These shortcomings call for reevaluation of the proposal. In the following chapter, an alternative, though a related account, will be developed, which is capable of overcoming the problems of the GBOIM-based approach. In addition, it will be shown that the GBOIM can in fact be dispensed with, and its effects derived from the interaction between the functional hierarchy and a phase-based locality theory in a more direct way.

Chapter 5

Deriving improper movement from the functional sequence

In the previous chapter, I have explored one possible way of relating improper movement to the functional sequence, namely by encoding restrictions on ordering of operations in terms of possible landing sites. The proposed account relied on the following constraint on movement:

- (1) *Generalized Ban on Improper Movement* (GBOIM)
A phrase cannot be moved to a position lower in terms of the functional hierarchy than the launching site of movement.

Setting aside the problems that the account faces, one advantage of the GBOIM-based approach to improper movement is that there is no need to postulate an extrinsic hierarchy of operations, as on some alternative approaches. The only hierarchy appealed to is the independently needed structural hierarchy. The goal of this chapter is to go one step further and ask whether we can also do without the GBOIM. In other words, does the GBOIM need to be stated as an independent constraint on movement, or can improper movement effects be derived from the functional hierarchy in a more direct way? In this chapter, I argue that this might be achieved by relying on the internal structure of the moving phrase itself, again in conjunction with a phase-based theory of locality. It will then be shown that this alternative account also overcomes the problems that were plaguing the GBOIM-based approach.

The chapter is organized as follows. As the analysis to be put forth relies on some ideas sketched in Svenonius (2004), I will start off by presenting Svenonius's proposal. The alternative account is then developed in section 5.2, and shown to be capable of accounting for improper movement effects in both identity and non-identity cases. Section 5.2.2 shows that long movement via the matrix vP is not a problem on this account, as it was on the GBOIM-based approach. The proposal also makes a number of predictions distinct from those made on an approach adopting the Hierarchy of Operations. These differences are discussed in section 5.2.5. Section

5.3 summarizes and closes off the chapter.

5.1 Svenonius (2004)

The account of improper movement proposed in Svenonius (2004) builds on the hypothesis that there is a close correspondence between the functional architecture of clauses and noun phrases. That there are parallels between nominal and clausal structure has long been noted, at least since Chomsky (1970) and Abney (1987). Over the years, numerous studies have offered further empirical support for the existence of such parallelism, as well as shown that the functional structure of both noun phrases and clauses is quite rich. Relying on these observations, Svenonius (2004) assumes that clausal projections have their counterparts in the nominal domain. On this view, a possible arrangement of features in the clausal and nominal structure might be as in (2). For the sake of illustration, it suffices to assume for now a rather coarse clausal structure, as in (2a).^{1,2}

- (2) a. C – T – *v* – V
 b. Q – K – *n* – N

Pursuing further the correspondence between the clausal and nominal domain, Svenonius (2004) assumes that there are (at least) two phase heads in the DP, corresponding to C and *v* in the clause. When the phase heads are merged, their complements spell out, as in the clause. This is illustrated below, with brackets indicating spelled-out domains and asterisks marking phasal heads. Since the label of the phase remains visible after its complement has undergone Spell-Out, it is still possible to attract or agree with the phase, but not with anything inside it.³

¹An alternative would be to assume that the organization of features in the nominal domain is the mirror image of the clausal featural organization, as suggested by Starke (2003). I will not explore this possibility here.

²The label ‘K’ is taken from an old tradition of postulating K[ase] projection in the noun phrase. However, it should be noted that it is not Svenonius’s intention to suggest that K can be straightforwardly identified with case. As he points out, on a fine-grained feature structure, it is to be expected that descriptive labels such as ‘case’ are too coarse. Morphology associated with the descriptive label ‘case’ need not reflect only features of K. See Svenonius (2004) for more details.

³Svenonius (2004) ends up adopting a slightly more complicated model, expanding (3) into (i). On this view, the trigger for Spell-Out of vP is an aspectual head Asp, and the trigger for Spell-Out of CP is a topical head Top.

- (i) a. Top* – [C – T – Asp* – [*v* – V]
 b. Op* – [Q – K – Num* – [*n* – N]

For the sake of simplicity and legibility of the tree representations, I will stick to (3), presenting the approach in terms of this model (note that the tree representations of the derivations to be discussed here are provided by me on the basis of Svenonius’s (2004) assumptions

- (3) a. $C^* - [T - v^* - [V$
 b. $Q^* - [K - n^* - [N$

Assuming then that A-type attractors attract features within a lower phase, while A'-attractors operate on features in a higher phase, this opens up a way of deriving the ban on improper movement. At the point in the derivation where A'-attractors are merged, the features in the lower phase are already spelled out and inaccessible for attraction. As a result, there can be no A-operation following A'-movement. Before I discuss in more detail how this works, let me first address a potential problem that the assumed phasehood of the nominal heads raises on the model sketched so far. If QP is a phase on a par with CP, then KP spells out as soon as Q is merged. This means that if a noun phrase as large as a QP was constructed and merged as the complement of V, then this noun phrase could never undergo A-movement out of VP, because KP would already be spelled out and inaccessible. Svenonius (2004) suggests that the problem might be circumvented by assuming that nominal phasal heads trigger Spell-Out of their complements only after being merged with the clausal phasal heads. According to Svenonius, the reason behind this might be that features of the nominal heads are uninterpretable until they are valued by the corresponding clausal head, so that a nominal phasal head does not function as a trigger for Spell-Out until its features are checked by the clausal phasal head.⁴ Given this assumption, a phrase as large as QP could be merged quite low without causing any of its parts to spell out too early to be attracted by higher probes.

A derivation involving long-distance *wh*-movement compatible with the model sketched so far is given below (phase heads are marked with an asterisk, the spelled out domains are marked with a strikethrough). (4a) illustrates the point in the derivation when the embedded vP enters the structure. QP is then merged in the Specifier of vP. Since *v* is a phase head, it triggers Spell-Out of its complement VP. It also values the features of the corresponding nominal phasal head, *n*. Once features of *n* are valued, *n* triggers Spell-Out of NP. Note that at this point QP does not function as a phase. QP will only trigger Spell-Out of its complement once it moves to CP and has its features checked. In the next step, represented in (4b), QP moves to the TP to check the K feature.

and are not to be found in his article).

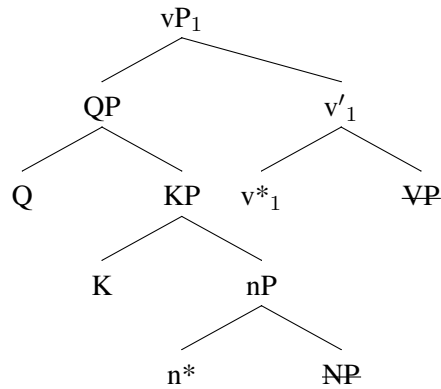
⁴Assuming a finer-grained functional hierarchy as given below, Svenonius (2004) suggests that this might be implemented in the following way.

- (i) a. $Top^* - [C - T - Asp^* - [v - V$
 b. $Op^* - [Q - K - Num^* - [n - N$

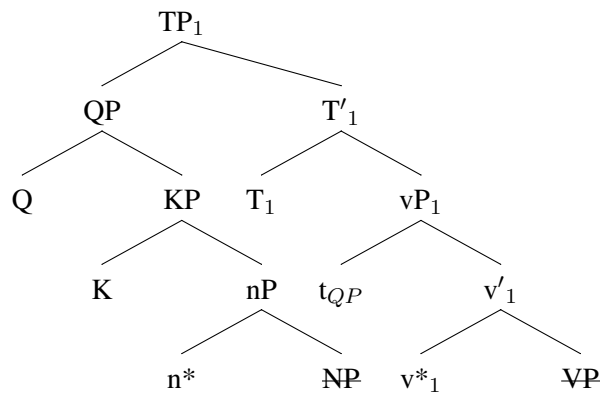
Assume for instance that *n* has {uAsp}, whereas Num has {uAsp, iNum}. Num probes to *n*, and their Asp features match, but since both instantiations of this feature are uninterpretable, no valuation takes place. When Asp is merged along the clausal projection, it comes with {iAsp, uNum}. It probes for Num, at which point the {uAsp} on both Num and *n* is valued and *nP* spells out.

(4) Who did you say left?

a.

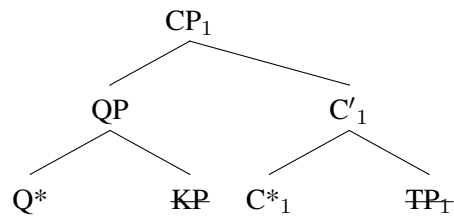


b.



Merger of C, the next phasal head, triggers spell out of TP and checks features of QP, activating it as a phasal head and thereby causing the complement of QP to Spell-Out. At this point, the only feature in the noun phrase that is accessible to further syntactic operations is Q, as all other features are in the spelled-out domain. QP will then move through the edge of the next phase, namely the matrix vP, before reaching its final landing site, the matrix CP.

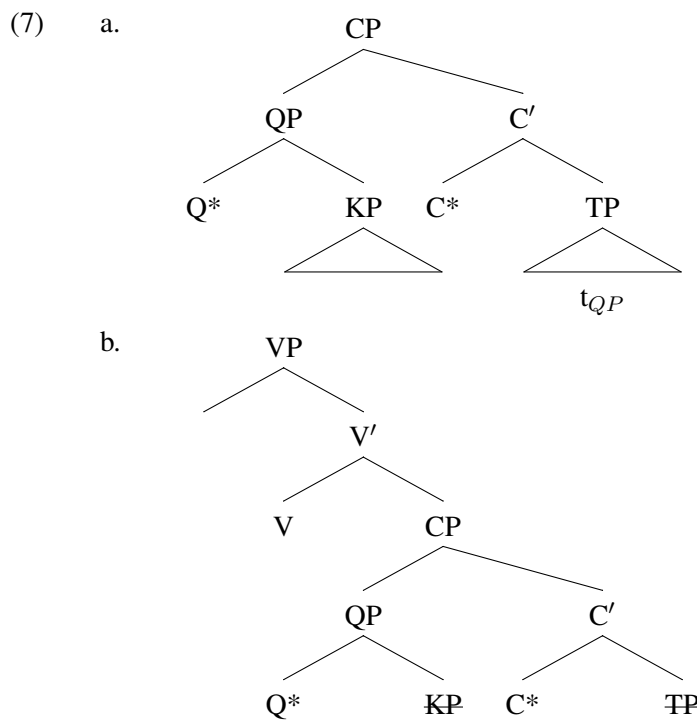
(5)



Consider now how the cases of improper movement such as those in (6) would be ruled out on the proposed model.

- (6) a. *Who seems that won?
b. *Who seems that it won?

Once the embedded CP is merged, QP will move to its specifier (see (7a)). When merged with CP, QP will begin to function as a phasal head, triggering Spell-Out of its complement KP. In the next step, shown in (7b), both TP and KP are spelled out and inaccessible for attraction. As a result, the features on the matrix T, and possibly the K feature of the noun phrase in (6b) will remain unchecked, causing the derivation to crash.



Consider next the cases of superraising such as those in (8).

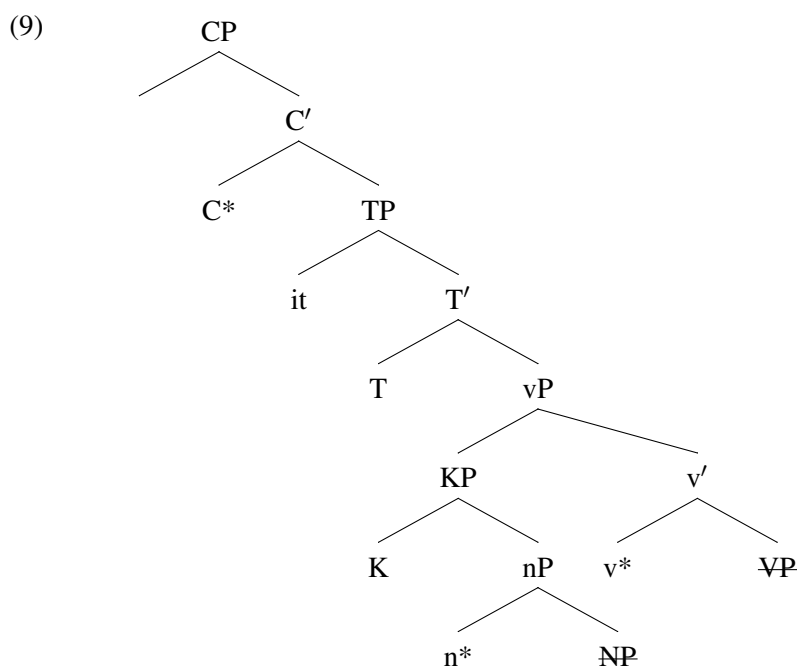
- (8) a. *John seems that it won.
b. *John seems that won.

What makes these examples different from those in (6) is that the moved phrase is a KP, rather than QP. This difference turns out to be quite significant under the current model. There are two derivations to be considered regarding examples such as these. When the embedded C is merged, it will trigger Spell-Out of its complement TP. As a result, everything embedded in the TP will become inaccessible for further syntactic operations. Therefore, if the KP remains embedded in TP, either in Spec of

vP, or in Spec of TP, it will be caught in the Spell-Out domain and thus will not be available for attraction by the matrix T, as desired.

However, there is still a derivation compatible with the current set of assumptions that would derive the ungrammatical examples in (8). If the KP moves to the escape hatch position, the Spec of CP, it will not be spelled out with the complement of the CP phase and could thus escape the embedded clause and target the matrix TP. Note also that with the QP missing, the potential trigger for Spell-Out of KP is also not available. Thus, additional assumptions are required to rule out cases such as (8).

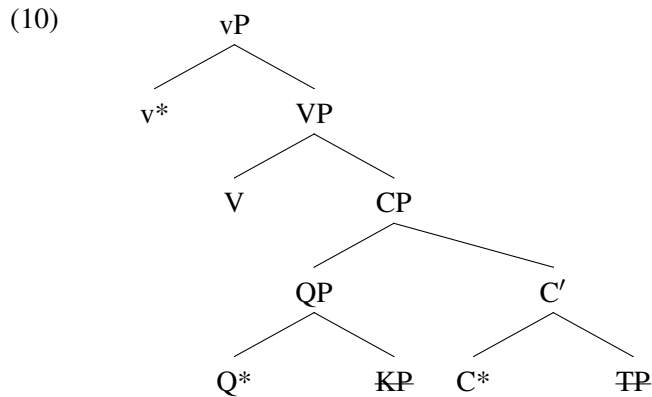
A potential way out of this problem could be to deny KP access to SpecCP. If KP is forced to remain inside the TP, it will inevitably be spelled out as part of the TP when C is merged. Svenonius (2004) suggests that KP might be prohibited from moving to SpecCP in the following way. Assuming strong identification of the two hierarchies, the clausal and their corresponding nominal heads might be too similar in a Relativized Minimality kind of way, giving rise to intervention effects.⁵ Thus, an attractor for Q might not be able to attract QP across the CP boundary, because Q and C are too similar. In the same vein, T might cause an intervention effect for K-movement. In the configuration such as (9), C would not ‘see’ and therefore could not attract a KP embedded in a TP. As a result, KP would not be able to move to CP. Only a phrase with the Q (or a higher) feature would be visible to C. This way movement of KP to SpecCP might be ruled out.



It is clear though that C does not cause an intervention effect for Q-movement when the QP occupies the specifier of CP, as in the configuration below. In this case,

⁵See Starke (2001), Taraldsen (2008) for similar ideas.

QP must be made visible to outside probes, otherwise there could be no long QP movement. In other words, vP must be allowed to attract QP out of the CP specifier.



This might be achieved by defining Attract Closest in such a way that the specifier of CP is equally close to a potential probe as the CP itself. Svenonius (2004) adopts the following definition:

- (11) α enters into agree relations with β iff α and β have features of type f and there is no closer γ with type f features (closer: γ c-commands β but not α)

Given these assumptions, examples like (8a) are successfully ruled out. However, cases like (8b), where KP might move to embedded TP, still remain problematic. If what makes movement of KP to SpecCP impossible is the fact that C does not see K past an intervening TP, then given the definition above, KP will actually be visible to C if it occupies SpecTP and would be able to move to SpecCP on this scenario. This means that examples of superraising as in (8b) must receive a different explanation from cases such as (8a).⁶

Summing up the discussion so far, the proposal outlined in Svenonius (2004) rules out improper movement by assuming a close correspondence between the nominal and clausal functional hierarchy and relying on the phase-based locality. Svenonius assumes that there are at least two phasal heads in the DP, corresponding to C and v in the verbal domain. Nominal phasal heads however are assumed to trigger Spell-Out only after their features have been valued by the clausal phasal heads. Once the complement of the phase head spells out, nothing embedded inside it can be attracted or agreed with. Assuming then that features involved in A-type operations reside in a lower phase than those involved in A'-operations, A-type features will be spelled

⁶(8b) might be ruled out by the Activity Condition (Chomsky (2001, 2000)), according to which XPs that have had all their features valued are rendered *inactive* and cannot enter into further syntactic dependencies. The proposal to be developed in the following sections will rule out all cases of superraising in the same manner, and no appeal to the Activity Condition will be necessary. For arguments against the Activity Condition see for instance Nevins (2004), Bošković (2007).

out early on and will not be accessible to syntactic computation once A'-features are merged. Consequently, there can be no A-movement following an A'-operation.

As we have seen, some additional assumptions are needed to rule out cases of superraising where the raised phrase is not a *wh*-expression (i.e. not a QP). In such configurations, the relevant phrase must be prohibited from moving to the Spec of CP, the escape hatch position, which is achieved by assuming that clausal heads cause intervention effects for extraction of the corresponding nominal projections. This is made to follow from a strong identification of the two hierarchies, in conjunction with a particular definition of Attract Closest. Nevertheless, the derivation where the moving phrase lands in the embedded TP still remains problematic as it leaves open a way of deriving improper movement configurations. To rule such derivations out, additional assumptions seem to be necessary.

5.2 The proposal

5.2.1 Outline of the proposal

In this section, I develop an alternative analysis of improper movement violations, building on and modifying somewhat Svenonius's proposal discussed in the preceding section. In particular, I will also argue that improper movement effects can be derived by relying on the internal structure of the moved phrase, in conjunction with a phase-based locality, though the two accounts will differ considerably in details of implementation. As will be shown, the proposed solution is capable of overcoming the problems that were plaguing the GBOIM-based account.

Like Svenonius's account, the proposal to be sketched also relies on the assumption that there is a close correspondence between the nominal and the clausal functional hierarchy. In other words, I will assume that features on the goal are hierarchically ordered in a way that parallels the ordering of these features in the clausal functional sequence. Given the clausal functional hierarchy $F_1 < \dots < F_n$ (where F_1 is at the bottom of the structural hierarchy and F_n takes F_{n-1} as its complement), if F_2 dominates F_1 , then the matching projection F'_2 in the nominal domain will dominate F'_1 . To illustrate the logic of the proposal, it will be sufficient for now to adopt a rather coarse functional structure, such as the one already employed in the previous section.

- (12) a. C – T – *v* – V
 b. Q – K – *n* – N

If the phase-based locality is to be adopted, then the first question that should be addressed is whether there are any phase heads in the nominal functional sequence. At first blush, it seems that there cannot be any, otherwise a phrase as large as QP could never undergo A-movement, as noted in the previous section. Recall that if Q were a phase on a par with C, then KP would spell out as soon as Q is merged.

This means that at the point when the QP is inserted into the clausal structure, KP would already be spelled-out and inaccessible. As discussed in the previous section, assuming that Q and *n* are phasal heads on a par with their clausal counterparts, Svenonius (2004) suggests that the problem of premature Spell-Out might be overcome if nominal phasal heads trigger Spell-Out of their complements only after being merged with their clausal counterparts. In other words, nominal phasal heads are ‘activated’ in their function as phases only after their features have been checked against the clausal phasal heads. Such an approach is however challenged by examples like (13), which seem to suggest that the phasehood of Q does not hinge on its ability to establish a meaningful feature-checking relation with C.

(13) *_{[CP [Which child_i [_{TP} t_i wonders [_{CP} t_iwho adopted t_i]]]]?}

To see this, consider first the pair of examples in (14).

- (14) a. ?Which child do you wonder who adopted?
 b. *Which parents do you wonder who adopted?⁷

Example (14a) is not perfect as it involves extraction out of a *wh*-island, but it is significantly better than (14b), which in addition involves a Superiority violation.⁸ The interesting thing about (14a) is that it contains two QPs. When the embedded C gets merged, both *wh*-phrases will move to its Spec. The subject *wh*-phrase, *who* in this case, will establish a feature-checking/valuation relation with the embedded C and will be interpreted in this position. Due to the PIC, the second *wh*-phrase will also have to stop off in the Spec of the embedded CP, but it is unclear whether it enters into any feature-checking/valuation relation with C, and even if it does, it would presumably not have all its features checked, making it eligible for further attraction to matrix CP. In any case, the two *wh*-phrases cannot establish the same kind of feature-checking/valuation relation with the embedded C, but crucially both must be ‘activated’ as phasal heads at this point. If the object QP did not start acting as a phase, then its complement KP would not be spelled out at this point and would remain accessible to the probes in the matrix clause. We would then expect the object QP to be able to undergo A-movement in the higher clause, since KP would still be accessible. The sharp ungrammaticality of (13) (which contrasts with the mild deviance of (14a)) shows that this would be an undesirable result. Examples of this

⁷Bad on the reading where it is the parents who adopt someone.

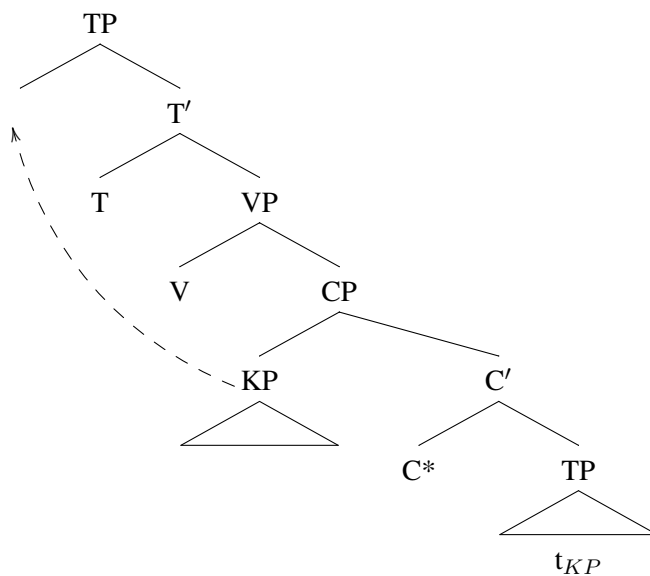
⁸The empirical generalization behind the Superiority Condition is that the highest *wh*-phrase prior to *wh*-fronting must move first to the SpecCP. It goes back to Chomsky (1973), where the pertinent condition is formulated in the following way:

- (i) No rule can involve X, Y in the structure ...X ... [...Z...WYV...] where the rule applies ambiguously to Z and Y, and Z is superior to Y. The category A is superior to the category B if every major category dominating A dominates B as well but not conversely.

type thus suggest that the Spell-Out of KP should not be made dependent on the relation between Q and C.

This conclusion is further corroborated by the fact that a KP occupying a SpecCP is unavailable for further syntactic operations even when there is no QP present, as evidenced by examples of superraising such as (15). In this case the moving phrase is a KP, not a QP. If it is to move out of the embedded clause, the KP must stop off in the Spec of the embedded CP (due to the PIC). Despite the fact that this move places the KP into a phase edge, the KP is not available to probes in the matrix clause and cannot be fronted to the matrix TP.⁹

- (15) a. *John seems that it likes cheese.
b. *John seems that likes cheese.



These observations lead to the conclusion that the Spell-Out of KP is not triggered by Q, but rather directly by C. I will therefore assume that clausal phasal heads can trigger Spell-Out of nominal projections as well, without the mediation of nominal phasal heads. This means that there are no phasal heads in the nominal domain, only clausal heads can ever be phasal. Note that this immediately solves the problem of premature Spell-Out, discussed earlier. Since Q is not a phase head on this view, there is no worry that KP embedded under a Q will Spell-Out even before QP is merged into the clausal structure.

Let me now address the question of how the ability of clausal phasal heads to

⁹Recall that on Svenonius's approach, the fact that KP is not visible to outside probes in (15) is due to the inability of embedded C to attract KP. In other words, KP cannot move to SpecCP in the first place. (15b) however remains problematic, as already discussed. A different way of tackling this issue will be developed presently.

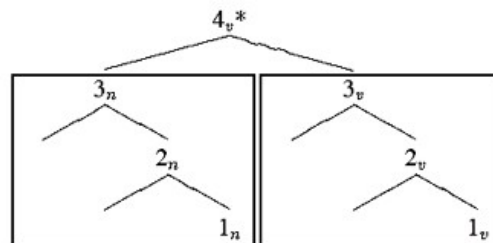
trigger Spell-Out of nominal projections can be implemented. If C is to trigger Spell-Out of a phrase sitting in its specifier, then it is clear that we can no longer hold on to the assumption that it is the complement of the phasal head which spells out, as is standardly done.¹⁰ In addition, we do not want every phrase in the Spec of CP to be spelled out since that would rule out all long movement out of the CP. Relying on strong identification between the nominal and clausal functional sequence, I suggest that Spell-Out is regulated by the following principle:

(16) *The Spell-Out Principle*

Given a functional hierarchy $F_n < F_{n+1} < F_{n+2} < \dots < F_{n+x}$, and F_m a phasal head, then merger of F_m triggers Spell-Out of all features/projections lower on the functional hierarchy than F_m .

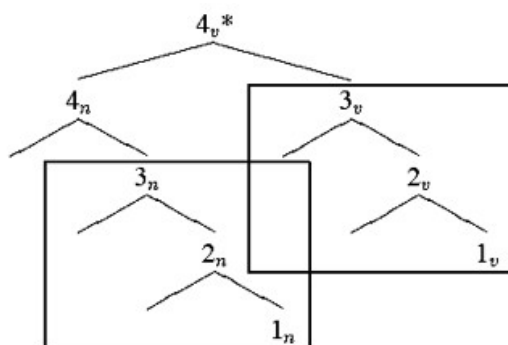
In other words if F_4 , let's say, is a phasal head, then F_1 , F_2 , and F_3 will spell out upon insertion of F_4 , but not F_4 or any higher feature. Assuming that the same kind of features (possibly with different values) comprise both the nominal and clausal functional hierarchy, merger of a phasal head will trigger Spell-Out of both nominal and clausal projections.¹¹ We might schematically represent this as in (17). Phasal heads are marked with an asterisk. When a phasal head enters the structure, it will trigger Spell-Out of all the features lower on the functional hierarchy. Features caught in the spelled out domains (boxed in the diagrams below) are then no longer accessible for further syntactic operations. Note that on these assumptions, the edge of the phase is no longer fully transparent.

(17)



¹⁰See chapter 1 for an outline of the standard approach to phase-based derivations.

¹¹In other words, the assumption here is that the set of functional heads that make up the clausal and the nominal functional hierarchy should be (at least partially) identified. Note also that the proposed Spell-Out Principle is not simply a restatement of the GBOIM. Since the GBOIM-based approach outlined in the previous chapter also relies on phase-based locality, it also requires an assumption regarding what constituent(s) spells out upon merging a phasal head. Following standard practice, I have been assuming in the previous chapter that it is the complement of the phasal head that spells out. The claim made in this chapter is that modifying this assumption along the lines of (16) suffices to derive the ordering restrictions on movement operations, and there is no need to appeal to an independent constraint on movement, such as the GBOIM.



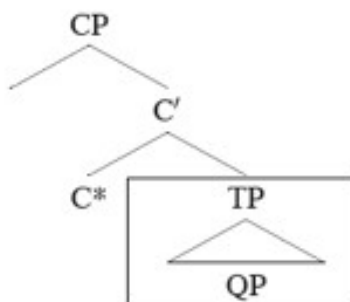
Consider now how this kind of approach tackles cases of superraising, starting with examples involving QPs.

- (18) a. *Who seems that it won?
- b. *Who seems that won?

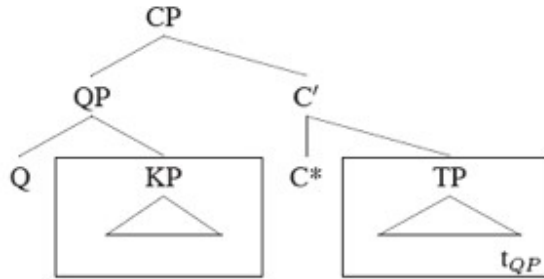
Again the crucial point in the derivation is the merge of the embedded CP, represented in (19). Given that C is a phase head by assumption, all features lower than C on the functional hierarchy will undergo Spell-Out at this point, in accordance with the principle stated in (16). Since T and its nominal counterpart labelled K here are lower on the functional sequence than C, these features will be spelled out and therefore be inaccessible to probes outside the CP. The ungrammaticality of the examples in (18) is thus straightforwardly derived. If the QP, *who* in these examples, remains in the embedded TP, it will be spelled out as part of the TP (see (19a)). If the QP moves to Spec of CP (i.e. the phasal edge), KP embedded inside the QP will still be spelled out, given that K is lower on the functional hierarchy than C (see (19b)). In either case, KP will not be visible to outside probes.

(19)

a.



b.

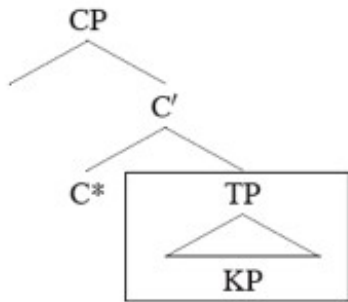


No additional assumptions are required to rule out examples of superraising in which the moved phrase is a KP, rather than QP, as in (20). Regardless of whether the KP *John* remains embedded in the TP, or moves to the escape hatch, the SpecCP, it will be spelled out when C is merged, as it is lower than C in terms of the functional sequence. The two options are illustrated in (21). Since whatever is spelled out can no longer be attracted, KP will not be accessible when the matrix T enters the structure. Consequently, examples like (20) cannot be derived.

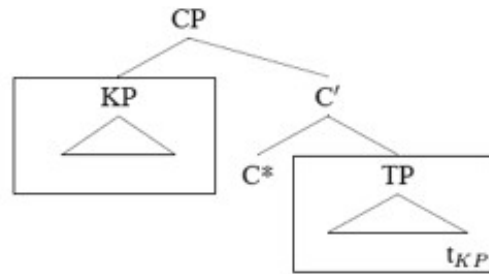
- (20) a. *John seems that it won.
- b. *John seems that won.

(21)

a.



b.



Hence, on the approach developed here, moving to the specifier of a phase head does not necessarily allow a phrase to escape from the Spell-Out domain. This seems to be a welcome result. Recall that on the model sketched in Svenonius (2004), the possibility of escaping the Spell-Out domain by targeting the specifier of a phase head opened up a way of deriving improper movement violations. The strategy that Svenonius adopted to tackle this problem was to deny KP access to the escape hatch position by assuming that clausal projections and their nominal counterparts are too similar in a Relativized Minimality kind of way. Consequently, nominal and clausal projections enter into competition for raising to some functional head, giving rise to intervention effects. Nevertheless, we have seen that even if these additional assumptions are adopted, the ungrammaticality of superraising examples such as (20b) is

unaccounted for. In this case, KP occupies the Spec of TP, from where it can move to the escape hatch position, Spec of CP, without incurring a relativized minimality violation. KP in the Spec of CP will then remain visible to outside probes since only the complement of CP spells out on this view.

The approach argued for here does not face these kinds of problems. Given the definition of the Spell-Out Principle in (16), KP will be spelled out as soon as C gets merged even if it occupies the Spec of CP and will therefore be unavailable for further syntactic manipulation.¹² In other words, a KP will be trapped inside the CP regardless of whether it moves to its Spec or not. On the other hand, a phrase which has an additional Q-feature will remain visible to probes in the higher clause, though only for purposes of Q-attraction, as all features lower than Q will be spelled out at this point. No additional assumptions are therefore needed to handle examples in (20).

5.2.2 Long movement out of a CP

Before turning to improper movement effects in cases involving remnant movement and subextraction out of moved phrases, let me first illustrate how a derivation involving long movement out of a CP proceeds on the current approach.

Recall that long-distance A'-dependencies pose a serious challenge to the GBOIM-based approach. The problem lies in the fact that the dashed movement step in (22), from the Spec of CP to the Spec of higher vP, is required by locality, but prohibited by the GBOIM. Namely, the PIC forces a phrase undergoing long movement to stop off in the Spec of every intermediate C and v/V head, assuming that both C and v/V are phase heads. However, once a phrase moves to the Spec of CP, GBOIM makes all projections that are lower than C on the functional hierarchy unavailable as potential landing sites. A phrase undergoing movement should thus be trapped in the embedded CP which will be spelled out as soon as the matrix v is merged, and consequently no movement out of the CP should be possible.

(22)

[_{CP}XP [C ... [_{vP} <XP> [v ... [_{CP} <XP> [C ... [_{vP} <XP> [_v [_{VP} V <XP>]]]]]]]]]

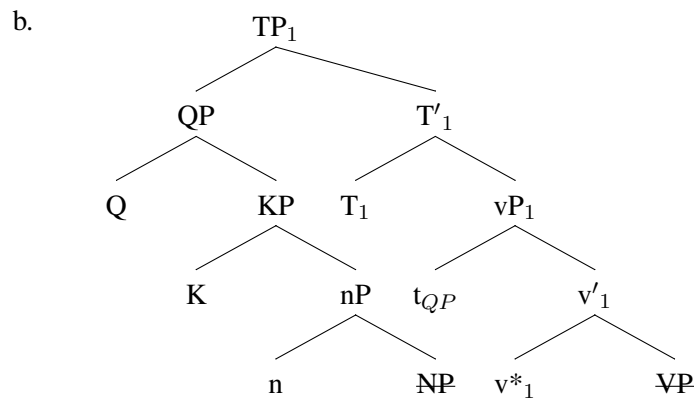
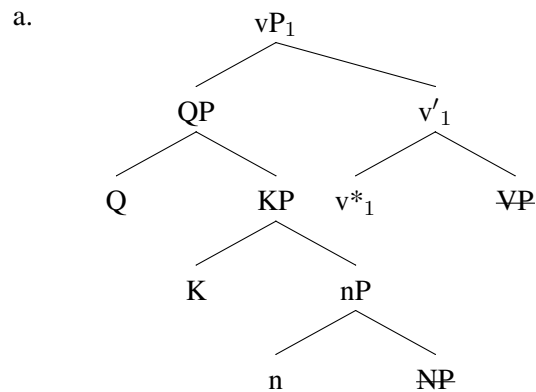
The analysis outlined in this chapter does not face this kind of problem. To see this, let us go step by step through a derivation involving long-distance *wh*-movement. Consider how an example like (23) might be derived. The tree in (23a) illustrates the point in the derivation when the embedded vP enters the structure.¹³ Since v is a phase

¹²Note that some authors assume that when a phrase spells out its label nevertheless remains accessible. I assume here that whatever is spelled out is opaque and can no longer be attracted or agreed with. See chapter 1 for further clarification.

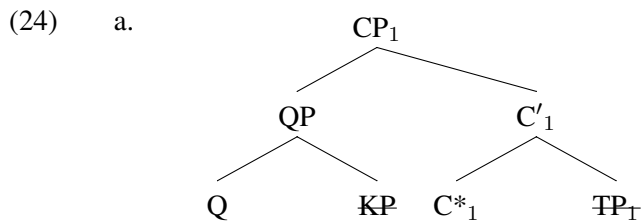
¹³Phase heads are marked with an asterisk, the spelled-out domains are marked with a strikethrough.

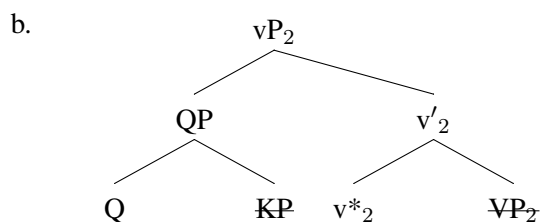
head, it triggers Spell-Out of all projections/features lower than v on the functional hierarchy. Thus, in accordance with (16), VP and its nominal counterpart will be spelled out at this point. In the following step, represented in (23b), QP moves to the Spec of TP to value the K feature.

(23) Who did you say left?



Merging of C, the next phasal head, triggers Spell-Out of both TP and KP. As a result, the only feature on the goal that is available for further syntactic operations is Q, as all other features are in the spelled-out domain. When the matrix v is merged in (24b), the QP will move to its Spec.





Note that there is no problem with the last movement step on the model sketched here. In principle any phrase can move to the Spec of the phase head, though some features on the goal become unavailable in the course of the derivation. Thus, while both Q- and K-features are still accessible on the phrase sitting in the Spec of embedded vP, a phrase which has reached the Spec of matrix *v* via the embedded CP only has the Q-feature still available for attraction. In the final step, QP moves from SpecvP to the Spec of matrix CP, where it is interpreted. The derivation converges with no problems encountered. A major stumbling block of the GBOIM-based approach is thus avoided, without losing the account of improper movement effects.

5.2.3 Non-identity cases

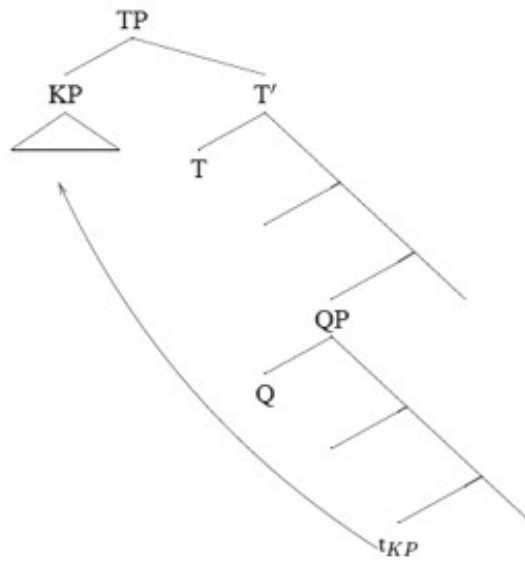
Consider now what predictions the approach developed here makes for cases of improper movement in configurations involving subextraction out of moved phrases and remnant movement. In order to illustrate the logic of the proposal, I will focus my attention for now on the interaction between K- and Q-movement, i.e. what would be the canonical instantiations of A- and A'-operations on an approach that assumes such a distinction.

Recall that K-movement (i.e. movement to a case position) can precede Q-movement (i.e. *wh*-movement), but that the reverse feeding relation is disallowed. That this is true for cases involving remnant movement can be shown on the basis of examples like (25), reproduced from chapter 2.

- (25) a. [How likely t_i to win the race] $_j$ is John $_i$ t_j ?
 b. *[A picture of t_i] $_j$ is known [which king] $_i$ to have been sold t_j .

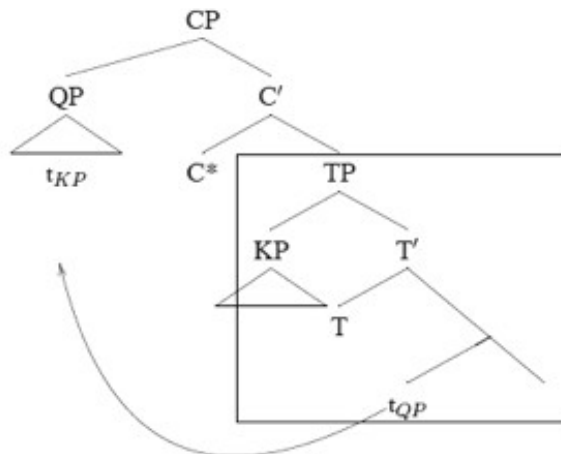
The licit sequencing of KP- and QP-movement in remnant movement configurations is schematized below. First, the TP attracts the KP from inside the QP. Recall that QP is not a phase on the approach developed here. Thus, KP will not be spelled out before it has a chance to be attracted by T.

(26)



When C gets merged, it will attract the remnant QP and trigger the Spell-Out of TP. Given our current set of assumptions, the derivation sketched below proceeds without problems.

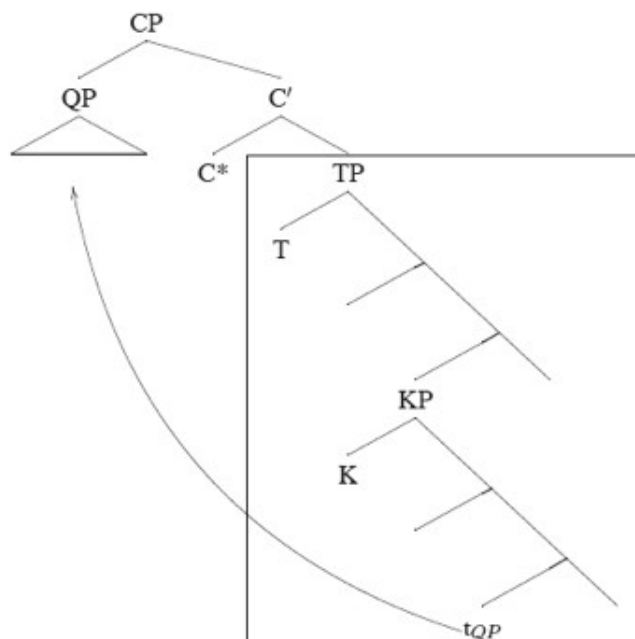
(27)



The illicit reverse ordering of these movement operations is straightforwardly ruled out. The relevant derivation is sketched below. It plays no role on the current approach whether KP moves to SpecTP or not. When the embedded CP is merged, it will attract QP from inside the KP. Since C is a phase head, all features lower than C on the functional hierarchy will undergo Spell-Out. This means that TP will be spelled out, together with the remnant KP embedded inside it. As a result, remnant

K-movement cannot follow Q-movement. Note that even if KP were to move to the Spec of CP, rather than remaining within the TP, it would nevertheless be spelled out, in accordance with the Spell-Out Principle in (16), and as already discussed in the previous section.

(28)

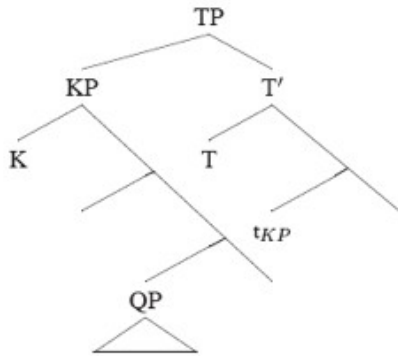


Let us now turn to cases of subextraction from moved phrases which remained unaccounted for under the GBOIM-based approach. The situation is different under the approach argued for here. Again, the licit ordering whereby K-movement feeds Q-movement is allowed, while the illicit one where K-movement follows Q-movement is ruled out.

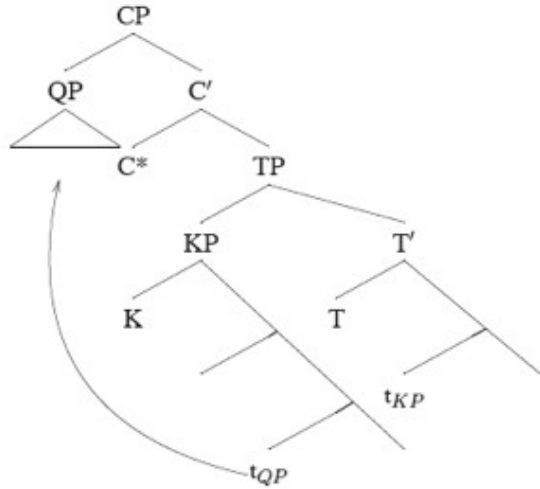
The derivation involving licit sequencing of these operations is sketched below. The KP will first move to TP (or AgrOP), carrying the QP along. Subsequently, C will attract QP from inside the KP. Only at this point, C will cause TP to spell out, making everything inside it unavailable for further attraction.

(29)

a.

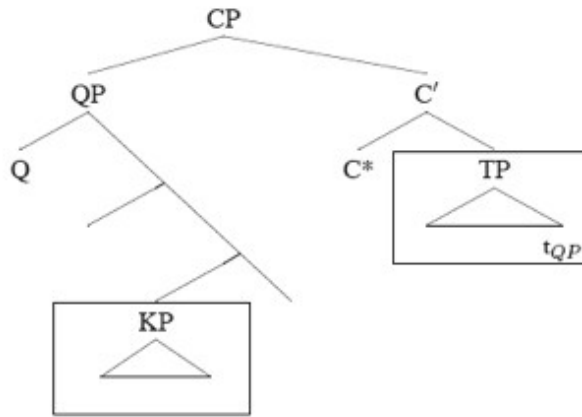


b.



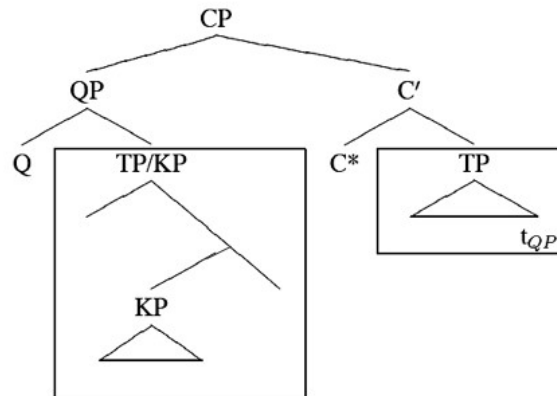
Consider now the reverse ordering of these two operations. (30) illustrates the point in the derivation when the embedded C is merged. QP will then move to the Spec of CP, dragging the KP along. C will then cause all features lower on the functional hierarchy to Spell-Out, which means that both TP and KP will be spelled out at this point. As a result, KP will no longer be visible for attraction at the point when a potential probe in the higher clause is merged.

(30)



In fact, a more likely scenario is that the KP in question is contained in another KP/TP, as illustrated below. A relevant example of this type is repeated in (31). Again, no K-movement out of the CP is possible.

(31) *Oscar was asked how likely to win it was.



As illustrated in this section, the approach argued for here accounts for improper movement violations in both types of non-identity configurations. What is more, no additional stipulations need to be made to capture these cases other than those required for identity cases.

So far, I have been focusing on the classic examples involving interaction between KP and QP movement (i.e. the paradigm cases of A- and A'-movement operations). In the rest of this chapter, I will expand the inventory of movement types beyond K- and Q-movement and explore how various movement operations are predicted to interact on the current model. I will first focus on scrambling and show that like an approach based on the Hierarchy of Operations, the analysis of improper movement developed here predicts that middle field scrambling (in German) should precede *wh*-movement/topicalization. I will then turn to cases where the two types of approaches, namely the one relying on the Hierarchy of Operations and the one developed here, make different predictions, investigating more closely the combination of scrambling and A-movement, and the combination of topicalization and *wh*-movement.

5.2.4 Middle field scrambling

Abels (2008b) argues that scrambling should be sandwiched between *wh*-movement and A-movement on the Hierarchy of Operations, restricting the term scrambling in this context to mean movement to a pre-subject position in a language like German.¹⁴ For the sake of convenience, the hierarchy proposed by Abels (2008a) is given again in (32).

(32) A-mvt \prec (middle field) scrambling \prec *wh*-mvt \prec topicalization

¹⁴For a somewhat different implementation, see Grewendorf (2003), as well as chapter 2 of this dissertation for discussion and comparison of Abels's and Grewendorf's proposal.

(32) states that middle field scrambling follows A-movement, but precedes *wh*-movement and topicalization. Let us now focus on this movement operation and investigate what locality restrictions the current model imposes on its application.

Consider first clause-boundedness of middle field scrambling in German. As already noted, scrambling in German cannot cross a CP-boundary (from Abels (2008b)).

- (33) *Gestern hat [jeden Studenten]_i Hans gesagt, dass seine Eltern t_i
 yesterday has every student.ACC Hans said that his parents
 unterstützen.
 supported

In this respect, the behaviour of scrambling in German parallels that of A-movement (or K-movement on our terms). Recall that on the standard view, the clause-boundedness of A-movement is captured in terms of improper movement. The derivation of examples like (34) involves a step of A-movement from Spec of CP, an A'-position.

- (34) a. *John seems that (it) likes cheese.
 b. *Who seems that (it) likes cheese.

We have seen in section 5.2.1 how examples of this type are handled under the current approach. The proposed analysis can easily be extended to examples involving scrambling. First of all, we need to slightly expand the clausal structure we have been operating with so far. Let us therefore assume that the scrambled phrase moves to the specifier of a functional head on top of TP, as we did on the GBOIM-based approach.¹⁵ The expanded hierarchy of functional projections is given in (35). Assuming the parallel arrangement of features in the nominal domain yields (35b).¹⁶

- (35) a. C – F_{scr} – T – v – V
 b. Q – F_{scr} – K – n – N

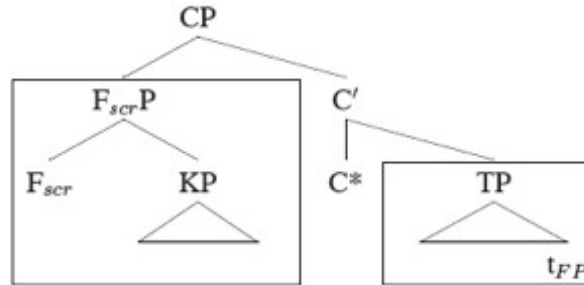
Given this arrangement of features, the observation that scrambling in German cannot cross the CP (i.e. is clause-bound) is derived in a way completely parallel to the case of K-movement. If a phrase bearing the scrambling feature remains inside the TP after the embedded C is merged, it will be spelled out together with the TP. Moving to the Spec of CP does not enable the phrase in question to escape the embedded CP. This is because F_{scr}P will be spelled out when C is merged, given that F_{scr} is lower on the functional hierarchy than C, as dictated by the Spell-Out Principle. Therefore, as

¹⁵There is another scrambling operation in German, often referred to as VP-scrambling, which places the moved phrase in a position following the subject. Though I focus here on scrambling to a pre-subject position, it should become clear as the analysis unravels that both types of scrambling are predicted to be clause-bound on the model argued for here. I will discuss some locality restrictions on VP-scrambling in the following section.

¹⁶I assume here that scrambling is a feature-driven movement operation, though I will remain agnostic about the exact nature of this feature. See for instance Sauerland (1996), Grewendorf and Sabel (1999), Müller (1998) for arguments in support of this view.

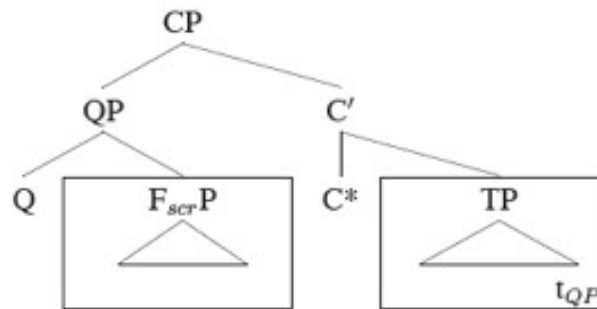
soon as C gets merged, $F_{scr}P$ becomes unavailable for further syntactic manipulation.

(36)



It should be clear by now that embedding an $F_{scr}P$ under a QP would not make long distance scrambling possible. Although a phrase with a Q-feature could potentially move to the higher clause, it could not do so by scrambling since the scrambling feature is spelled out inside the embedded CP (as illustrated below) and is therefore not visible to the probes in the upper clause.

(37)



Therefore, *wh*-movement could not be followed by a scrambling step that would place the moving phrase in the matrix middle field position. The factual correctness of this prediction is supported by ungrammaticality of examples like (38), repeated from chapter 2 and taken from Abels (2007).

- (38) *Gestern hat [welchen Studenten] nur ein Meisterdetektiv wissen
 yesterday has which student only a master-detective know
 können, Hans beim Schummeln erwischt hat.
 could Hans at.the cheating caught has
 'Yesterday only a world-class detective could have known which student
 Hans caught cheating.'

The same results are replicated for the interaction between scrambling and *wh*-movement in cases involving subextraction and remnant movement. The observation

that scrambling necessarily precedes *wh*-movement in both of these configurations is straightforwardly derived on the current model.¹⁷ The logic of the solution is the same as already illustrated for the case of *K*-movement, with no additional stipulations required.¹⁸ The proposed account therefore straightforwardly captures the following observations: (i) that a phrase with a scrambling feature cannot escape out of a CP, and (ii) that *wh*-movement must strictly follow scrambling. The interaction between scrambling and case-driven movement will be discussed in the following section.

5.2.5 Symmetric feeding relations

We have seen thus far how the proposal developed here accounts for the asymmetric ordering relations between case movement and scrambling on the one hand, and *wh*-movement (and topicalization) on the other. The observed ordering restrictions are a consequence of there being a phase boundary between the positions targeted by case movement and scrambling and those targeted by *wh*-movement and topicalization. Note that the presence of a phase boundary is crucial. If two operations target positions which are not separated by a phase boundary, no ordering is imposed on them in cases of cross-clausal movement. To illustrate this more clearly, let us assume a functional hierarchy as given in (39).

(39) [XP X [YP Y [ZP Z [WP W]]]]

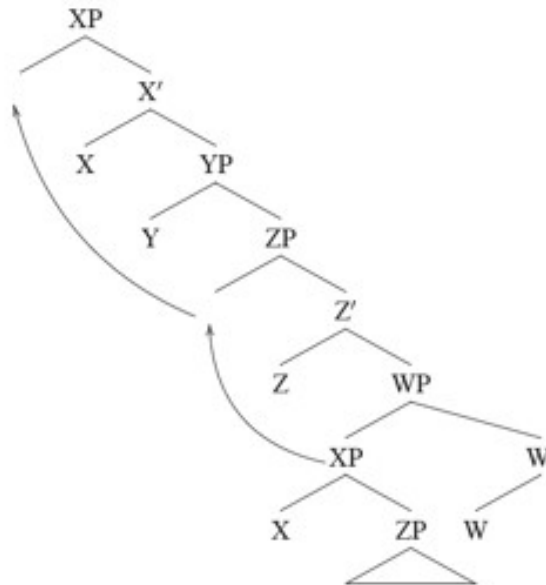
For clause-internal movement, we predict that movement to ZP must precede movement to XP regardless of the presence or absence of phase boundaries. The standard assumption that movement is always to a c-commanding position rules out the reverse ordering. The predicted feeding relations are given in (40), and the licit sequencing of the two operations in question is schematized in (41).

- (40) a. *X-mvt \prec Z-mvt (no lowering)
 b. Z-mvt \prec X-mvt

¹⁷For relevant examples illustrating this observation I refer the reader back to chapter 2.

¹⁸The same predictions are made for the interaction between scrambling and topicalization, assuming that topicalization also targets a position in the C-domain.

(41)

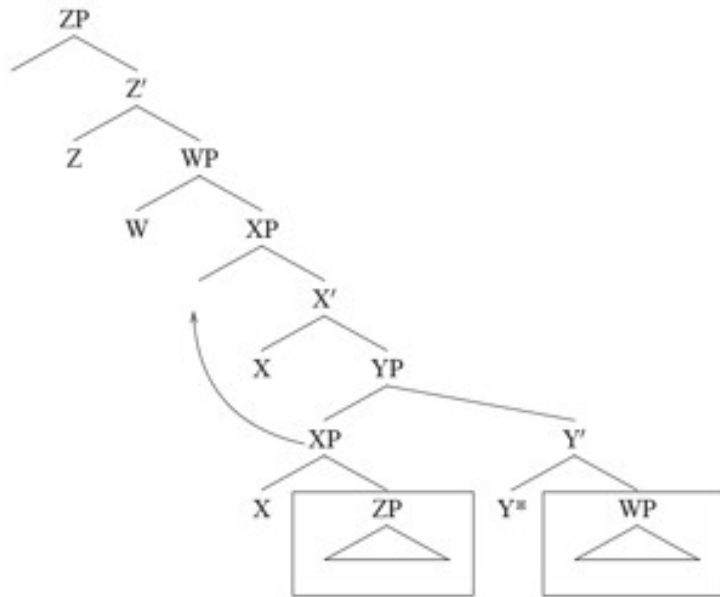


Consider now a configuration involving embedding. If Y is a phasal head, then we predict that Z-movement should precede X-movement also in cases of cross-clausal movement. This is because ZP will be spelled out as soon as Y is merged, given that Z is lower on the functional sequence than Y. X will remain accessible and could be attracted, but crucially XP movement cannot be followed by ZP-movement. This is schematized in (43).¹⁹

- (42) a. *X-mvt \prec Z-mvt
 b. Z-mvt \prec X-mvt

¹⁹Whether W is a phasal head is immaterial on this view, since both Z and X are higher in terms of the functional hierarchy than W, and will therefore remain accessible after merger of WP.

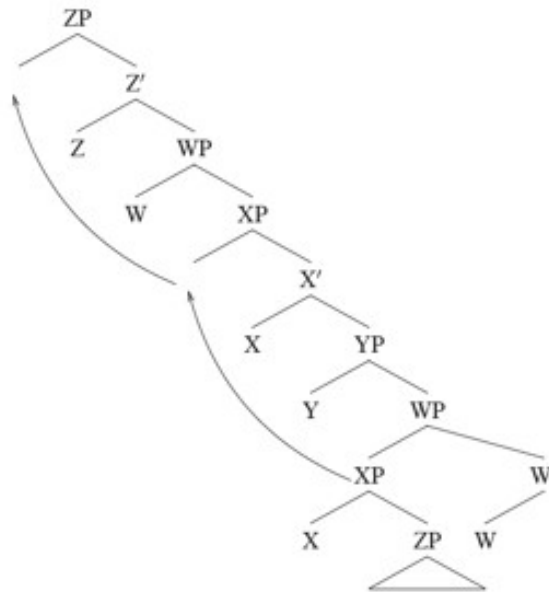
(43)



However, if Y is not a phasal head, neither XP nor ZP will be spelled out when Y is merged. As a result, both XP and ZP will remain accessible to the probes in the higher clause, and no ordering on the application of these two operations is imposed. In particular, the ordering whereby X-movement precedes Z-movement, which was ruled out on the previous two scenarios, is allowed. This is illustrated in (45). Note also that movement is always to a c-commanding position, as required.

- (44) a. X-mvt \prec Z-mvt
- b. Z-mvt \prec X-mvt

(45)



Therefore, the account developed here does not always impose asymmetric ordering relations on movement operations in cases of cross-clausal movement. Under certain well-defined conditions, symmetric feeding relations between movement operations are allowed.

Contrast the predictions of the current model to the one endorsing a Hierarchy of Operations. The proponents of this view argue that, at least for the subset of operations they consider, the ordering is always asymmetric. An example of such a hierarchy is repeated below (taken from Abels (2008b)):

(46) A-mvt \prec (middle field) scrambling \prec *wh*-mvt \prec topicalization

It is clear that even confining our attention to the operations listed in (46), the approach argued for here does not always impose asymmetric feeding relations. The potential points of disagreement concern the interaction between A-movement (i.e. case movement) and scrambling, as well as between *wh*-movement and topicalization. Let me discuss each of these in turn.

5.2.5.1 Combining scrambling and A-movement

Consider first the interaction between A-movement and middle field scrambling in German. Both A-movement and scrambling in German target a domain between the C and ν phase boundaries. Recall that what was meant by A-movement for the purposes of the Hierarchy on Abels's (2008b) approach is movement to a case position, while scrambling was defined as movement to a pre-subject position in German. Given that the landing site of middle field scrambling c-commands all case positions,

we expect that this type of scrambling should not be able to feed A-movement within a single clausal domain. In other words, both an approach relying on the Hierarchy of Operations and the one pursued here predict an asymmetric feeding relation between A-movement and middle field scrambling in cases of clause-internal movement. This prediction is borne out. The relevant examples can be found in chapter 2.

The situation is different however when we turn to cases involving cross-clausal movement. An approach relying on the Hierarchy of Operations still states that there should be an asymmetric feeding relation between A-movement and scrambling, but the approach pursued here does not make the same prediction. In principle, as long as the embedded clause is smaller than a CP, scrambling should be able to feed A-movement, at least if by A-movement we mean fronting of the subject to a case position in the middle field (standardly, SpecTP).²⁰ We know that both scrambling and A-movement are possible out of certain infinitival clauses, such as those selected by the verb *versuchen* ‘try’.²¹ In these contexts, the current proposal predicts that there should be no ordering on the application of these two movement operations. However, as already noted in chapter 2, this prediction is difficult to test. Recall that Grewendorf (2003) provides the following examples to show that scrambling cannot feed A-movement. In (47), the pronoun scrambles out of the infinitive to a position in the matrix clause. Subsequently, the infinitival clause undergoes A-movement.

- (47) a. *weil [t_j zu küssen]_i ihn_j von Maria vergeblich t_i versucht wurde.
 since to kiss him by Mary in.vain tried was
 b. *weil [t_j zu küssen]_i von Maria ihn_j vergeblich t_i versucht wurde.
 since to kiss by Mary him in.vain tried was

However, I have already noted that there are several observations that cast doubt on the relevance of (47) for the issue at stake. Let me repeat these briefly. First of all, the scrambled phrase ends up in a position in the matrix clause which is lower than the subject position. If the scrambling position is higher than the subject position, then examples like (47) are expected to be ungrammatical. Secondly, A-movement in passive constructions in German is not obligatory, as illustrated in (48) (from Müller (1995)).

- (48) a. daß [IP [vP dem Peter das Buch gegeben] wurde]
 that the.DAT Peter the.NOM book given was

²⁰A situation might be different regarding movement to an object case position. If this position is below the vP phase boundary, or if object case is assigned in the base position, then we again expect middle field scrambling to strictly follow case-assignment. In fact, in the following chapter, I will argue on the basis of the binding data that the relevant phasal heads are C and Asp, rather than C and v/V. If so, then there might be even a case position for an internal argument above the base-generated position of the subject but still within the domain of the lower phase, again forcing middle field scrambling to strictly follow object case-assignment.

²¹Refer back to chapter 2 for relevant examples.

- b. daß [_{IP} das Buch_i [_{VP} dem Peter t_i gegeben] wurde]
 that the.NOM book the.DAT Peter given was

Hence, examples such as (47) might in fact involve two instances of scrambling, namely scrambling of the pronoun, followed by scrambling of the remnant infinitive. It is well-known that a phrase cannot be scrambled if another phrase has scrambled out of it (see Müller (1998) for a list of relevant references). The ungrammaticality of (47) might then be due to the same restriction. Consequently, it is questionable whether examples like (47) can tell us anything about a possible feeding relation between A-movement and scrambling. In fact, it seems to be extremely difficult to construct examples that would either prove or disprove such a claim. Therefore, although the approach pursued here makes a prediction different from the one that an approach relying on the Hierarchy of Operations does, it seems to me that no compelling evidence for either approach can, or at least has been offered by now.

Note that there is another scrambling operation in German, referred to often in the literature as VP-scrambling, which targets a position lower than the surface subject position. This scrambling operation inverts the order of the direct and indirect object, and displays certain properties which are distinct from middle field scrambling. For instance, the scrambled accusative can bind a reflexive from the derived position, as (49a) illustrates. Middle field scrambling, on the other hand, does not create new binding possibilities (see also chapter 3).²²

- (49) a. dass der Arzt_j den Patienten_i sich_{i/j} t_i im Spiegel
 that the.NOM doctor the.ACC patient self in.the mirror
 zeigte.
 showed
- b. *weil den Studenten_i [die Lehrer von sich_i] zweifellos t_i in guter
 since the student.ACC the teachers of self undoubtedly in good
 Erinnerung behalten haben.
 memory kept have
 intended: 'The teachers of himself have undoubtedly kept the student
 in good memory.'

The approach pursued here predicts that this type of scrambling should be able to feed case movement even within a single clausal domain.²³ Again, it is difficult to confirm the validity of this prediction. However, that such a feeding relation is in principle

²²(49a) is from Müller (1995), (49b) is from Grewendorf and Sabel (1999).

²³The proposal does not necessarily predict that the reverse feeding relations should be possible, i.e. that case movement (to SpecTP) should feed this type of scrambling in cases of cross-clausal movement (clause-internal instances of the corresponding feeding relation are independently ruled out by the Extension condition). If this type of scrambling targets a position in the lower phase (vP or Asp phase), then the relevant feature will be spelled out too early to be able to follow an operation targeting a position in the middle field, such as case movement to SpecTP.

possible can be more easily demonstrated on the basis of the corresponding Icelandic examples. Icelandic permits passivization of a direct object only with a subset of triadic verbs. Thus, the verb *gefa* ‘give’ allows its direct object to be passivized, but not a verb like *skila* ‘return’. The contrast is illustrated below. Note that the indirect object can be passivized in both cases (examples from Thráinsson (2007)).

- (50) a. (i) Konunginum var gefin ambáttina.
the.king were given maidservant
‘The king was given the maidservant.’
(ii) Ambáttin var gefin konunginum.
the.maidservant was given the.king
‘The maidservant was given to the king.’
b. (i) Foreldrunum var skilað börnunum.
the.parents was returned the.kids
‘The parents were returned the kids.’
(ii) *Börnunum var skilað foreldrunum.
the.kids was returned the.parents
out in the sense: ‘The parents were returned the kids.’

The same class of verbs that allows direct object passivization also allows the direct and the indirect object to invert their order. Thus, verbs of the *gafa*-class permit both the (base) order IO DO, and the (inverted) order DO IO. With verbs of the *skila*-class, there is no inversion construction. Thus, (51aii) is ungrammatical on the reading where the kids were returned to the parents (and is only acceptable if it is the parents that were returned to the kids) (from Thráinsson (2007)).

- (51) a. (i) Hann gaf konunginum ambáttina.
he gave the.king.DAT the.maidservant.ACC
‘He gave the king the maidservant.’
(ii) Hann gaf ambáttina konunginum.
he gave the.maidservant.ACC the.king.DAT
‘He gave the maidservant to the king.’
(i) Mannræninginn skilaði foreldrunum börnunum.
the.kidnapper returned the.parents.DAT the.kids.DAT
‘The kidnapper returned the kids to the parents.’
(ii) *Mannræninginn skilaði börnunum foreldrunum.
the.kidnapper returned the.kids.DAT the.parents.DAT

Holmberg and Platzack (1995) (following earlier suggestions by Falk (1990)) argue that the two observations are related and that the passive which promotes the direct object originates from the inversion construction. This conclusion is further corroborated by the fact that constructions involving direct object passivization such as (50aii) display the same properties as the inversion constructions. For instance, both configurations require the indirect object to be focused, while examples where indirect objects are passivized are acceptable regardless of where the focus is placed.

One way to interpret these observations is to assume that a direct object can be passivized only if it has scrambled over the indirect object (see for instance Ottósson (1993)).²⁴ Constructions such as these thus might be used to argue that the scrambling operation which inverts the order of internal arguments can in principle feed case movement, as predicted on the approach pursued here.²⁵

Let me point out here that VP-scrambling has not been discussed by the proponents of approaches appealing to a Hierarchy of Operations. The fact that VP-scrambling might feed A-movement (i.e. case movement) is not necessarily a problem for such approaches though. Recall that on Abels's (2008b) view for instance, the placement of scrambling with respect to case movement indicated in (46) holds only of middle field scrambling. It is plausible to assume that VP-scrambling is an operation distinct from middle field scrambling and occupies a different position in the Hierarchy of Operations, though again the exact conditions that determine what counts as an instance of a particular operation on such an approach remain to be clarified.

5.2.5.2 Combining *wh*-movement and topicalization

Another pair of interactions where the approach developed here predicts a symmetric feeding relation to be possible concerns *wh*-movement and topicalization. Both movement operations target a projection in the (split) CP domain. If there is one phase boundary in the left periphery, C for instance, and both *wh*-movement and topicalization target a projection above this phrase boundary, then we again expect no ordering between these two operations to be imposed in cases of cross-clausal movement.²⁶ This prediction conflicts with Abels's (2008b) claim that the two operations

²⁴This is not the conclusion that Holmberg and Platzack (1995) reach though. They assume that both possible orders with *gefa*-type verbs are base generated. For German, the standard assumption has been that the inverted construction is derived by scrambling (e.g. Webelhuth (1989), Müller (1995), etc.). Ottósson (1993) who argues that inverted order in Icelandic is also derived by movement, shows that the conditions on inversion in Icelandic and German are essentially the same, suggesting a unified analysis of inversion in these two languages. The discussion in the main text presupposes a movement analysis of inversion constructions.

²⁵The same conclusion might be drawn on the basis of object shift data. Only verbs which permit the inverted order allow the direct object to shift over the in-situ indirect object. With verbs of the *skila*-class, which don't have the inverted alternative, this is prohibited. Collins and Thráinsson (1996) thus argue that the direct object undergoes object shift from the inverted position. Assuming then, as Collins and Thráinsson (1996) do, that object shift in Icelandic is A-movement to a case-assigning position, such facts suggest again that inversion can feed case movement. I refer the reader to Collins and Thráinsson (1996) for further discussion and relevant examples.

²⁶Rizzi (1997) has argued that the CP should be split into a sequence of hierarchically ordered functional projections. If this view of the left periphery is adopted, the question arises which one of the Rizzian functional projections constitutes a phase head. For now, I will assume that the lowest head in the left periphery is the phasal head. An interesting

in question are asymmetrically ordered as given in (52).

(52) *wh*-movement < topicalization

Let us scrutinize this claim more closely, focusing first on cases of remnant movement in German. Examples such as those below seem to lend support to the ordering given in (52) (from Grewendorf (2003)).²⁷

- (53) a. ??[*t_i* zu überreden]_{*j*} weiß ich nicht [*wen_i* sie versucht *t_j* hat]
 to persuade know I not who.ACC she tried has
 ‘I don’t know whom she tried to persuade.’
 b. ??[Ein Buch *t_i*]_{*j*} weiß ich wohl [*worüber_i* ich *t_j* schreiben sollte]
 a book know I well about-what I write should
 ‘I know well about what I should write a book.’

The situation is however a bit more complex. Observe that extraction in (53a) takes place out of a coherent infinitive, the type of infinitive out of which scrambling is allowed. If we adopt Abels’s claim that *wh*-phrases in German in principle can scramble, then the *wh*-phrase in (53a) might have actually scrambled out of the infinitive and then undergone a step of *wh*-movement. Similarly, in (53b), the *wh*-phrase might have first scrambled out of the DP, and then subsequently *wh*-moved to the embedded C-domain. In this case, the examples in (53) would show that topicalization can affect a remnant XP out of which another phrase has been scrambled. In other words, these examples might be taken to show that scrambling can feed topicalization, not that *wh*-movement can feed topicalization. This kind of conclusion seems to be supported by examples like those in (54). Recall that scrambling cannot cross the finite clause boundary in German. We know therefore that *wh*-phrases in (54) could not have been scrambled out of the embedded clause, but must have undergone a step of *wh*-movement. Interestingly, when the remnant is subsequently topicalized, the result is ill-formed. Both examples in (54) involve extraction out of a *wh*-island, but it has been argued that they contrast sharply with their counterparts in (55), where no remnant movement takes place (see Müller (1998), Hinterhölzl (2006)).

- (54) a. *[Daß Fritz *t_i* liebt]_{*j*} weiß ich nicht [*wen_i* er *t_j* gesagt hat]
 that Fritz.NOM loves know I not who.ACC he said has
 b. *[Gesagt daß Fritz *t_i* liebt]_{*j*} weiß ich nicht [*wen_i* er *t_j* hat]
 said that Fritz.NOM loves know I not who.ACC he has

alternative to pursue would be to assume that each head in the left periphery constitutes a phase (for an argument along these lines see for instance van Craenenbroeck and van Koppen (2007)). On the latter view, various movement operations targeting the left periphery would in fact be ordered.

²⁷Note that all the examples to be discussed are quite degraded, so the conclusions are based on contrasts in grammaticality.

- (55) a. ??[Daß Fritz Caroline liebt]_i weiß ich nicht wer t_i gesagt
that Fritz.NOM Caroline.ACC loves know I not who said
hat.
has
- b. ??[Gesagt daß Fritz Caroline liebt]_i weiß ich nicht [ob
said that Fritz.NOM Caroline.ACC loves know I not whether
er t_i hat]
he has

Müller (1998) however reports that there is a contrast in acceptability between examples in (54) and the one in (56a). In (56a), extraction takes place out of an incoherent infinitive. Scrambling out of this type of infinitival clause is impossible. In this respect, incoherent infinitives pattern with embedded finite clause. We know therefore that the *wh*-phrase, *who*, in (56a) could not have scrambled out of the infinitive, but must have been *wh*-moved out of it. Thus, as in (54), the topicalized phrase in (56a) should contain a trace of *wh*-movement. Interestingly however, Müller (1998) reports that topicalization of the remnant infinitive in cases like (56a) does not exhibit the severe deviance that can be observed with analogous cases involving finite clauses in (54), and is on a par with the corresponding sentence in (56b) that do not involve a trace in the topicalized category.

- (56) a. ??[PRO t_i zu lesen]_j weiß ich nicht was_i der Fritz t_j abgelehnt hat.
to read know I not what the Fritz refused has
- b. ??[Das Buch zu lesen]_j weiß ich nicht warum der Fritz t_j abgelehnt
the book to read know I not why the Fritz refused
hat.
has

If examples such as (56a) are indeed acceptable, then we are led to the conclusion that *wh*-movement can indeed feed remnant topicalization. The severe deviance of (54) cannot then be due to the interaction between topicalization and *wh*-movement, but must be attributed to an independent factor.²⁸

Abels (2008b) uses the pattern in (57) to further argue that the hierarchical ordering of *wh*-movement and topicalization is as stated in (52).

²⁸Recall that Grewendorf (2003) treats short *wh*-movement in German and long *wh*-movement as movements of different type, targeting different landing sites. According to him, short *wh*-movement is a non-operator movement as it does not show Weak Crossover effects, while long *wh*-movement is operator movement. He then takes examples in (53) as showing that short *wh*-movement can feed topicalization. Though he does not discuss cases such as (56a), his account predicts that they should be illicit. If the facts are as Müller (1998) reports, then these examples present a clear challenge to Grewendorf's analysis.

- (57) a. ??[Über Karl den Großen]_i weiß ich nicht, [was für ein Buch t_i]_k er t_k
 about Charlemagne know I not what for a book he
 schreiben will.
 write wants
 ‘About Charlemagne I don’t know what kind of book he wants to
 write.’
- b. *[Über welchen deutschen Kaiser]_i sagt er [ein fertiges
 about which German emperor says he a completed
 Manuskript t_i]_j hat keiner t_j anzubieten?
 manuscript has nobody to.offer
 ‘About which German emperor does he say that a completed manuscript
 nobody can offer?’

According to Abels (2008b), (57a) which involves topicalization out of *wh*-moved constituent is degraded, but is still better than (57b) where *wh*-movement takes place out of a topicalized constituent. Let us accept that the judgements are as given and conclude that *wh*-movement and topicalization are indeed asymmetrically ordered. There is however another interfering factor that should be considered before we draw any firm conclusions on the basis of the patterns given above. Consider the interaction between *wh*-movement and topicalization in cases where neither moving phrase contains the other. Example (58a) shows that a topicalized constituent can skip over a *wh*-intervenor. On the other hand, topics create strict islands for *wh*-movement in German, as illustrated in (58b) (examples are from Müller and Sternefeld (1993)).

- (58) a. ??Radios_i wie_j ich nicht [CP wie_j (daß) [TP man t_j t_i repariert]]
 radios know I not how that one repairs
- b. *Was_i glaubst du [gestern_j hat Ede t_j t_i repariert?]
 was believe you yesterday has Ede repaired

In light of this, it is quite plausible to assume that the same restriction which gives rise to the contrast in (58) is also responsible for the contrast in grammaticality between the examples in (57). In (58b), there is no feeding relation between *wh*-movement and topicalization, rather *wh*-movement crosses a topicalized constituent and the result is ungrammatical. It might be argued then that the reason why (58b) and the corresponding examples involving remnant movement are deviant is because there is an intervenor along the path of movement and not because topicalization feeds *wh*-movement.²⁹ Therefore, we again don’t find here compelling evidence that

²⁹It is interesting to note that a topic intervenes for *wh*-movement only when it is in its licensing position. Consider the cases of remnant movement, such as those in (53). Here the first movement step involves extraction of a *wh*-phrase out of a constituent bearing a topic feature. The second movement step involves topicalization of the remnant and the result is acceptable. In (57b), the first movement step involves topicalization. The second movement step involves extraction of a *wh*-phrase across a constituent bearing topic feature and the result is ungrammatical. Thus, in both configurations there is a step where a *wh*-phrase extracts out

wh-movement and topicalization should generally be asymmetrically ordered. The data reviewed here is compatible with the view that there is no ordering on the application of *wh*-movement and topicalization in cases of cross-clausal movement, though some feeding relations in German might be ruled out by independent principles.

The same pattern of interactions is replicated in the corresponding Norwegian examples, further suggesting that the ungrammaticality observed in examples like (59b) and (60b) might have a common cause (examples courtesy of Kristine Bentzen).³⁰ I assume that the pattern given below is analyzable along the same lines as argued for German.

- (59) a. Om Karl den Stor vet jeg ikke hva slags bok han vil skrive.
 about Charlemagne know I not what kind book he will write
 b. *Om hvilken tysk keiser sa han (at) et ferdig
 about which German emperor said he (that) a completed
 manuskript hadde ingen å tilby?
 manuscript had noone to offer
- (60) a. Radioer vet jeg ikke hvordan man reparerer.
 radios know I not how one repairs
 b. *Hva tror du at igår hadde Jon reparert?
 what you think that yesterday had Jon repaired

With these considerations in mind, let us turn to the interaction between *wh*-movement and topicalization in English. According to Lasnik and Saito (1992), *wh*-movement out of topicalized constituents is at least marginally possible, as already noted in chapter 2. The relevant examples are repeated below.

- (61) a. ??Who_i do you think that [pictures of t_i]_j, Mary believes t_j are on sale?
 b. ??Who_i do you think that [pictures of t_i]_j, John wanted t_j?

(62) is another example illustrating the same point.

- (62) (??) [Which emperor]_i did you say that [even pictures of t_i]_j John worships t_j?

a constituent with a topic feature. The difference is that only in the latter case, the topicalized constituent occupies a licensing position.

³⁰Note that cases like (58) or (60) where two distinct constituents undergo *wh*-movement and topicalization respectively, do not fall under the scope of Abels's (2008b) GenPIM. Under his definition of *affectedness*, the *wh*-phrase in (58b) would not be affected by the previous step of topicalization. Hence, the contrast observed in (58) and (60) would not be related to the one in (57) and (59), and would have to receive an independent explanation.

These examples suggest that topicalization can feed *wh*-movement in English.³¹ The relative acceptability of (63) suggests further that the reverse ordering of these operations is also in principle licit. In these cases, topicalization takes place out of a *wh*-moved constituent.

- (63) a. ?[That medicine]_i I don't know [how much of t_i]_j I should take t_j.
 b. ?[That emperor]_i I don't know [which picture of t_i]_j we should frame t_j.

Taken together with examples like (61) and (62), cases like (63) seem to suggest that *wh*-movement and topicalization in English enter into a symmetric feeding relation. In other words, these patterns might be taken to indicate that topicalization and *wh*-movement are in fact not ordered with respect to each other, at least in cases involving cross-clausal movement, as expected on the approach pursued here.

For the sake of completeness, consider reported judgements regarding the interaction between these two movement operations in cases where neither moving phrase contains the other. Given our analysis of German facts, we expect that both topicalization across a *wh*-moved constituent, and *wh*-movement across topicalized phrases should in principle be possible. Though judgements are subtle, this expectation does seem to be borne out. Grimshaw (1993) has argued that *wh*-movement across intervening topics is marginally possible.³²

- (64) ??Under which/what circumstances did you say that [these books] they would give to children? (from Grimshaw (1993))

Examples involving topicalization out of *wh*-islands are not perfect, but again not fully unacceptable:

- (65) a. ??That book I wonder where John put. (from Lasnik and Saito (1992))
 b. ?That kind of job even the government must wonder who could be happy with. (from Grimshaw (1993))

5.2.5.3 Interim summary

To conclude this section, we have seen that the proposal argued for in this chapter does not exclude symmetric feeding relations between movement operations under certain well-defined conditions. In particular, when positions targeted by two dis-

³¹The examples are not fully acceptable, but neither are the German examples discussed earlier.

³²Many reported examples of this type are ungrammatical, but they involve crossing dependencies.

- (i) a. *What does John think that Bill Mary gave to? (from Rochemont (1989))
 b. *Which books did Lee say that to Robin she will give? (from Culicover (1992))

tinct movement operations are not separated by a phasal boundary, the operations in question are not predicted to be ordered with respect to each other in cases involving cross-clausal movement. As we have seen, no compelling evidence against this prediction can be adduced from the patterns discussed in this section. In addition, the relative acceptability of examples involving the interaction between *wh*-movement and topicalization in English might be taken to suggest that these operations are not ordered, conforming to the predictions of the proposed approach.

It is interesting to compare the current proposal with the classic analysis of improper movement in terms of the A/A'-distinction and the one appealing to a Hierarchy of Operations. The classic analysis groups movement operations either under the label of A-, or under the label of A'-movement and then imposes an ordering restriction on those two classes of operations. Note that on this view, no ordering is imposed on members of the same class, thus basically allowing symmetric feeding relations between movement operations of the same type (A or A'). For the subset of movement operations discussed in this section, such an approach predicts something like (66). Whether middle field scrambling should be placed before or after the *precede*-sign depends on its status with respect to the A/A'-distinction, which as we have seen in chapter 2, is a matter of debate.

(66) case-mvt, VP-scrambling, middle field scrambling \prec *wh*-mvt, topicalization

On an approach such as Abels's (2008b) which appeals to the Hierarchy of Operations, it is assumed that no symmetric feeding relations are permitted and all the operations listed (66) are ordered, yielding (67).³³

(67) (VP-scrambling) \prec case-mvt \prec middle field scrambling \prec *wh*-mvt \prec topicalization

The analysis pursued here thus shares some aspects of the classic approach, although no appeal to the A/A'-distinction is made. Mutual feeding relations are not (always) ruled out, and operations targeting positions within the same domain are predicted to pattern alike as far as possible feeding/bleeding relations with operations targeting a distinct domain are concerned. By *domain* I mean here a part of the clausal structure demarcated by (a) phasal boundary(ies). For the set of movement operations considered in this section, the predictions of the current proposal might be summarized as in (68).

(68) VP-scrambling \prec (subject) case movement, middle field scrambling \prec *wh*-movement, topicalization

³³I have placed VP-scrambling in brackets as this operation is not discussed by the proponents of the Hierarchy of Operations.

5.3 Summary

In this chapter, I have argued that a broad range of violations which might be regarded as falling under the rubric of improper movement can be captured by relying on the internal featural make-up of the moving phrase, in conjunction with the phase-based locality. The account crucially relies on the assumption that there is a close correspondence between the nominal and the clausal functional hierarchy. In particular, I assume that features on the goal are hierarchically ordered in a way that parallels the ordering of these features in the clausal functional sequence. The ordering restrictions on movement operations then fall out as a consequence of the Spell-Out Principle stated below, which requires that, when a phasal head is merged, all projections/features lower on the functional hierarchy than the phasal head get spelled out, rather than only the complement of the phase head, as on standard assumptions.

(69) *The Spell-Out Principle*

Given a functional hierarchy $F_n < F_{n+1} < F_{n+2} < \dots < F_{n+x}$, and F_m a phase head, then merger of F_m triggers Spell-Out of all features/projections lower on the functional hierarchy than F_m .

With these assumptions in place, restrictions on the ordering of various movement operations in both identity and non-identity cases are derived. In addition, the clause-boundedness of certain movement operations is also accounted for. As we have seen, the account does not exclude symmetric feeding relations between movement operations under certain well-defined conditions. Namely, if functional projections targeted by two movement operations are not separated by a phasal boundary, the current proposal imposes no ordering on such operations in cases of cross-clausal movement. Some potential evidence against this prediction has been reviewed and shown not to be compelling.

We have also seen that the proposed analysis overcomes a number of problems that the GBOIM-based approach faced. In particular, improper movement steps in configurations involving subextraction out of moved phrases, which were not captured by the GBOIM-approach, are accounted for in the same manner as identity cases or cases of remnant movement. In addition, long distance A'-dependencies no longer pose a problem, since successive cyclic A'-movement via intermediate vPs is not ruled out in the current model. Finally, not only are these results achieved without an appeal to an extrinsic hierarchy of operations, but also no condition on movement, such as the GBOIM or Abels's GenPIM, need be stated. The ordering restrictions fall out directly from the way movement relates to the functional sequence and from constraints on syntactic derivations imposed by the phase-based locality. Both factors involved, namely the functional hierarchy and the phase-based locality, are needed independently of the improper movement effects discussed here.

Part II
Anaphora

Chapter 6

Setting the stage

The preceding chapters have focused on the well-known ordering asymmetry between A and A' movements, namely the observation that A-movement can feed A'-movement, but A'-movement cannot feed A-movement. I have argued that the observed ordering restrictions on movement operations can be made to follow from the functional hierarchy, in conjunction with a theory of locality, without recourse to ordering statements and/or representational filters. The notion of phases and timing of Spell-Out play a key role in the account. It has been shown that relying on these factors, we can achieve broad empirical coverage and rule out illicit combinations of movement operations not only in configurations involving stepwise movement of the same category, but also in cases of remnant movement and subextraction from moved phrases. In the following chapters, I will argue that the same factors determine the nature of anaphoric relations. In particular, I will claim that the phase constitutes the local domain for binding, offering thereby a unified analysis of locality restrictions on movement and binding dependencies.

The idea that common syntactic factors constrain both movement operations and relations between anaphoric expressions and their antecedents is an old one, going back at least to Chomsky (1973), where it was first proposed that constraints on binding could be reduced to those on movement. In the Chomsky (1973) system, the distribution of reciprocals, such as *each other* in English, was explained by the same principles that constrained transformational operations, since such constructions were thought to involve movement of *each*.¹ In his seminal work, *Lectures on Government and Binding*, Chomsky maintains a connection between movement and binding, though with an interesting modification: rather than binding possibilities being constrained by conditions on movement, in Chomsky (1981) it was binding principles that constrained movement operations. This was achieved by treating traces of movement as elements with referential properties and therefore subject to conditions of the Binding Theory. As already noted in section 3.1 of chapter 3, traces of A'-movement shared the feature specification of R-expressions and were thus subject to Principle

¹Chomsky (1973) offers no account for the distribution of reflexives.

C, while traces of A-movement were classified as anaphoric elements and were therefore under the purview of Principle A of the Binding Theory. Improper movement violations were then reduced to violations of conditions on binding.² However, in his later work (see for instance Chomsky (1993)), Chomsky abandons this line of research, dissociating movement from binding and leaving it therefore unexplained why locality restrictions on these two types of dependencies appear to be so similar. The following chapters will reestablish this link, by arguing that both movement and binding relations are sensitive to phases. Note that an approach which posits a single locality domain for movement and binding is also more theoretically parsimonious than an approach relying on binding and movement-specific locality domains and thus preferable given the aims and methodology of the Minimalist programme.

Extending the relevance of phases from movement to binding phenomena also provides further support for the notion of the phase as a valid theoretical tool for capturing the properties of natural language. In the preceding chapters, I have slightly altered the standard conception of phases and timing of Spell-Out. This modification has then been shown to have important empirical consequences for the way movement operations interact. Throughout the following chapters, I will use the same conception of phases that I relied on in accounting for improper movement and show that it also makes some novel empirical predictions regarding possible anaphoric relations.

In addition to phase-based locality, the internal featural make-up of phrases undergoing movement has played an important role in our analysis of improper movement. I will follow this line of thinking in subsequent chapters as well, especially in chapter 8, linking the distribution of anaphoric expressions to their featural composition. Consequently, no specific conditions regulating the distribution of anaphoric expressions will need to be stated.

Finally, the discussion of the binding phenomena will lead us to address another well-known difference between A and A'-movements, namely that in contrast to A-moved elements, elements that have undergone A'-movement cannot bind anaphors from the derived position. Since the preceding chapters no longer appeal to the binary distinction between A/A'-movement types, this observation must now be restated. A way of capturing this contrast in the model pursued here will be outlined in the following chapter.

The discussion will proceed as follows. The first chapter lays out the hypothesis that phases constitute local domains for anaphora and then tests the plausibility of this hypothesis against the empirical evidence, focusing mainly on English. This kind of approach will lead us to distinguish local from nonlocal binding relations, following some earlier work on anaphora (e.g. Lebeaux (1984), Pollard and Sag

²For some alternative ways of formalizing the relation between movement and binding phenomena refer back to chapter 3, especially the proposals of Williams's (2003) in 3.3 (cf. Riemsdijk and Williams (1981)), and Neeleman and van de Koot (2010) in 3.5. Several other proposals that establish a link between movement and binding will be reviewed in the following two chapters.

(1992), Reinhart and Reuland (1993)). We will however go one step further and argue that there are in fact two, semantically rather distinct, nonlocal binding relations. Considerable part of this chapter will be devoted to the discussion of the prohibition on A'-binders, suggesting a particular way of capturing this restriction and exploring its consequences. In chapter 8, I explore how anaphoric dependencies might be encoded. Following some earlier proposals, I will argue that an Agree relation must be established between an anaphor and its antecedent in the course of the derivation, explaining thereby why phases should act as the relevant locality domains in binding dependencies. The featural make-up of anaphors and pronouns will be shown to play a key role in determining both their interpretation and their syntactic distribution.

Chapter 7

Phases as local binding domains

7.1 Introduction

This chapter concerns itself with the distribution of anaphors (mainly reflexives). A striking property of anaphors is that they require the presence of an antecedent, i.e. another NP with which they corefer, within some local domain. This requirement is met in (1a) for instance. In (1b), even though pragmatic factors strongly point to *Mary* as the intended referent for *herself*, *Mary* is not local enough to bind the anaphor. Since there is no other potential antecedent available, the sentence is ungrammatical.¹

- (1) a. $Mary_i$ criticized herself_{*i*}.
b. *As soon as $Mary_i$ arrived, they asked herself_{*i*} to leave.

This behaviour distinguishes anaphors from pronominals. Pronominals can, but need not, find an antecedent in the same sentence. They simply need their referent to be sufficiently prominent in the discourse for the speaker and the hearer to be able to identify it. Thus, the instance of *her* in (2b) could refer to *Mary* (marked here by coindexation), but it could also equally well pick up another referent from the discourse. In addition, the antecedent must not be ‘too close’ to the pronominal. Hence, replacing the anaphors in (1) with bound pronominals produces the opposite judgements:

- (2) a. * $Mary_i$ criticized her_{*i*}.
b. As soon as $Mary_i$ arrived, they asked her_{*i*} to leave.

This chapter argues for a particular conception of the local domain relevant for binding relations. Specifically, I will try to substantiate the claim that the phase is the relevant locality domain for binding of anaphors. In doing so, I will appeal to the conception of phases as outlined in the previous chapter and motivated by the analy-

¹Following standard practice, the shared subscript index on the anaphor/pronominal is used to indicate that the two phrases enter into referential dependency.

sis of improper movement. Whether the distribution of pronominals can be captured by appealing to the same local domain will be briefly discussed in the following chapter, though in general the main focus will remain on binding relations involving anaphors.

The chapter is structured as follows. I start off by laying out the hypothesis that the relevant local domain for binding is the phase, i.e. the same locality domain that movement operations are sensitive to. In section 7.3, I then step through a number of syntactic configurations, testing the plausibility of this hypothesis against the empirical evidence from English. The relevant data will lead us to draw a distinction between locally and nonlocally bound anaphors in English in section 7.4, and nonlocal binding will be further examined in some detail in section 7.6. Section 7.5 discusses the predictions that the current proposal makes regarding binding into so called picture-NPs. In section 7.7, I explore how the restriction requiring that the potential antecedent should occupy an A-position might be captured in the model of binding developed here. Section 7.8 focuses on binding relations established across a PP-boundary, and particularly on the issue of whether prepositional phrases constitute local binding domains or not. Section 7.9 contrasts the current proposal with another phase-based approach to binding, namely the one developed in Hicks (2009). I close off with a short summary of the main conclusions.

Before we proceed, a note on terminology is in order since the terms found in the literature and the way they are employed from one paper to another yield many opportunities for confusion. Henceforth, I will use the term *anaphor* (with the plural *anaphors*) as a cover term for reflexives and reciprocals, though we will focus mainly on reflexives in this and the subsequent chapter. Since the traditional notion of *pronoun* applies to reflexive (e.g. *myself*, *himself*, etc.) and non-reflexive pronouns (e.g. *he*, *it*, *I*, *us*, *your*, etc.) alike, I will use the term *pronominal* to refer to the latter class (following essentially Chomsky (1981)). For pronominals coindexed with an antecedent in the same sentence I will often employ the term *bound pronominals*. I will also at times make use of the term *anaphora* as it is employed in the semantic literature to refer to the relationship between a referentially dependent expression and a referentially independent expression that serves as its antecedent. Thus, both anaphors and bound pronominals would fall under the scope of this notion.

7.2 Laying out the hypothesis

On the basis of patterns such as those in (1) and (2), Chomsky (1981) formulated the following binding conditions, regulating the distribution of anaphors and pronominals:

- (3) *Binding Condition A*
An anaphor is bound in its governing category.

- (4) *Binding Condition B*
A pronominal is free in its governing category

In this context, ‘bound’ is understood as in (5), while in each case, binding refers to A-binding, i.e. by an element occupying an A-position.

- (5) An anaphor α is *bound* in β if there is a category c -commanding it and coindexed with it in β ; otherwise, α is *free* in β .

Chomsky’s (1981) proposal achieved considerable empirical coverage and provided the most coherent and complete account of the GB era for the distribution of anaphors and pronominals. In some version or other, it is still adopted, or tacitly assumed in much recent work on various aspects of syntactic theory. Therefore, I will henceforth refer to it as *the canonical binding theory*.

The formulations of binding conditions given above appeal to the governing category as the local domain relevant for binding. Roughly, we may characterize the government relation as holding between a lexical head and its complements, and between INFL and its subject when it contains AGR, which it must contain whenever it is tensed (for technical details see Chomsky (1981), pp. 162-170).

- (6) *Governing Category*
 α is the governing category for β if and only if α is the minimal category containing β , a governor of β and a SUBJECT accessible to β .
- (7) SUBJECT includes the following: the subject of an infinitive, an NP or a small clause, AGR in (i), but not NP in (i) if INFL contains AGR.
(i) NP INFL VP, where INFL = [[+/- Tense], (AGR)]
- (8) *Accessibility*
 α is accessible to β if and only if β is in the c -command domain of α and assignment to β of the index of α would not violate (9).
- (9) * $[\gamma \dots \delta \dots]$, where γ and δ bear the same index.

Over the years, a number of modifications to Chomsky’s (1981) definition of the local domain have been suggested in order to meet the empirical challenges that the initial formulation faced (see for instance Huang (1983), Chomsky (1986b), Hestvik (1991)). Some of these empirical challenges will be pointed to in the course of this chapter. From the theoretical perspective, the definition in (6) remains problematic even if it can be made to fit the empirical facts since it crucially relies on the notions of *government* and *accessible SUBJECT*. The arbitrary nature of these theoretical concepts has become a serious cause of concern, especially with the advent of Minimalism. Consider the notion of *government* for instance. In the GB theory, the government relation played a central role since every module of grammar made reference to it. Thus, government was implicated in Case- and θ -role assignment,

trace licensing, in determining the distribution of PRO, and in establishing binding domains. In the Minimalist framework, the status of government has become suspect as it does not meet the Minimalist demands for elegance, naturalness and parsimony. Organization of phrases in an X'-format provides a set of privileged relations, namely the head-complement and the specifier-head relation. Considerations of theoretical parsimony thus urge us to reexamine whether introducing the notion of government, alongside the two primitive X'-theoretic relations, is truly needed to capture the empirical facts. This kind of reasoning led to a series of novel proposals, showing that the relevant grammatical phenomena can be reanalyzed without appealing to government (see Hornstein et al. (2005) for an overview).

The notion of accessible SUBJECT raised many worries of both empirical and conceptual nature already in the GB era. Thus, Huang (1983) has argued that accessibility of a SUBJECT has implications only for the distribution of anaphors, but is irrelevant as far as pronominals are concerned. On the other hand, Bouchard (1985) and Chomsky (1986b) propose ways of eliminating the notion of *accessibility* altogether. Chomsky (1986b) for instance suggests that this might be achieved by assuming that anaphors move to INFL at LF, thereby capturing the empirical facts that previously required appealing to the presence of an accessible SUBJECT. However, as Hicks (2009) notes, it is unclear whether the stipulation of LF-movement of anaphors to handle the relevant facts is theoretically more appealing. It imposes in fact a quite enormous redundancy, with anaphor movement required in all cases where an anaphor is present, but yet only accounting for ungrammaticality in a very limited set of cases (see section 7.4 for the cases in question). Even if the notion of accessibility can be done away with, considerations of theoretical parsimony suggest that the concept of SUBJECT, argued to be relevant for the binding theory, should be derived from more general principles/properties, rather than stated as a primitive.²

In light of these concerns, a redefinition of the relevant local domain for binding is called for. Ideally, such a redefinition should appeal to no binding-specific theoretical concepts and its relevance should not be limited to binding phenomena. In this chapter, I therefore explore the hypothesis that the pertinent local binding domain (at least for anaphors) is the only derivational domain employed in the Minimalist framework, namely the phase. The idea is theoretically appealing as it brings together locality domains of both binding and movement. If phases determine locality and cyclicity of movement, then discovering that other properties of grammar exploit the same fundamental architectural design would provide additional support for the notion and nature of phases, in addition to dispensing with binding-specific locality domains. This chapter will argue that identifying phases as relevant local binding domains has not only theoretical appeal, but empirical bite as well. The idea that the phase might constitute the relevant domain for binding, though relatively recent, is not novel. It has been developed most thoroughly by Hicks (2009). His proposal will be reviewed and contrasted with mine at the end of this chapter.

²See Huang (1983) and Chomsky (1986b) for attempts at simplifying this notion.

If we are to argue that phases are the relevant local binding domains, then we have to be explicit about the identity of phasal heads. Following Chomsky (2001), it is commonly assumed that phase heads are C and v^* which introduces an external argument. However, as already noted in chapter 4, standardly employed phasehood diagnostics suggest that unaccusative and passive verbs also introduce phase boundaries. We have also seen evidence for a phase boundary in VPs headed by raising verbs in the course of the discussion of successive cyclic A-movement in section 4.3 of chapter 4. In light of these observations, the question arises as to which projection in the verb phrase should be identified as phasal. Before answering this question, let me add another observation made in Svenonius (to appear). Svenonius points out that no matter which verb class a particular verb belongs to (e.g. unaccusative, causative etc.), the possibility of combining it with tense, aspect, and modality remains unchanged. I will follow Svenonius (to appear) in assuming that this is because the verb phrase spells out as a phase separately from the material in the T-domain. If the entire VP should spell out, regardless of the number and nature of projections it consists of, then the phasal head must be the one immediately dominating the highest projection in a decomposed VP. I will assume that that projection is AspP.³ Identifying Asp as the relevant phasal head also captures the observation that all types of verb phrases (causative, unaccusative, etc.) are associated with a phase boundary, the observation which here reduces to the fact that all types of verb phrases are dominated by AspP. The proposal to be put forth in the rest of this chapter will therefore be phrased in terms of C and Asp being the relevant phasal heads. However, the choice between Asp and v^* as phases will not in fact be crucial until section 7.7.⁴

Another modification to the standard view of phases that will be adopted here is the one made in the previous chapter. It is generally assumed that merging a phasal head triggers Spell-Out of its complement. Hence, while the phasal head and its specifier remain accessible to further syntactic operations, its complement and anything embedded inside it do not. In the preceding chapter, I have abandoned the assumption that it is the complement of the phasal head which spells out, and assumed instead

³The analysis of improper movement developed in the previous chapter can be easily restated in terms of C and Asp being the phasal heads, instead of C and v/V as assumed there. As far as I can see, nothing crucial in the account changes if this modification is adopted.

⁴Attributing the phasehood status to Asp rather than to a projection within VP allows us also to remain neutral with respect to how the distinction between different verb classes should be encoded. If the differences between unaccusative, unergative and causative verbs are to be encoded syntactically, then one way to think about it would be in terms of different flavours of v (Harley (1995), Folli and Harley (2002)). The distinction between a causative and an unaccusative verb for instance, could then be interpreted as reflecting the choice between the external-argument-selecting v meaning CAUSE, and an unaccusative little v meaning BECOME. Alternatively, we might think of different types of verbs as reflecting differences regarding the number and type of verbal projections, with unaccusatives containing a subset of projections present in causative verbs and crucially lacking the external-argument-introducing head (e.g. Ramchand (2008)). The choice of Asp as a phasal head is then compatible with either of these approaches.

that what spells out is all projections/features lower on the functional hierarchy than the phasal head, as stated in (10).

- (10) *The Spell-Out Principle*
 Given a functional hierarchy $F_n < F_{n+1} < F_{n+2} < \dots < F_{n+x}$, and F_m a phase head, then merger of F_m triggers Spell-Out of all features/projections lower on the functional hierarchy than F_m .

Assuming a tight correspondence between the nominal and clausal functional hierarchy, it has been shown that a broad range of restrictions on ordering of movement operations can be derived by appealing to the principle in (10). Investigation of binding patterns that will be carried out in this chapter will not force us to abandon this conception of phases and Spell-Out. Quite on the contrary, it will be demonstrated that adopting this view leads to some novel predictions, which I will argue are empirically supported.⁵

With this background on phases in mind, let us now turn to empirical facts regarding anaphoric binding and test the plausibility of treating phases as local binding domains.

7.3 Binding in finite and non-finite clauses

Consider first a simple case of binding in finite clauses in (11). The antecedent of the anaphor, *Bill*, is merged in SpecvP, as illustrated below. Since both the anaphor and the antecedent are within the same phase, *Bill* can locally bind the reflexive at this point. Subsequently, AspP will be merged. As Asp is a phasal head, the external argument, which still has a case feature to check, will move to the edge of the phase and from there to SpecTP, its final landing site.

- (11) *Bill_i likes himself_i.*
 [_{vP} Bill [_{vP} likes himself]]
 [_{TP} Bill [_{AspP} <Bill> [_{vP} <Bill> [likes himself]]]]

Contrast this derivation with the one in (12). In this case, the antecedent *Bill* is merged too late to locally bind the anaphor. *Bill* will enter the derivation in the Spec of matrix vP. At this point however, the anaphor will already have been spelled out and therefore inaccessible for purposes of local binding.⁶

- (12) **Bill_i said that Mary likes himself_i.*
 [_{vP} Bill [_{VP} said [_{CP} that [_{TP} Mary [_{AspP} ... [_{VP} likes himself]]]]]]]

⁵For ease of exposition, the internal structure of the relevant phrases will generally not be represented in fine detail that an approach appealing to (10) would require unless making reference to (10) would lead to predictions distinct from the ones arising on the standard view of phases.

⁶I continue to mark the spelled out domain with a strikethrough.

Observe also that anaphors in English need not be bound by the external argument, as evidenced by binding patterns in double object constructions. (13a) shows that an NP in object position can act as a binder (examples from Barss and Lasnik (1986)).

- (13) a. I showed John_i himself_i (in the mirror).
 b. *I showed himself_i John_i (in the mirror).

The possibility of establishing a binding relation between the two objects is expected on the current approach given that they are not separated by a phase boundary.⁷

- (14) [_{vP} I showed [_{VP} John [_{VP} himself]]]]]

The ungrammaticality of (13b) falls out assuming, as is standardly done, that the indirect object asymmetrically *c*-commands the direct object in double object constructions. Since a binding relation can only be established in cases where the potential antecedent *c*-commands the anaphor (at some point in the course of the derivation), (13b) is correctly predicted to be illicit given that this basic structural requirement on binding is not met.⁸

⁷The account of binding patterns in double object constructions (DOCs) argued for here is compatible with a number of proposals regarding the analysis of DOCs. Presenting these proposals at this point would take us too much afield and I will therefore limit the discussion to a couple of comments made in this footnote. One way to think about the multiple VP-shells in the representation in (14) is in terms of verbal decomposition proposed in Ramchand (2008) (see also Pesetsky (1995), Harley (2003) for similar proposals). On Ramchand's approach, the notion of verb is a composite involving maximally three projections: Init(iation)P, Proc(ess)P, and Res(ult)P. The external argument would be introduced in the Spec of InitP. The Goal in the DOCs (*John* in this case) would be merged in Spec of ProcP, while the Theme would occupy the Spec of ResP, which is in the case of DOCs associated with abstract possessional semantics. The aspect of Ramchand's analysis that might prove particularly important in the current context is the fact that the Goal is base generated in a position *c*-commanding the Theme. On Larson's (1988) classic VP-shell analysis, the Goal comes to *c*-command the Theme as a result of NP-movement. If we adopted Larson's approach, then we might expect (13b) to be grammatical given the assumption adopted here that reflexives can be bound at any point in the derivation as long as they are within the same phase domain as the binder. If the Goal is initially merged below the Theme, then it could be bound by the Theme in this position, prior to subsequent movement, and we would derive the ungrammatical (13b). However, whether the approach to binding developed here is ultimately incompatible with Larson's analysis of DOCs depends on the treatment of Condition C. If Condition C also applies in the course of the derivation, then it would rule out (13b) quite independently of Condition A effects. As Condition C effects will be left outside the scope of this study, I won't discuss this issue further at this point.

⁸The *c*-command requirement on binding will be further discussed in subsequent sections.

Note that the external argument is also available as an antecedent in double object constructions. This is again expected considering that the external argument occupies the same phase domain as the two objects when it enters the structure in SpecvP. Once AspP is merged, vP will spell out, but crucially all relevant local binding relations will have already been established by this point.

- (15) John_{*i*} showed Mary himself_{*i*} (in the mirror).
 [_{vP} John showed [_{VP} Mary [_{VP} himself]]]]]

Let us now turn to non-finite clauses which have played a central role in determining binding domains since the earliest formulations of the binding theory. The challenge posed by such constructions is that the antecedent for the anaphor surfaces in the matrix clause, not in the embedded clause which contains the anaphor, as in the case of subject-to-subject raising in (16). Classical GB accounts responded to this challenge typically by extending the local domain to include the matrix clause, so that the matrix subject can act as the local binder.

- (16) Bill_{*i*} seems to have hurt himself_{*i*}.

I assume here that anaphor binding takes place in the course of the syntactic derivation (the idea goes back to Belletti and Rizzi (1988)).⁹ What is required on the approach pursued here is that the anaphor and its antecedent occupy the same phasal domain at some point in the course of the derivation. If the anaphor and a potential antecedent never occupy the same phase, as in (12), a local binding relation cannot be established under the assumption that anaphors are bound in syntax. The reason for this is that, given the PIC, operations that take place in the syntactic component cannot reach across a phase boundary and access elements buried within the phase.

With this in mind, consider the example in (16). It is standardly assumed that the subject in raising constructions moves to the matrix clause from a θ position in the embedded clause. At the point at which the subject enters the derivation it is in the same phase as the anaphor, making it possible to establish a local binding relation, as illustrated in (17). The fact that the subject subsequently moves to the matrix TP is immaterial since binding takes place as soon as the antecedent is merged and before it undergoes further movement.

- (17) [_{vP} Bill_{*i*} [_{VP} hurt himself_{*i*}]]
 [_{TP} Bill [_{AspP} <Bill> [_{VP} seems [_{TP} to have [_{vP} <Bill> [_{VP} hurt himself]]]]]]]

In all the cases discussed so far, binding relations are established between a binder and a bindee in their base positions, before displacement takes place. Con-

⁹This view will be refined in the following chapter where I will argue that a syntactic dependency is established between an anaphor and its antecedent in the course of the syntactic derivation, which then translates into a variable binding relation at the interface. For now, the simplified view assumed in the main text will do.

sider next the cases where movement feeds anaphor binding. A relevant case in point are ECM constructions. The following examples illustrate the relevant patterns:

- (18) a. Bill_i believes himself_i to love Mary.
 b. Mary wants Bill_i to educate himself_i.
 c. *Bill_i wants Mary to educate himself_i.

To capture these patterns, a theory of binding must be able to predict that the local domain for the embedded object should be limited to the embedded clause, while the local domain for the ECM-subject should extend to the matrix clause. In the canonical binding theory, these domains are correctly computed by the definition of governing category. The embedded verb governs the embedded object; hence its governing category is the embedded clause. The ECM-subject, on the other hand, is governed by the matrix verb, and therefore the governing category for the ECM-subject extends to the matrix clause.

Consider now how the binding patterns in ECM constructions are captured on the current approach. (18b) is straightforward and parallel to cases discussed earlier. The ECM subject enters the derivation in SpecvP of the embedded clause. At this point it occupies the same phase as the anaphor, and local binding is established before any movement takes place.

- (19) [_{vP} Bill [_{VP} educate himself]]

The really interesting case is the one in (18a). Here the ECM-subject takes the form of the anaphor and is bound by the matrix-clause subject. How does this come about? The ECM-subject enters the derivation in SpecvP of the embedded clause. However, it is not spelled out within the embedded vP. Following standard assumptions and as discussed in chapter 4, the ECM-subject moves to the matrix clause. As a result, during the course of the derivation it comes to occupy a position within the same phase as the matrix-clause subject, making local binding possible. Assuming that all movement out of a phase proceeds via the phase edge, the first opportunity for establishing a local binding relation arises in the configuration given in (20). At this stage of the derivation, the ECM-subject is located in the phase edge, from where it will subsequently undergo movement to the matrix clause, while the matrix-clause subject occurs in SpecvP, its base generated position.¹⁰

- (20) Bill_i believes himself_i to love Mary.
 [_{vP} Bill [_{VP} believes [_{TP} to [_{AspP} himself [_{vP} <himself> [_{VP} love Mary]]]]]]]

¹⁰Since both the matrix and the ECM subject will continue to move, they will come to occupy positions within the same phase also at later stages of the derivation. This is however irrelevant, considering that already at the stage illustrated in (20), they are in a sufficiently local configuration for binding to take place. I also continue assuming that there is no stop-over in embedded TP in ECM constructions, as discussed in chapter 4. However, nothing hinges on this assumption in the present context.

It is clear from the example just discussed that there is no requirement that the anaphor must be bound within the phase in which it is merged. What is important for local binding to take place is that the anaphor not be spelled out before it can be bound. Therefore, by moving the anaphor out of the Spell-Out domain it becomes possible to establish a binding relation outside of the phase where the anaphor has originally been merged. If the anaphor stays within the Spell-Out domain, then it can only be accessible to potential binders within the same domain. Thus, a binding relation between *Bill* and the anaphor can never be established in a configuration such as (18c), repeated below. Here the anaphor is spelled out within the lowest phase in the embedded clause, while the pertinent binder enters the derivation in the matrix clause, at the point when the anaphor is no longer accessible.

- (21) [_{vP} Bill [_{VP} wants [_{TP} to [_{Asp} Mary [_{vP} <Mary> [_{VP} educate himself]]]]]]]

Let me summarize the discussion so far. I have assumed that anaphor binding takes place in the course of syntactic derivation. As a result, a binding configuration must be established within the phase. If an anaphor does not find an antecedent within the same phase, it will be spelled out and inaccessible to binders merged at later points in the derivation. An anaphor however does not necessarily have to be bound within the phase where it is initially merged. If the anaphor moves out of the Spell-Out domain, it may be locally bound at later stages of the derivation.¹¹ What is important therefore is that both the anaphor and the potential binder be within the same phase at some point in the derivation for local binding to be possible.

So far I have been careful to use the term *local binding*, rather than simply *binding* to refer to cases where an anaphor is bound within the phase domain. The reason for this is that there are contexts in English where it seems that a binding relation between an anaphor and its antecedent can span across phasal boundaries. The claim that I will attempt to substantiate in the following section is that cases of binding across phases are in fact different from cases of binding within the phase domain and that excluding these contexts from the scope of local binding is an advantage rather than a drawback of the proposal.

7.4 Non-local binding

It has long been observed that there are environments in English where anaphors can be bound across rather long distances (Chomsky (1981), Lebeaux (1984)). Consider the following examples where the anaphor finds an antecedent outside the finite clause in which it is contained (from Lebeaux (1984)):

- (22) a. John_i knew that some pictures of himself_i would be on sale.
b. John_i knew that there were some pictures of himself_i inside.

¹¹This view will be refined in section 7.5.

In order to accommodate such examples into the canonical binding theory, Chomsky (1981) incorporates the so called *i-within-i* condition, stated in (24), into his definition of *accessibility*. Recall that the relevant binding domain in Chomsky (1981) is the governing category, the definition of which is repeated below.

- (23) β is a governing category for α if and only if β is the minimal category containing α , a governor for α , and a SUBJECT accessible to α .
- (24) $*[\gamma \dots \delta \dots]$, where γ and δ bear the same index.
- (25) α is accessible to β if and only if β is in the c-command domain of α and assignment to β of the index of α would not violate (24).

Given the definitions above, the governing category for the reflexive in the environments under discussion is appropriately extended to the matrix clause. In (22a), repeated below, NP is not the governing category since it contains no accessible subject. The embedded clause is not a governing category either, under the assumption that AGR element of S is coindexed with the NP it governs. As a result, S is not accessible to the anaphor because of condition (24). The AGR element of the matrix clause is however accessible and therefore the matrix clause defines the local binding domain within which the anaphor must be bound.¹²

- (26) John_{*i*} knew that [S [NP some pictures of himself_{*i*}] would be on sale]

Example (22b) is also ruled out by (24) under the assumption that the expletive and its associate are coindexed.

Thus, the strategy that Chomsky (1981) adopts in dealing with examples such as (22) is to subsume them under the cases of local binding. This kind of strategy would be more difficult to implement on the approach pursued here, where the phase is argued to be the local binding domain. In both examples, the binding relation is established across the CP phasal boundary. It is difficult to imagine how we might go about modifying the notion of phase so as to include these binding environments. Note also that such a move would have far-reaching consequences since, as we have seen in the preceding chapters, the notion of phase plays a crucial role in constraining other syntactic relations as well. The conclusion we are led to therefore is that binding relations in (22) are not instances of local binding.

¹²The crucial assumption on this account is that AGR serves as a kind of antecedent to the anaphor, i.e. it acts as the SUBJECT in the sense relevant to (23). Thus, in (i), the subject is *John*, but the SUBJECT is AGR:

- (i) John [_{INFL} past AGR] win.

In subsequent work (Chomsky (1986b)), Chomsky tries to eliminate this, as he himself notes, “somewhat artificial assumption that AGR can be a binder” by assuming that anaphors move to the INFL position at LF. I will not go into details of this proposal.

That binding relations in examples like (22) are of a different nature than binding relations in local binding configurations has been independently argued for by a number of authors (see Bouchard (1982), Lebeaux (1984), Pollard and Sag (1992), Reinhart and Reuland (1993), among others).¹³ It has been pointed out that locally and nonlocally bound anaphors contrast formally in a number of ways and that the proper characterization of binding conditions A and B should start by differentiating local from nonlocal binding domains. Lebeaux (1984) (see also Bouchard (1982)) identifies an array of syntactic properties which he argues distinguish local and non-local binding configurations. For instance, he points out that only nonlocally bound anaphors allow split antecedents. An antecedent is split if it consists of (at least) two NPs occupying distinct argument positions, as in (27).

- (27) a. John_i told Mary_j that there were some pictures of themselves_{i+j} inside.
 b. *John_i told Mary_j about themselves_{i+j}.

Secondly, the requirement that the antecedent c-command the anaphor seems to hold only of local binding configurations:

- (28) a. John_i's campaign required that pictures of himself_i be placed all over town.
 b. *John_i's wife loves himself_i.

Thirdly, nonlocally bound anaphors give rise to strict/sloppy ambiguity under VP-ellipsis. Consider (29a). What is of concern here is the interpretation of the anaphor in the elided VP. The elided *himself* can be coreferential with either *John* ('strict' reading) or *Bill* ('sloppy' reading). Locally bound anaphors allow only sloppy readings.

- (29) a. John_i thought that there were some pictures of himself_i inside, and Bill did too.
 (=pictures of John, *or* =pictures of Bill)
 b. John hit himself, and Bill did too.
 (=Bill hit Bill, *but* ≠ Bill hit John)

Finally, nonlocally bound anaphors are apparently in free variation with pronominals; locally bound anaphors are not.

- (30) a. John_i knew that some pictures of himself_i/him_i would be on sale.
 b. John_i knew that there were some pictures of himself_i/him_i inside.

- (31) John_i likes himself_i/*him_i.

¹³Pollard and Sag (1992) use the term *exempt anaphora* for anaphors that are not locally bound, while Reinhart and Reuland (1993) term these *logophors*. I will continue to use the more neutral notion *nonlocal anaphora*.

If the binding relation between an anaphor and its antecedent is always local in nature, then the non-complementarity between pronominals and anaphors in environments such as (30) is surprising on the standard model of the binding theory. Given that Conditions A and B are exact mirror images of each other, as is clear from the definitions given below, this view leads to the expectation that pronominals and anaphors should be in strict complementary distribution. This prediction is however not borne out in cases like (30).

(32) *Condition A*
An anaphor is bound in its governing category

(33) *Condition B*
A pronominal is free in its governing category.

It appears then that pronominals and anaphors are in complementary distribution only in cases of local binding. In nonlocal binding environments, both pronominals and anaphors are licit.¹⁴

Given these considerations, examples such as those in (22) no longer pose a challenge to our proposal that the phase is the relevant local binding domain. In fact, the phase-based approach makes the correct prediction that these cases should behave differently from standard cases of local binding.¹⁵

It should be noted at this point that cases of nonlocal binding discussed in this section involve so called *picture-NPs*, i.e. contexts where the anaphor is embedded inside an NP. These environments introduce complications that might make the intervention of CP between the antecedent and the anaphor irrelevant. Namely, a number of authors have claimed that binding into picture-NPs is always nonlocal in nature, even when no clausal boundary is crossed. For Pollard and Sag (1992) and Reinhart and Reuland (1993), the reflexive is nonlocally bound even in cases like (34):

(34) John_i saw a picture of himself_i.

Therefore, binding relations established with an anaphor embedded in an NP merit a more detailed discussion, which I turn to in the next section.¹⁶

¹⁴This claim will be scrutinized in more detail in the following chapter.

¹⁵The nature of nonlocal binding will be discussed in section 7.6.

¹⁶Note that we cannot avoid complications that arise with picture-NPs by using simply reflexives, rather than reflexives embedded in a noun phrase. This is because bare reflexives cannot appear in subject positions in English, as shown in (i-a) and (ii-a). Lebeaux (1983) also notes that there is a contrast regarding the use of reciprocal vs reflexive in subject positions, with reciprocals being marginal in most cases, and reflexives impossible. He thus concludes that reflexives must be subject to a principle beyond Condition A that would distinguish them from reciprocals. There are a number of different approaches as to how the ban on nominative anaphors might be captured (see for instance Everaert (1991b) for an overview). I will leave this puzzle for another occasion.

(i) a. *John thinks that himself will win.

7.5 Binding into picture-NPs

Chomsky (1986b) proposes an analysis of anaphora in subjectless picture-DPs which has the potential of accounting for their non-local properties in contexts like (22), while at the same time treating them as cases of local binding. Namely, he suggests that subjectless picture-NPs contain an implicit argument with the properties of PRO. If so, then the possibility arises that the reflexives in contexts like (22), which I argued to be nonlocally bound, are in fact locally bound by PRO, as shown in (35). The nonlocal properties of the reflexives would then follow from PRO's long distance relationship to its controller.

- (35) a. John_i knew that there were [PRO_i pictures of himself_i] inside.
 b. John_i thought that [those PRO_i pictures of himself_i] were nice.

Although an intriguing proposal, Lebeaux (1984) argues that it is untenable. One piece of evidence against it concerns the interpretation of these examples, which is not what would be expected if an implicit subject is present. Namely, it is not the case that pictures in these contexts are necessarily John's pictures of himself, as the PRO analysis would lead us to expect, rather than someone else's pictures of John. In addition, if the pronominal can be used instead of the reflexive only when PRO is missing, as Chomsky assumes, then pictures in (36) should necessarily be interpreted as someone else's pictures of John. In other words, there should be a contrast in meaning in these environments depending on whether a pronominal or an anaphor appears as the object of the picture noun phrase. According to Lebeaux, no such contrast in meaning is discernible.

- (36) John_i thought that those pictures of him_i were nice.

An additional argument against the PRO analysis is given by Hicks (2005). He points out that NPs with possessive and agentive subjects are always definite, and as such should be illicit as the associate of the expletive *there* due to the definiteness restriction on this construction:

- (37) a. There are [some/several/∅ pictures of Lee Trundle] on every Swansea City fan's wall.
 b. *There are [the/John's/someone's pictures of Lee Trundle] on every Swansea City fan's wall.

If there is indeed PRO in the NP-internal subject position in (35a), we would predict that like (37b), the definiteness effect should render the sentence ungrammatical.

-
- b. ??John and Mary think that each other will win.
- (ii) a. *John didn't know what himself had done.
 b. John and Mary didn't know what each other had done.

In light of these considerations, I conclude that the PRO-based analysis of examples like (35) is not viable. The nonlocal properties of anaphors in these constructions do not follow from long-distance binding of PRO, but rather from the fact that the binder and the anaphor are not within the same local domain. In other words, the cases in (35) are true instances of nonlocal binding. Some additional evidence in favour of this conclusion will be presented in the course of this section. Thus, at least in some cases, binding into picture-NPs must be nonlocal in nature. On the view pursued here, this is true when an anaphor and its antecedent are separated by a phase boundary, as in (35) where a CP phase boundary intervenes.

The question I wish to address now is whether anaphors in picture-NPs can ever be locally bound, or whether binding of anaphors embedded in picture-NPs might in fact always be nonlocal in nature. If the latter turns out to be the case, then the intervention of CP in cases like (35) would be irrelevant.

Consider the binding possibilities into picture-NPs in cases where no clausal boundary intervenes.¹⁷

- (38) a. Mary saw John_i's picture of himself_i.
 b. Mary_i saw a picture of herself_i.
 c. *Mary_i saw John's picture of herself_i.

The approach pursued here predicts that all the licit binding relations indicated in (38) should be local in nature. Recall that the analysis of improper movement in the previous chapter led us to the conclusion that there are no phasal heads in the nominal domain. Assuming a close correspondence between the nominal and clausal functional hierarchy, I have argued in chapter 5 that clausal phasal heads, C and Asp on our view, trigger Spell-Out of nominal functional projections as well, as dictated by the Spell-Out Principle repeated below for the sake of convenience.

- (39) *The Spell-Out Principle*
 Given a functional hierarchy $F_n < F_{n+1} < F_{n+2} < \dots < F_{n+x}$, and F_m a phase head, then merger of F_m triggers Spell-Out of all features/projections lower on the functional hierarchy than F_m .

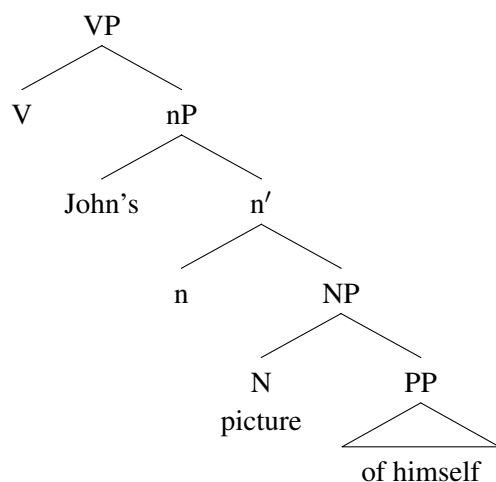
On this view, parts of nominal structure may incrementally become inaccessible in the course of the syntactic derivation, despite the fact that no phasal heads are present in the nominal domain. This allowed us to capture a broad range of restrictions on ordering of movement operations.

If this view is on the right track, then we now predict that a local binding relation should be possible between a possessor and an anaphor embedded in the NP since the two relevant phrases occupy positions within the same phase and no parts of the

¹⁷There is a lot of speaker variation in judgements regarding binding into picture-NPs. While I do feel that ultimately it is important to address such a variation, I won't tackle this issue at this point or try to determine what factors might play a role. For the time being, I will stick to idealized judgements as commonly done in the literature.

nominal structure are spelled out at this point. The relevant part of the derivation is sketched in (40).¹⁸ I assume here that ‘*of*-PPs’ are merged as complements of the noun.

(40)

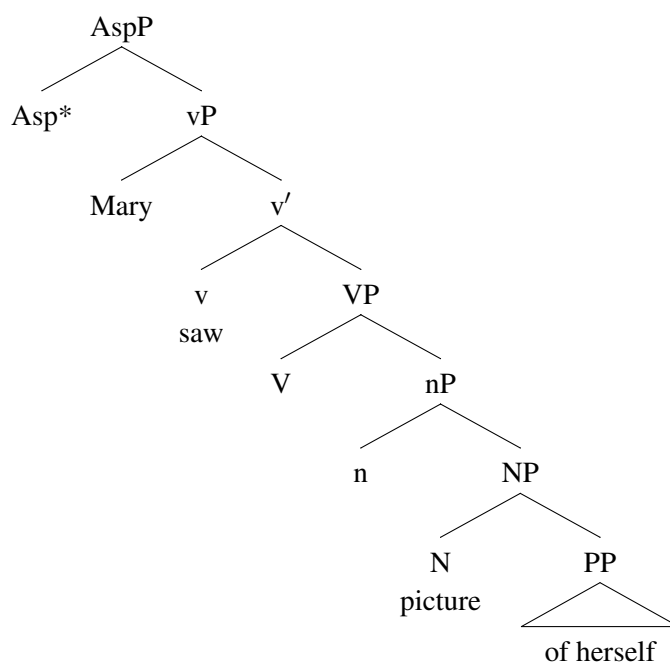


The same prediction is made regarding examples like (38b), where the anaphor is bound by the clausal subject. On our assumptions, the first phasal head to enter the structure will be AspP, as represented in (41). When Asp is merged, all projections/features lower on the functional hierarchy than Asp will be spelled out, as dictated by the Spell-Out Principle in (39). However, since the clausal subject is introduced in a projection lower than Asp, the relevant binding relation between the subject and the anaphor can be established before Asp is merged.¹⁹

¹⁸Irrelevant projections within the nominal domain have been omitted in the following representations.

¹⁹When the picture-NP is displaced, the predictions change. We will return to this point shortly.

(41)



Note that the current approach leads to the expectation that (38c) should be licit as well. The blocking effect of possessors has featured prominently in the alternative analyses of binding phenomena, as will become clear shortly. However, the presence or absence of the possessor is not crucial on the approach sketched here, given the assumptions adopted so far. What matters is whether the anaphor is still accessible at the point when the antecedent enters the structure. Since (38c) parallels (38b) in this respect, we predict that the example should be licit with the coindexing as indicated. I will leave the ungrammaticality of (38c) unresolved for now and return to it in the following chapter after I have discussed the mechanisms involved in encoding binding relations. It will then be shown how the pertinent binding relation can be ruled out on the current model of binding.

This said, let me now turn to predictions made on alternative proposals regarding the status of examples in (38). In the system of Chomsky (1981), the noun phrase containing the anaphor in (38a) and (38c) is the governing category within which the anaphor must be bound, as it contains an accessible SUBJECT, namely the possessor. In fact, the binding patterns in picture-NPs constituted the main reason for assuming that the presence of a subject plays a role in determining binding possibilities. If the noun phrase counts as the local binding domain when the possessor is present, then potential binders outside of this domain are not accessible to the reflexive. This accounts for the contrast in grammaticality between (38a) and (38c). When the possessor is absent, as in (38b), the noun phrase contains no subject and therefore the governing category is extended to the clause. The clausal subject can then locally

bind the anaphor.

Recall however that Chomsky (1981) makes no distinction between local and nonlocal binding relations, subsuming all instance of referential dependency under local binding. The influential proposal by Reinhart and Reuland (1993) that does assume such a distinction makes different predictions regarding the nature of binding into possessor-less NPs. Reinhart and Reuland (1993) treat Condition A as a condition on the distribution of reflexive predicates, rather than a condition on the distribution of anaphors.²⁰ A predicate is considered reflexive if two (or more) of its arguments are coindexed. On this view, a (SELF) anaphor can be logophoric (i.e. nonlocally bound) when it does not reflexive-mark a predicate. Otherwise, Condition A, as stated below, will rule it out.

(42) *Condition A*

A reflexive-marked syntactic predicate is reflexive.

(43) *Definitions*

- a. The *syntactic predicate* formed of (a head) P is P, all its syntactic arguments, and an external argument of P (subject).
- b. The *syntactic arguments* of P are the projections assigned θ -role or Case by P.

Given these assumptions, the binding relation between a clausal subject and an NP-internal anaphor in possessor-less noun phrases, as in (38b) repeated in (44), is predicted to be nonlocal in nature. In such cases, the anaphor does not reflexive-mark the matrix verb since it is not an argument of the verb, the entire picture-NP is. In addition, the definition of syntactic arguments in (43) ensures that the picture-N does not form such a predicate since it lacks an external argument.

(44) Mary_i saw a picture of herself_i.

The blocking effect of the possessor, which was incorporated into the definition of the governing category in the canonical binding theory, is here captured within the definition of syntactic predicates. This is how the difference between (38a) and (38c), repeated below, is accounted for. When the subject (i.e. the possessor) is present, the N forms a syntactic predicate. The anaphor reflexive-marks this predicate and thus Condition A applies to require that it be reflexive. This requirement is met in (45a), where two arguments of *picture* are coindexed, but not in (45b). The latter derivation is therefore ruled out.

- (45) a. Mary saw John_i's picture of himself_i.
- b. *Mary_i saw John's picture of herself_i.

²⁰A more detailed presentation and critical discussion of this proposal can be found in the following chapter.

To sum up, we have seen that there is a general consensus in the literature concerning contexts such as (45a), where the anaphor is bound by the possessor. All the approaches known to me, including the one developed here, predict that such a binding relation should be licit and local in nature. Consensus however breaks down regarding cases like (44). The canonical binding theory and the proposal developed here predict that anaphors in possessor-less noun phrases are locally bound by the clausal subject. On the other hand, on Reinhart and Reuland's (1993) approach (see also Pollard and Sag (1992)), anaphors in these contexts are predicted to be logophoric. Can we adduce some evidence in favour of one or the other view?

First of all, it is important to note that on Reinhart and Reuland's analysis, binding into subjectless picture-NPs is always nonlocal, regardless of the syntactic environment in which this NP occurs. Thus, for them picture-NPs in contexts like (38), receive the same treatment as picture-NPs in examples like (22), repeated here for convenience.

- (46) a. John_i knew that some pictures of himself_i would be on sale.
b. John_i knew that there were some pictures of himself_i inside.

This is not the case on the approach pursued here. While anaphors in subjectless picture-NPs are locally bound in (43), they are not in (46), where a phase boundary intervenes between the binder and the anaphor. Therefore, Reinhart and Reuland's argument that anaphors in subjectless picture-NPs are logophoric which is based on examples in (47) loses its force. These are all examples where the binder and the anaphor are not within the same phase.

- (47) a. A picture of myself would be nice on that wall.
b. Lucie thought that a picture of herself would be nice on that wall.
c. The queen demands that books containing unflattering descriptions of herself will be burned.

We must therefore restrict our attention to configurations such as (38). One way we might go about determining the nature of binding relations in such contexts is to check whether anaphors in these environments show properties of nonlocal binding identified by Lebeaux (1984). An attempt along these lines has been made by Hicks (2009). The conclusion he reached is that results are far less convincing than in cases of binding across a clause boundary into an NP. Recall that according to Lebeaux (1984), locally bound anaphors do not allow split antecedents, must be c-commanded by their antecedents, allow only sloppy readings under VP ellipsis and are in complementary distribution with pronominals. Consider for instance the possibility of taking split antecedents. While (48a) seems fine, Hicks (2009) provides (48b) to show that other types of NPs assumed to be in the class of picture-NPs produce more deviant results:

- (48) a. Bush_i showed Kerry_j every picture of themselves_{i+j}.
b. *?Bush_i told Kerry_j a story/rumour/lie about themselves_{i+j}.

Similarly, testing the possibility of taking a non-c-commanding antecedent yields inconsistent results:

- (49) a. Max_i's eyes watched eagerly a new picture of himself_i in the paper. (Reinhart and Reuland (1993))
 b. *Bush_i's opponents spread malicious rumours about himself_i involving pretzels. (Hicks (2009))

In VP-ellipsis contexts, reflexives in possessor-less picture-NPs sometimes, but again not always, pattern as nonlocal anaphors in giving rise to both strict and sloppy readings (from Hicks (2009)):

- (50) a. Bush wouldn't show the reporters pictures of himself in a pretzel factory, but Kerry would.
 = show the reporters pictures of Kerry, or:
 = show the reporters pictures of Bush
 b. Bush told a great joke about himself, and Kerry did too
 = tell a great joke about Kerry, but
 ≠ tell a great joke about Bush

Finally, the availability of pronominals in place of an anaphor in subjectless picture-NPs seems to suggest that the anaphor is nonlocally bound.²¹

- (51) Mary_i saw a picture of her_i.

The results of applying Lebeaux's diagnostics to environments like (44) thus seems to be inconclusive. Until we understand the nature of the diagnostics in question and the factors that regulate the distribution of nonlocally bound anaphors, it is unclear how these results should be interpreted. I will therefore refrain from drawing any firm conclusions on the basis of such patterns.²²

Note that the difficulty in determining the nature of binding into subjectless picture-NPs in English stems from the fact that anaphors in this language can in principle be both locally and nonlocally bound. Turning to other languages might offer us a clearer picture. Consider German for instance. There is no evidence that the German reflexive *sich* can be used logophorically. In environments which we argued allow nonlocal binding of anaphors in English, the German reflexive *sich* is

²¹It should be noted though that speaker judgements regarding the acceptability of pronominals in these cases are far from clear or consistent. Chomsky (1981) for instance assigns a star to (i) (Chomsky, 1981, p. 217, his 92iii):

- (i) *John_i saw a picture of him_i.

²²Note also that the examples contrasted above differ along several parameters, such as the type of the verb, the type of the prepositional phrase, etc. Some of these factors might play a role in determining binding possibilities. I will leave such issues aside for now, but hope to return to them in future work.

illicit ((52a) is from Buring (2005a), (52b) from Kiss (1993)).

- (52) a. *Martell_i hofft, dass eine Reportage über sich_i im Radio gespielt
 Martell hopes that a report about self in-the radio played
 wird.
 becomes
 ‘Martell is hoping that a report about himself is going to be aired.’
 b. *Gernot_i erinnerte sich daran, daß die ZEIT ein Bild von sich_i
 Gernot remembered self there.on that die ZEIT a picture of self
 veröffentlicht hatte.
 published had
 ‘Gernot remembered that die ZEIT had published a picture of himself.’
 (translation mine)

Nevertheless, *sich* is licit in picture-NPs in contexts like (53).

- (53) a. Ulrich_i las ein Buch über sich_i. (Kiss (1993))
 Ulrich read a book about self
 b. Hans_i soll ein Bild von sich_i einsenden. (Lucius Bader, p.c.)
 Hans should a picture of self send

The same is true in Serbian. The reflexive *sebe* cannot be used in nonlocal binding configurations like (54), but can occur in picture-NPs, as illustrated in (55).

- (54) a. *David_i je mislio da će članak o sebi_i biti interesantniji.
 David aux thought that will article about self be more.interesting
 ‘David thought that the article about himself would be more interest-
 ing.’
 b. *Petar_i se nadao da će ona slika sebe_i sa premijerom biti
 Peter refl. hoped that will that picture self with Prime Minister be
 objavljena.
 published
 ‘Peter hoped that that picture of himself with the Prime Minister will
 be published.’
 (55) a. David_i je izgubio članak o sebi_i.
 David aux lost article about self
 ‘David lost the article about himself.’
 b. Petar_i je prodao onu sliku sebe_i sa premijerom.
 Peter aux. sold that picture self with Prime Minister
 ‘Peter sold that picture of himself with the Prime Minister.’

Facts such as these suggests that anaphors in picture-NPs can be locally bound, as predicted on the approach pursued here.²³

The account pursued here makes another prediction that is not shared by either of the two alternative approaches considered. Since on the current view the nature of a binding relation hinges on the presence/absence of an intervening phasal boundary, anaphoric dependencies in configurations like (56) are predicted to be distinct from those in (57). Only in the latter set of cases are the binding relations nonlocal in nature, given that there is a phase boundary intervening between the anaphor and its antecedent. Anaphors in examples like (56) are instances of local anaphora.

(56) John_i saw a picture of himself_i.

- (57) a. John_i knew that some pictures of himself_i would be on sale.
b. John_i knew that there were some pictures of himself_i inside.

On the other hand, both the canonical binding theory and the reflexivity theory of R&R predict that binding relations in (56) and (57) are uniform in nature. For R&R in both types of configurations we are dealing with nonlocal binding, while the canonical binding theory treats all these examples as cases of local binding. These approaches thus offer little hope of understanding the contrast in grammaticality between German examples in (52) and (53), or the parallel cases in Serbian, given in (54) and (55). The observed contrast is however expected on the approach pursued here, under the assumption that anaphors in these languages must be strictly locally bound. If that is the case, then (52) and (54) are correctly predicted to be unacceptable since a local binding relation cannot be established due to the presence of a phase boundary. When binding into a picture-NP does not cross a phasal boundary, as is the case in (53) and (55), anaphors become licit in these languages. In English, the contrast is blurred by the fact that anaphors in this language can also be nonlocally bound. However, as we have seen, even in English there are some indications that the two configurations must be distinguished, given that they do not equally consistently display properties of nonlocal binding. These considerations suggest that the position that the picture-NP occupies in the syntactic structure matters in determining the binding possibilities and/or the nature of the binding relations established, as predicted on the approach pursued here.

Another set of cases that are relevant in this respect are examples like (58), noted by Barss (1986).

- (58) a. John_i wondered which picture of himself_i Mary saw.
b. John_i wondered which picture of himself_{i/j} Bill_j saw.

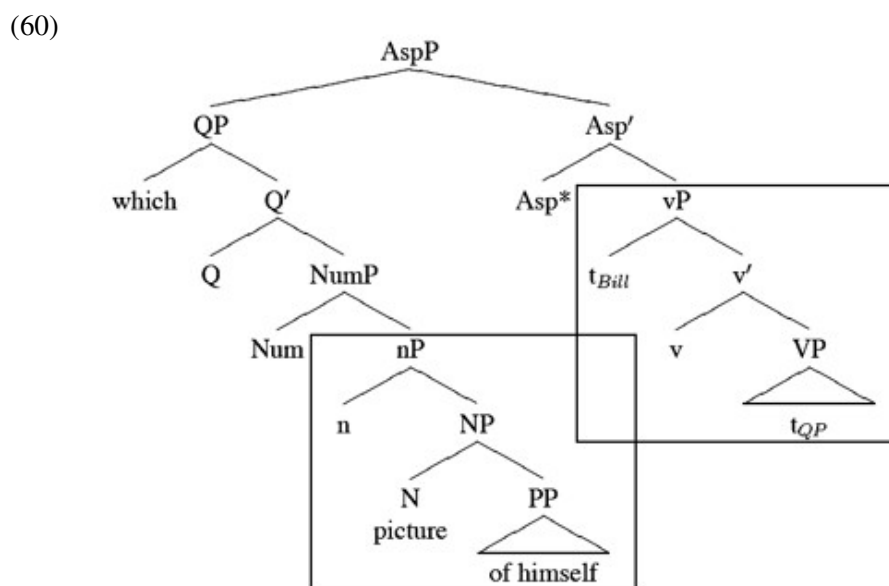
²³Whether anaphors can in addition be nonlocally bound in such contexts in languages like English which allow nonlocally bound anaphors depends on one's view of nonlocal binding. I won't discuss at this point possible restrictions on the distribution of nonlocally bound anaphors. Some aspects of nonlocal binding relations will be discussed in the following section.

- c. Which picture of himself_i does John_i think that Mary saw?

It is often argued that sentences like (58) show that *wh*-movement can feed Condition A (e.g. Chomsky (1995)). According to this view, *wh*-movement of *which picture of himself* would place the anaphor in SpecCP, where its binding domain would include the matrix subject. The examples are of particular interest for a phase-based approach like the one pursued here because the picture-NP in this case occupies the phase edge, the SpecCP. As a result, the picture-NP and the matrix clause subject are within the same phase at one point in the derivation, as schematized below:

- (59) [Asp* [vP SUBJ [vP V [CP picture-NP [C* ...]]]]]

Nevertheless, the binding relation between the anaphor and the matrix clause subject is predicted to be nonlocal in nature on the approach developed here. The reason is that the anaphor embedded in the NP will be spelled out before the picture-NP reaches the SpecCP. To see this, we need to focus on the embedded clause. The first phasal head to enter the structure is Asp. Recall that at this point projections lower than Asp in the functional sequence will be spelled out, including the nominal projections in the Spec of AspP. In other words, on our assumptions, both the vP and its nominal counterpart, nP, will be spelled out at this stage, as illustrated below (the spelled out domains are again boxed).



Given that the anaphor is contained in the spelled out domain, it will not be accessible to potential binders outside of the AspP. As a result, any binding relation between the anaphor and a potential binder introduced at the later stages of the derivation will have to be nonlocal in nature.²⁴

²⁴It might be argued that different types of PPs (e.g. *of*- vs. *about*-PPs) attach at different heights in the nominal structure. On the view pursued here, such differences in the height of

The approach pursued here thus predicts that the syntactic position of the picture-NP bears on possible binding relations. As we have already seen, when the picture-NP is in the base-generated object position, an anaphor embedded inside it can be locally bound by an antecedent outside the NP, as in cases like (61).

(61) Bill_i saw a picture of himself_i.

However, as soon as AspP is merged, the anaphor embedded inside the picture-NP will be spelled out and consequently, once the picture-NP moves out of the AspP, no local binding relations with the anaphor contained inside it can be established.

That anaphors in these contexts are nonlocally bound has been independently argued for by Reinhart and Reuland (1993) and Pollard and Sag (1992). Note also that the comparative data point to the same conclusion. Reflexives in German and Serbian, which we noted cannot be nonlocally bound, are hopeless in these contexts (example (62a) is from Büring (2005a), (62b) is from Kiss (2001)).

- (62) a. Hans_i fragte [welche Bilder von ihm_i/*sich_i] ich gesehen hatte.
Hans asked which pictures of him/self I seen had
- b. [Das Buch über sich_i/*sich_j] glaubt der Gernot_j mag der Ulrich_i.
the book about self believes the Gernot likes the Ulrich
- (63) a. David_i je pitao [koju sliku sebe_{*i}/_j] je Petar_j izgubio.
David aux asked which picture self aux Peter lost
- b. *[Koji članak o sebi_i] je David_i rekao da su novine objavile?
which article about self aux David said that aux papers published

Observe also that in (62b) and (63a), the anaphors can be bound by the subject of the embedded clause, pointing again to the conclusion that these binding relations are different from the ones established after the displacement of the picture-NP. The approach developed here thus correctly predicts that the position of the picture-NP in the syntactic structure has consequences for the possibility of establishing local binding relations. Note that no such prediction is made on alternative approaches

attachment might affect binding possibilities. Note that if a PP is merged in a position higher than nP in (60) it would not be spelled out when AspP is merged. An anaphor embedded in such a PP should then at least be accessible to potential antecedents in the middle field. Investigation of such issues would require a careful study of the properties of different types of PPs, taking us too much afield. I will therefore set these issues aside for now. Note however that for cases discussed here such considerations do not seem to be crucial. The anaphor is presumably a KP, and thus will be spelled out as soon as the embedded C is merged (given that K is lower on the functional hierarchy than C). Therefore, the anaphor should be inaccessible to potential antecedents in the matrix clause, and we again predict that the binding relation between the anaphor and the matrix clause subject in examples like (58) should be nonlocal in nature.

considered.²⁵ Summarizing the discussion of picture-NPs, we have seen that by relying on the conception of phases motivated in chapter 5 on the basis of improper movement violations, we end up making some novel predictions in the domain of anaphoric binding. While all three approaches discussed here correctly account for the fact that the possessor can function as a local binder, predictions diverge regarding the nature of binding into possessor-less NPs. These are summarized in the table below.

	Chomsky (1981)	R&R (1993)	current model
nature of the binding relation	local	nonlocal	local/nonlocal
syntactic position of the picture-NP matters	no	no	yes

Table 7.1: Binding into possessor-less picture-NPs

The current model predicts that an anaphor embedded in a possessor-less NP can be either locally or nonlocally bound depending on whether it occupies the same phase as the potential binder at the point when binding takes place. If it does, as in cases where the picture-NP occupies its base-generated object position, then a local binding relation can be established. Once the picture-NP leaves the AspP, only a nonlocal binding relation can be established. Consequently, only in languages which allow anaphors to be nonlocally bound will such constructions be licit. This is distinct from both Chomsky's (1981, 1986) and Reinhart and Reuland's (1993) view. While for Chomsky, binding into subjectless picture-NPs is always local, for R&R anaphors in these contexts are always logophoric. I have argued here that the empirical facts, and particularly the comparative data, speak in favour of the current model, i.e. that binding into subjectless picture-DPs is not always uniform in nature.

In the following section, I focus on nonlocal binding relations, examining their nature in more detail and arguing that the typology of binding relations should be further refined.

7.6 Zooming in on non-local binding

In the previous sections, I have followed the work of Lebeaux (1984), Pollard and Sag (1992) (henceforth P&S), Reinhart and Reuland (1993) (henceforth R&R), among others, in drawing a distinction between local and nonlocal binding environments. However, as we have already seen, the set of environments where I have argued non-local binding obtains is not identical to the one predicted on alternative approaches.

²⁵Hicks (2009), who also develops a phase-based account of binding phenomena, distinguishes configurations such as those in (56) from those in (57), but not from those in (58), as is the case on the current view. See section 7.9 for further elaboration on this point.

In this section, I spell out my assumptions regarding the nature of nonlocal binding relations, which will mark a further departure from these earlier accounts.

7.6.1 Refining the typology of binding relations

Having identified nonlocal binding relations, an important question that arises is what the nature of this type of binding relation is. For P&S and R&R, coreference in nonlocal configurations is not subject to principles of grammar and is regulated solely by discourse considerations. We might follow P&S and R&R in assuming that all uses of anaphora in nonlocal configurations are logophoric. If so, syntax would have nothing interesting to say about these cases and we could relegate all the potential questions to a theory of discourse. I will however pursue here a different kind of approach. I have argued in the previous sections that anaphoric binding takes place in the course of the syntactic derivation, making phases the relevant local binding domains. Once the phase is complete, the domain of the phase is shipped to the interfaces for interpretation. The possibility then arises that an anaphor which has not been bound within a phase in the course of the syntactic derivation, gets bound at the semantic interface. In other words, a reflexive in English which, as we have seen, can find an antecedent outside the local binding domain, might not be discourse-bound, but rather bound at the interface. There might in addition be also logophoric uses of reflexives in English, but not all nonlocal binding configurations would be logophoric on this view. Some potentially good candidates of true logophoric uses of reflexives in English are given below:

- (64)
- a. A picture of myself would be nice on that wall. (from Reinhart and Reuland (1993))
 - b. What about yourself? (from Hicks (2005))
 - c. John_i was furious. The picture of himself_i in the museum had been mutilated. (from Pollard and Sag (1992))

While not denying that anaphors in (64) might be discourse-bound, I will argue here that some instances of nonlocal binding take place at the semantic interface, rather than at the level of discourse. I will show that such an approach avoids certain complications of the theory of variable binding that arise if all cases of nonlocal binding are treated as logophoric. It also reduces the class of examples relegated to the discourse component by placing a subset of nonlocal relations under the purview of grammar proper.

Thus, in the system just outlined, three possibilities arise for establishing shared reference between an antecedent and an anaphor. A binding relation might be established:

1. in syntax
2. in semantics

3. at the discourse level

Anaphors in particular languages then might differ in whether they allow all, or only some of these binding options. English reflexives seems to be relatively flexible in allowing all three options, while anaphors in German and Serbian appear to be more restricted and require a local antecedent, as we have seen in the previous section.

In the rest of this section, I will probe the nature of nonlocal binding a bit further, arguing in favour of the proposed two-way distinction in nonlocal binding relations. Given the grammaticality of examples like (64) where there is no potential antecedent available in the sentence, we already know that discourse-binding must in general be allowed as an option (at least for reflexives in English). The question therefore is whether there is some reason to assume another type of nonlocal binding relations, distinct from discourse-binding. In this section I argue that there is.

That a binary distinction in binding relations is not sufficient has also been argued by Williams (2003), who considers as one of the fundamental problems of R&R's approach the fact that there is nothing intermediate between the grammatical and logophoric anaphora. He observes that the behaviour of reciprocals cannot satisfactorily be resolved on this approach. This is because reciprocals can be nonlocally bound but do not escape the utterance altogether in finding their antecedent.

(65) [John and Mary]_i think pictures of [each other]_i are in the post office.

If we conclude, on the basis of the grammaticality of (65), that *each other* can be logophoric, then we might expect (66) also to be licit, on a par with examples in (64).²⁶

- (66) a. [John and Mary]_i called on each other at the same time. *[Each other]_i's houses consequently had a forlorn and deserted look.
 b. [John and Mary]_i were furious. *A picture of [each other]_i in the museum had been mutilated.

Williams (2003) concludes that what is needed is something intermediate between grammatical and discourse-level anaphora. In the current model, the contrast between (65) and (66) might be interpreted as suggesting that reciprocals in English can be nonlocally bound at the interface, but do not have truly logophoric uses.

7.6.2 Variable binding vs coreference

Another advantage of assuming a three-way distinction in binding relations along the lines argued for here is that it allows us to retain the standard assumptions concerning variable binding. It is well-established in the semantic literature on binding that there

²⁶Example (66a) is from Williams (2003), example (66b) is from Peter Svenonius, p.c.

are two strategies of anaphora resolution: variable binding and coreference.²⁷ In cases like (67a), we can think of the embedded pronominal as a referring expression of its own, which just picks out the same individual as its antecedent. We will refer to this relation as one of *coreference*. In (67b), the pronominal functions as a variable, bound by the quantified NP.

- (67) a. John_{*i*} said that he_{*i*} was OK.
 b. No woman_{*i*} doubts that she_{*i*} is OK.

It is important to note that the syntactic environments allowing coreference are not identical to those allowing a bound variable interpretation, suggesting that the two interpretations must be kept apart. Generally, a bound variable interpretation is possible only when the antecedent c-commands the pronominal, as in (67b). This condition is not met in (68) and (69) (examples from Reinhart and Grodzinsky (1993)), which explains the ungrammaticality of (68b) and (69b). Nevertheless, the pronominals in (68a) and (69a) can refer to the same individual as *Lucie*. This suggests that there must be another strategy of anaphora resolution apart from variable binding, and that the strategy in question is not sensitive to the c-command requirement.

- (68) a. Most of her_{*i*} friends adore Lucie_{*i*}.
 b. *Most of her_{*i*} friends adore every actress_{*i*}.
 (69) a. A party without Lucie_{*i*} annoys her_{*i*}.
 b. *A party without every actress_{*i*} annoys her_{*i*}.

The relation between *Lucie* and the pronominal in the examples above has been argued to be one of coreference, not variable binding. The coreferential interpretation is not available in (68b) and (69b). A pronoun cannot *corefer* with a quantified NP since quantified NPs are not referential expressions in the first place. Thus, when the antecedent is a quantified NP only the bound variable interpretation is possible. Since the binding relations in (68b) and (69b) must involve variable binding, but the potential binder does not c-command the pronominal, the examples are simply ungrammatical on the coindexing indicated.

While the coreferential use is sensitive to the semantic type of the antecedent, the bound variable interpretation is not. As Reinhart (1983b) and Reinhart and Grodzinsky (1993) show, the bound variable interpretation is available with all types of NPs, including the referential ones. A sentence like (70), where *Alfred* and *he* refer to the same individual (marked here by coindexing), is in fact formally ambiguous between the two readings represented in (70a) and (70b).

²⁷Note that the term *anaphora* is used here in the broad sense that encompasses both bound pronominals and reflexives. See footnote (2) for clarification. I also continue to use the term *pronominal*, as a way of distinguishing non-reflexive from reflexive pronouns.

- (70) Alfred_i thinks he_i is a great cook
- a. Alfred (λx (x thinks x is a great cook))
 - b. Alfred (λx (x thinks he is a great cook))

In (70a), where the pronominal is construed as a bound variable, it is the property of considering oneself to be a great cook that is attributed to Alfred, whereas in (70b), on the coreference interpretation, it is the property of considering Alfred to be so. In the context of (70), these two interpretations are equivalent. However, there are contexts where this is not the case. One of these is given in (71), which is ambiguous with respect to whether (70a) or (70b) is attributed to Alfred only.

- (71) Only Alfred_i thinks that he_i is a great cook.

Reinhart (1983b) has argued that the failure to acknowledge the distinction between coreferential and bound variable uses of pronominals has led to many misconceptions and complications in the canonical binding theory. Under the standard view, there is no difference between (68a) and (68b) as far as the binding theory is concerned. The ungrammaticality of (68b) is then attributed to a peculiar property of quantification in natural language, formulated as Bijection Principle, which filters it out at LF.²⁸ The principle prohibits an operator from locally binding more than one variable. After QR has applied to (68b), the operator *every* locally binds both the trace and the pronoun. Although the Bijection Principle rules out (68b), it seems like a rather ad hoc solution since it is not, as Reinhart and Grodzinsky (1993) note, reducible to any independent property of operators.

Reinhart (1983b) suggests an alternative way of approaching these facts. Note that both anaphors and pronominals can function as bound variables, in which case they need to be syntactically bound, i.e. coindexed with a c-commanding antecedent. Pronominals can however also choose their reference freely from the discourse and it is this property that distinguishes them from other anaphoric expressions. These considerations lead to the conclusion that coreference is in fact the exceptional case. Reinhart (1983b) thus argues that binding conditions regulate only bound variable anaphora, while coreference is governed by distinct principles as a peculiar property of reference resolution in natural language. Pronominals do fall under the scope of the binding theory, but only with respect to their bound variable interpretation. On this view, the ungrammaticality of (68b) is not unexpected, but rather conforms to the general requirement on anaphoric relations.

If coreference is not subject to the binding theory, then requirements imposed on syntactic binding have no consequence for coreferential interpretations. Variable binding (or *semantic binding*, as it is often referred to) on the other hand, requires syntactic binding (see also Bach and Partee (1980), Heim and Kratzer (1998)).

²⁸The Bijection Principle was proposed by Koopman and Sportiche (1982) and subsequently adopted in the GB framework.

(72) *Syntactic binding*

A node α is *bound* by a node β iff α and β are coindexed and β c-commands α .

This means that c-command is not a prerequisite for a coreferential relation to be established, nor is syntactic coindexing necessary. Reinhart (1983b) in fact argues that coreference is just a subcase of the broader process of reference resolution that takes place in the discourse component.²⁹ It is clear that we need mechanisms to regulate reference resolution in cases of cross-sentential anaphora where the syntactic binding requirements are not met. As Reinhart and Grodzinsky (1993) note, there is no reason why the same mechanisms could not apply sentence-internally.³⁰

With these considerations in mind, let us now focus on examples in (73). By turning a potential antecedent into a quantified NP, it can be shown that anaphors in nonlocal binding contexts can function as bound variables.

- (73) a. Every boy_{*i*} asked which picture of himself_{*i*} Lucie likes.
 b. Which picture of himself_{*i*} does [every boy]_{*i*} think that Lucie likes?

Acknowledging this, R&R conclude that the relation between an antecedent and a nonlocally bound anaphor (logophor on their view) can be either that of coreference, or of variable binding. But recall that for them logophoric uses of anaphors are governed purely by discourse principles. This means that variable binding can also be achieved in the discourse component. If so, it is difficult to see how the standard c-command requirement on variable binding should be stated, given that the discourse component should not have access to such structural notions as c-command. In fact, it seems that on such an approach, the well-established c-command requirement must be abandoned altogether. Consider (73b). In this case, the antecedent does not c-command the anaphor at surface structure. Nevertheless, the sentence is grammatical. This is noted by R&R as a problem for the theory of variable binding under their approach. If variable binding can be established without c-command, then a host of ungrammatical cases remain unaccounted for, including those in (68b) and (69b).³¹

²⁹A similar view of coreference is presented in Partee (1978) who views *coreferential anaphora* as just one subcase of a more general phenomenon of *pragmatic anaphora*.

³⁰A terminological remark is in order. It is important to note that the use of the notion *coreference* in the GB-style binding theory differs from the way semantic literature has been employing this term. In the GB framework, the term is used to characterize the relation between an antecedent and an anaphor/pronominal in contexts where they overlap in reference. Coindexation is standardly used to mark such a relation. The use of the term is then extended also to cases involving quantified NPs, since these also involve coindexation, although, as we already noted, quantified NPs do not strictly speaking refer. The main text assumes a much narrower use of this term. In fact, on Reinhart's view coreference obtains in environments where there is no coindexation, insofar as coindexation is a signal of a syntactic binding relation.

³¹Even if we resorted back to something like a Bijection Principle to rule out (68b) and (69b), as R&R seem to suggest, it is still unclear to me why the same principle does not rule

Note that these problems might be avoided if nonlocal binding is allowed to operate at the semantic interface. If we can appeal to the semantic representation of (73b), then, after reconstruction, the quantified NP would c-command the anaphor, and (73b) would thus conform to the c-command requirement on variable binding. There is thus no need to allow variable binding to take place in the discourse component in order to accommodate cases of nonlocal binding such as those in (73), and in fact by not doing so we circumvent the problems related to the c-command requirement noted above. I will therefore assume that discourse binding involves only coreference, never variable binding. Given that coreference is not subject to principles of grammar proper, it is thus not surprising that it is insensitive to structural conditions such as c-command.

These considerations lead to the following typology of binding relations:

	domain	semantic relation	c-command required	beyond principles of grammar proper
binding established in syntax	phase	variable binding	yes	no
binding established in semantics	sentence	variable binding	yes	no
binding at the level of discourse	discourse/text	coreference	no	yes

As stated in the table above, both binding in the course of the syntactic derivation and binding at the semantic interface involve variable binding.³² They differ however with regard to locality: only binding in syntax is constrained by phases. On the other hand, discourse binding is semantically distinct from the previous two in that it involves coreference, not variable binding. Discourse binding should therefore be impossible with quantified antecedents. We thus end up with two distinct binding relations, both of which can be characterized as nonlocal binding relations.

Since the current proposal builds on Reinhart's (1983b) view that the binding theory is only concerned with the bound variable, but not the coreferential reading of anaphoric expressions, let me end this section by noting that in her later work (see Reinhart (2000)), Reinhart argues that the traditional distinction between binding

out (73b).

³²This view will be refined in the following chapter. For now, we will stick to this coarse characterization.

and coreference should be modified to distinguish binding and covaluation. Covaluation is a cover term which includes coreference and so called cobinding. That another anaphoric relation termed cobinding must be distinguished was argued by Heim (1998) on the basis of examples like (74).

(74) Every wife thinks that only she respects her husband.

(74) can mean either that every wife thinks that other wives do not respect their husbands, or that every wife thinks other wives do not respect her husband. The question is what gives rise to this ambiguity. The second reading cannot be due to coreference since coreferential interpretation is not available with quantified NP antecedents. The two readings are then argued to arise as a consequence of two different binding patterns. On both construals, *she* is bound by *every wife* (or more precisely by the λ -operator that *every wife* is a sister of). However, *her* is bound by *she* on the first reading, and by *every wife* on the second. Thus, on the latter construal, *she* and *her* end up cobound, i.e. both bound by *every wife*.

In what is to follow, I will ignore cobinding and stick to the traditional distinction between variable binding and coreference. Cobinding is only relevant in configurations involving two or more pronominals linked to the same antecedent. We will leave such complex examples out of the scope of our study. Furthermore, as Heim (1998) and Fox (2000) argue, the main difference between regular binding and cobinding is in the degree of locality involved. Both involve operator-variable dependencies, but in cases of covaluation, the λ operator that binds the variable is not the closest one available. Coreference however still remains distinct from the other two binding relations in not requiring c-command and being restricted to referential antecedents.³³

7.6.3 A fresh look at nonlocal binding contexts

The previous sections have argued that a relation between an anaphor and its antecedent can be established either in the course of the syntactic derivation, at the semantic interface, or at the level of discourse. Only the first type of binding relation must be established within the local domain, leaving us with two types of nonlocal binding relations. Unlike previous approaches which relegate all nonlocal anaphoric relations to the discourse component, we have now taken a subset of nonlocal relations back under the wing of grammar proper. Thus, we have considerably reduced the number of occurrences of discourse-bound or logophoric uses of anaphors.³⁴ Hopefully, this will allow us not only to avoid the problems noted in the

³³For further discussion of cobinding see Heim (1998), Reinhart (2000), Fox (2000), Büring (2005b).

³⁴A note on terminology. The term *logophoric* may be somewhat misleading given that it was coined by Clements (1975) to describe pronouns in the West-African language Ewe which can take as their antecedents an individual whose belief or attitude towards a certain state of affairs is being reported. R&R use the term to mean all instances of reflexives which are grammatical despite the fact that they appear bound beyond what they consider to be the

previous sections, but also to ultimately get a sharper picture of discourse-binding. It is thus worth revisiting examples of nonlocal anaphora given in the literature, in light of our two-way distinction in nonlocal binding relations.

Clearly, examples where there is no sentential antecedent available for the anaphor must be treated as cases of logophoric use also on the approach pursued here. Examples of this type are particularly frequent with first and second person reflexives in English.

- (75)
- a. There were five tourists in the room apart from myself. (from Reinhart and Reuland (1993))
 - b. Physicists like yourself are a godsend. (from Reinhart and Reuland (1993), citing Ross (1970))
 - c. She gave both Brenda and myself a dirty look. (from Reinhart and Reuland (1993), citing Zribi-Hertz (1989))
 - d. Both John and myself knew the answer but didn't dare say it. (from Hicks (2009))
 - e. No-one misbehaved, myself excepted. (from Hicks (2009))

Though it seems that examples of this type are more difficult to construct with third person reflexives, it is claimed that they are not impossible. Two such examples are given below.

- (76)
- a. Physicists like herself are rare. (from Hicks (2009), citing Fiengo (1977))
 - b. John_{*i*} was furious. The picture of himself_{*i*} in the museum had been mutilated. (from Pollard and Sag (1992))

There are however many examples where third person reflexives are bound beyond the local domain. In the rest of this section, we will focus on such examples.

Consider the following examples of picture-NPs:

- (77)
- a. John_{*i*} found a picture of himself_{*i*}.
 - b. John_{*i*} heard stories about himself_{*i*}.

Reinhart and Reuland (1993) and Pollard and Sag (1992) claim that anaphors in these contexts are nonlocally bound. I have already argued that there are reasons to doubt this claim, particularly in light of the cross-linguistic data. Recall that reflexives in languages like German or Serbian are licit in these contexts, although they in general cannot be nonlocally bound. I repeat the relevant contrast from Serbian.

local binding domain. P&S employ the term *exempt* anaphora, instead of logophoric, in the same sense. I will use the term logophoric to mean discourse-bound. Defining the term in this way has no significant consequences for R&R's, or P&S's approaches since for them nonlocally bound reflexives are always bound in the discourse component. It does however have significance for the approach developed here since nonlocally bound does not equal discourse-bound on our view.

- (78) a. *David_i je mislio da će članci o sebi_i biti interesantniji.
 David aux thought that will articles about self be more.interesting
 'David thought that articles about himself would be more interesting.'
 b. David_i je čuo priče o sebi_i.
 David aux heard stories about self
 'David heard stories about himself.'

Patterns such as the one in (78) led us to the conclusion that binding into picture-NPs need not always be nonlocal in nature as R&R and P&S argue. On the approach pursued here, cases such as (77) involve local binding relations. If this conclusion is on the right track, then regardless of how we treat nonlocal binding contexts, we already have a narrower set of discourse-bound anaphora than R&R or P&S assume.

Let us now turn to examples which would feature nonlocal binding relations both on R&R/P&S's view and on the approach developed here. First of all, there are cases like (79), discussed already by Lebeaux (1984) and Chomsky (1986a) (examples are from Lebeaux (1984)):

- (79) a. John_i knew that some pictures of himself_i would be on sale.
 b. John_i knew that there were some pictures of himself_i inside.

The binding relations in (79) are clearly nonlocal in nature as they cross a finite clause boundary. However, on our view they still need not be relegated to the discourse component. The antecedent, *John*, c-commands the reflexive and the binding relation can be established in semantics.

However, as P&S argue, there are many grammatical cases where the antecedent does not c-command the reflexive. Given that on our assumptions, c-command is a necessary requirement for establishing a successful binding relation at the semantic interface,³⁵ such examples would be potentially good candidates of discourse anaphora. One set of examples that P&S mention are cases involving *wh*-movement:

- (80) a. Which picture of himself_i does John_i think Mary sold?
 b. Which picture of herself_{i/j} does Susan_i think Mary_j prefers?

Although the antecedents do not c-command the anaphors at surface structure in these cases, they do so after the phrase containing the anaphor has reconstructed either to its base position or to one of the intermediate positions on its movement path.³⁶ Hence, not only can these cases be analyzed as involving binding at the semantic interface, but as I have already argued (see the discussion regarding examples like (73b)), adopting such a view has clear advantages over a discourse-based analysis. Such an approach is further supported by the contrast between (81) and (82) (the examples are from Pesetsky (1987)). In (82) c-command is not satisfied even after

³⁵See the following chapter for further discussion of this point.

³⁶Note that on our view, the 'j'-cases might be instances of local anaphora, so it is the 'i'-cases which are particularly significant.

reconstruction, and such examples are therefore ungrammatical as expected. If on the other hand cases like (80) involved logophors and c-command was irrelevant, then one would expect that (81) and (82) should pattern alike.

- (81) a. Pictures of himself_i I know John_i likes.
 b. Which pictures of herself_i did you say Mary_i bought?
- (82) a. *Pictures of himself_i I know John_i's friends like.
 b. *Which picture of herself_i did you say the company that employs Mary_i bought?

Let me now turn to other examples discussed by P&S where binding seems possible despite the apparent lack of c-command. The examples they give are listed below:^{37,38}

- (83) a. The fact that there is a picture of himself_i hanging in the post office is believed (by Mary) to be disturbing Tom_i.
 b. The picture of himself_i in the museum bothered John_i.
 c. The picture of herself_i on the front page of the Times made Mary_i's claims seem somewhat ridiculous.
 d. The picture of each other_i with Ness made [Capone and Nitty]_i somewhat nervous.
 e. John_i's campaign requires that pictures of himself_i be placed all over town.
 f. John_i's intentionally misleading testimony was sufficient to ensure that there would be pictures of himself_i all over the morning papers.
 g. The agreement that [Iran and Iraq]_i reached guaranteed each other_i's trading rights in the disputed waters until the year 2010.
 h. The picture of herself_i on the front page of the Times confirmed the allegations Mary_i had been making over the years.

Examples (83a) and (83b) involve object experiencer verbs. This class of verbs shows many interesting properties, one of which is the long-noted binding puzzle that (83a) and (83b) illustrate: anaphors contained within the subject of such verbs

³⁷I have reordered their examples to facilitate subsequent discussion. Examples (83a) and (83e) are originally from Jackendoff (1972).

³⁸They also give the following two examples, although as far as I can see, the antecedent does c-command the anaphor in these cases. I assume that these examples can be handled in the same way as those in (79), i.e. as involving binding at the interface.

- (i) a. They_i made sure that nothing would prevent each other_i's pictures from being put on sale.
 b. [Kim and Sandy]_i knew that *Computational Ichthyology* had rejected each other_i's papers.

may be bound by the object, in violation of the usual c-command condition on bound anaphora. However, Belletti and Rizzi (1988) have argued that the surface subject in such constructions is c-commanded by the Experiencer object at D-Structure.³⁹ Assuming then that Condition A can apply at D-Structure, the Experiencer can properly bind the anaphor, satisfying the c-command requirement. The odd binding properties of Experiencer verbs are thus just another instantiation of the same problem we observe in (80) and (81): in both contexts, movement of the phrase containing the anaphor reverses the original c-command relations. Given that I have been assuming that Condition A can be satisfied at any point during the syntactic derivation, we thus need only adopt Belletti and Rizzi's proposal that the surface subject starts out lower than the Experiencer and we can bring (83a) and (83b) back under the purview of grammar proper. (83b) might then instantiate a local binding relation. The anaphor is base-generated in a position lower than *John*, where it is bound prior to undergoing movement to the subject position. The anaphor in (83a) might be bound at the interface, after the clause containing it reconstructs to its base position where it is c-commanded by *Tom*. There is no need to appeal to discourse binding to account for these cases.

Unlike object experiencer verbs, subject experiencer verbs fail to display the backward binding pattern (examples are from Hale and Keyser (2002)):

- (84) a. *Each other_{*i*}'s mothers love [Bill and Hank]_{*i*}.
 b. *Each other_{*i*}'s students respect [Noam and Morris]_{*i*}.

This observation calls for an explanation but is not addressed by P&S. On the other hand, the ungrammaticality of (84) is expected on Belletti and Rizzi's view since at no point in the course of the syntactic derivation is the experiencer c-commanded by the antecedent in (84).

Note further that licit binding configurations with object experiencer verbs allow quantified NPs as antecedents. On our view, this means that cases like these cannot involve discourse-binding since the only strategy of anaphora resolution at the discourse level is coreference. It also further supports the conclusion that the required c-command relation holds at a certain point in the course of the syntactic derivation.⁴⁰

- (85) a. Jokes about herself_{*i*} amuse every philosopher_{*i*}. (Reinhart and Reuland (1993))
 b. A picture of himself_{*i*} in the papers would bother any mafia boss_{*i*}. (Peter Svenonius, p.c.)

³⁹Belletti and Rizzi's proposal was aimed at solving another puzzle posed by experiencer predicates, namely the problem that these constructions raise for the *Uniformity of Theta Assignment Hypothesis (UTAH)*.

⁴⁰The grammaticality of (85) is not unexpected on R&R's view, given their assumption that discourse-binding involves both coreference and variable binding. However, we have already seen that this assumption requires considerable weakening of the c-command requirement on variable binding.

Pesetsky (1995) notes that the same kind of binding puzzle observed with object experiencer verbs arises in periphrastic causatives:

- (86) a. Each other_i's remarks made [John and Mary]_i angry.
 b. These stories about herself_i made Mary_i nervous.
 c. Pictures of each other_i caused [John and Mary]_i to start crying.

According to Pesetsky (1995), even (87) are fairly acceptable compared to the similar agentive examples in (88).

- (87) a. ?Each other_i's stupid remarks eventually killed [John and Mary]_i.
 b. ?Rumours about herself_i always plunge Mary_i into a deep depression.
 (88) a. *Each other_i's stupid friends eventually killed [John and Mary]_i.
 b. *Each other_i's swimming coaches plunged [John and Mary]_i into the pool.

Let us add another set of examples. Consider (89):

- (89) Nixon gave Mailer a book.

As discussed in Pesetsky (1995), (89) can be interpreted as asserting that Nixon performed a particular action, i.e. giving a book to Mailer. However, the sentence has another reading which might be paraphrased as in (90):⁴¹

- (90) "Mailer wrote a book which he wouldn't have been able to write if it hadn't been for Nixon."

Interestingly, the second reading, which Pesetsky refers to as *causative*, is missing in the alternant with *to*. (91) is thus fine, but has only the "perform an action"-reading. This can be made particularly clear by replacing *Nixon* with a non-Agentive DP, in which case the *to*-construction lacks any non-deviant reading. (93) illustrates the same pattern (examples are from Pesetsky (1995)):

- (91) Nixon gave a book to Mailer.
 (92) a. The war years gave Mailer his first big success.
 b. *The war years gave his first big success to Mailer.
 (93) a. Katya taught me Russian.
 b. Katya taught Russian to me.
 c. Lipson's textbook taught me Russian.
 d. *Lipson's textbook taught Russian to me.

Observe now that backward binding is possible on this causative reading of verbs like

⁴¹The example and observation go back to Oehrle (1976).

give (from Pesetsky (1995)):

- (94) a. Each other_{*i*}'s remarks gave [John and Mary]_{*i*} a book.
 b. Those books about himself_{*i*} taught Bill_{*i*} the meaning of caution.

Pesetsky (1995) argues that all these cases, including the object experiencer verbs, can be unified under the following generalization:

- (95) A Causer argument of a predicate π may behave as if c-commanded by an argumental DP governed by π .

Building on Belletti and Rizzi's proposal, he then argues that the availability of backward binding (and other puzzling properties) in these constructions is due to the fact that the Causer argument starts out lower than (other) internal arguments. This initial "lowness" allows anaphors inside the Causer argument to be bound by apparently non-commanding antecedents.⁴²

If this conclusion is on the right track, then examples (83c) and (83d) that P&S provide can be dealt with without resorting to discourse binding.⁴³ We may debate the correctness of Pesetsky's analysis, but it remains to be shown how the analysis of backward binding in these constructions in terms of discourse-binding would capture the relevant facts and whether it could offer a way of relating the binding patterns to other puzzling properties of these constructions, such as the unavailability of the *to*-alternate on the causative reading of double object verbs.

Let us now turn to the examples in (83e) and (83f) where the antecedent is a possessor embedded in a DP, instantiating another violation of the c-command requirement. Note that the problem in this case is in fact more general and extends also to bound variable uses of pronominals:

- (96) a. Everyone_{*i*}'s mother kissed him_{*i*}. (from Reinhart (1983b))
 b. Every girl_{*i*}'s father thinks she_{*i*}'s a genius. (Kayne (1994))

This has long been noted as a problem for the c-command requirement on variable binding and various ways around this problem have been suggested in the literature (see for instance Reuland (1998)). In fact, as Hicks (2009) points out, other interpretative phenomena assumed to require c-command in order to be licensed show the same behaviour, such as Negative Polarity Items:

- (97) Noone_{*i*}'s ticket will be worth anything if the manager decides to rest all the best players. (from Hicks (2009))

⁴²Pesetsky's analysis is naturally more complex than reported here. For relevant details, I refer the reader to the work in question.

⁴³Even if the phrase containing the anaphor in (83c) reconstructs to a position below the antecedent, the problem with c-command persists given that *Mary* does not c-command out of the DP. We will return to this issue shortly.

P&S argue that the correct generalization has to do with viewpoint rather than with c-command. According to them, the contrast in grammaticality between (98a) and (98b) is due to the fact that the bearer of the experiencer role is the individual whose viewpoint is being reflected: in (98a) the experiencer is *John*, but in (98b) it is *John's father*, not *John*. Similarly, (98c) and (98d) are fine because in all these cases it is *John* whose viewpoint is reflected.

- (98) a. The picture of himself_{*i*} in *Newsweek* bothered John_{*i*}.
 b. *The picture of himself_{*i*} in *Newsweek* bothered John_{*i*}'s father.
 c. The picture of himself_{*i*} in *Newsweek* dominated John_{*i*}'s thoughts.
 d. The picture of himself_{*i*} in *Newsweek* made John_{*i*}'s day.

Nevertheless, even if this proposal were adopted to handle the reflexives, the more general problem for variable binding would persist. On the other hand, if a satisfactory solution to the c-command puzzle that possessor DPs pose could be found, then again there might be no need to deal with cases like (83e) and (83f) in terms of discourse binding.⁴⁴

This leaves us with examples in (83g) and (83h) as the only potential cases of discourse-binding in (83). All other examples, as we have seen, might be reanalyzed as cases of nonlocal binding at the interface.

7.6.4 Further considerations

In the previous section, I have argued that a large set of cases that P&S (and R&R) treat as discourse binding could in fact involve binding at the semantic interface, and thus be under the purview of grammar proper. Two additional factors must be considered when discussing logophoric uses of reflexives, which I will just briefly mention in this section.

P&S note that the written register seems to tolerate violations of Principle A more easily. Zribi-Hertz (1989) illustrates this by giving a number of examples attested in the works of various writers, two of which are reproduced below:

- (99) a. Not till she had, with difficulty, succeeded in explaining to him_{*i*} that she had done nothing to justify such results and that his_{*i*} wife was equally incredulous of her innocence and suspected himself_{*i*}, the pastor, to be the cause of her distress, did his face light up with understanding. [William Gerhardie]

⁴⁴P&S also provide the following example with the possessor acting as a binder (originally from Lebeaux (1984)):

- (i) A fear of himself_{*i*} is John_{*i*}'s greatest problem.

I will leave the discussion of copular constructions for some future occasion. Let me just note that P&S acknowledge that their account does not explain why anaphors in copular constructions should be exempt.

- b. Clara_i did not know whether to regret or to rejoice at their arrival; she_i did not get on well with either of them (...), and yet, on the other hand their presence did not intensify the difficulty, but somehow dissipated and confused it, so that at least its burden did not rest upon herself_i alone. [Margaret Drabble]

Examples such as these are predicted to be ungrammatical on P&S's and R&R's analyses because according to these analyses anaphors cannot be used logophorically when they appear in an argument position. According to P&S, it is instructive to note that examples such as these are uniformly judged ungrammatical by American speakers. They thus conclude that grammaticality of cases such as (99) is a special property of the written register. It seems that the grammatical constraints on anaphora can sometimes be relaxed in highly stylized narrative.

R&R argue that focus is another factor which seems to facilitate logophoric uses of anaphors in argument positions. They distinguish two logophoric uses: perspective and focus logophors. Perspective logophors are the by now familiar to us point-of-view logophors, exemplified in R&R's view by the following examples:

- (100) a. There were five tourists in the room apart from myself.
 b. Physicists like yourself are a godsend.
 c. She gave both Brenda and myself a dirty look.
 d. It angered him that she . . . tried to attract a man like himself.
 e. Max boasted that the queen invited Lucie and himself for a drink.

In all these cases, anaphors do not occupy an argument position and are thus predicted to be logophoric. On R&R's view, (101) are crucially different from (100). In (101), a free SELF anaphor occurs in an argument position at S-structure. According to R&R, this is only possible when the anaphor is focussed. They then assume that focus expressions undergo movement at LF and are thus no longer in argument positions at LF. If Condition A applies also at LF, then it follows that focus anaphors should appear exempt from Condition A despite occupying argument positions at S-structure.

- (101) a. This letter was addressed only to myself.
 b. Why should the state always take precedence over myself?
 c. Bismarck's impulsiveness has, as so often, rebounded against himself.

In R&R's view thus, although both (100) and (101) involve logophoric uses of anaphors, they are significantly different. Namely, those in (101), where the anaphors occur in argument positions at S-structure, can only be occurrences of focus logophors. R&R then go on to claim that "this is why such examples are more marked: they are possible only when the context clearly signals focus, or a contrastive reading, so they are highly context-dependent." (Reinhart and Reuland (1993), p. 673). On the other hand, logophors not occurring in argument positions at S-structure, like (100), "do

not require any special accommodation and are easily judged acceptable with no context” (Reinhart and Reuland (1993), p. 673). I leave to future research to determine whether such a stark contrast as R&R suggest really exists between (100) and (101).

Given that we have reanalyzed a substantial number of examples featuring anaphors in non-argument positions as involving nonlocal binding at the interface, rather than discourse binding, the question arises whether the right cut is really between examples like (83) and (100) versus those in (101). On the approach pursued here, discourse-binding constitutes a far more limited phenomenon in English than R&R and P&S lead us to expect. If factors such as register and focus-marking need to be taken into account, we could then ask whether the remaining cases of nonlocal binding (at least in case of third person reflexives) which cannot be reanalyzed as involving binding at the interface might in fact involve uses of reflexives in special registers or uses of reflexives as focus markers.

7.6.5 Interim summary

In this section, I have argued that binding relations which are not established within the local domain need not necessarily be analyzed in terms of discourse binding. I have distinguished two types of nonlocal binding relations: nonlocal binding established at the interface and discourse binding. Semantically, the two types of binding relations are significantly different: the former one involves variable binding, while the latter one involves coreference. Consequently, only binding at the interface requires c-command and allows quantified NPs as antecedents. Binding in the discourse component, on the other hand, is not sensitive to c-command relations and is limited to referential antecedents.

7.7 The ban on A'-binders

Since the earliest formulations of the binding theory, it has been assumed that a binder must occupy an A-position for a local binding relation to be possible. In all the cases we have looked at up to now, this requirement has indeed been met. That elements in A'-positions cannot serve as potential antecedents can be shown on the basis of examples like (102). In the first example it is the topicalized phrase which fails to serve as a potential antecedent, while the remaining examples involve *wh*-movement ((102a–c) are from Grewendorf and Sabel (1999), citing Chomsky, class lectures; (102d) is from Williams (2003)).

- (102) a. *The guests_{*i*}, [each other_{*i*}'s dance partners] criticized <the guests>.
 b. *Whose friends_{*i*} did [each other_{*i*}'s pictures] convince the director that he should interview <whose friends>?

- c. *Which actors_{*i*} did [pictures of themselves_{*i*}] convince the director that he should interview <which actors>?
- d. *John wondered [which man]_{*i*} pictures of himself_{*i*} convinced Mary that she should investigate <which man>.

In the canonical binding theory, it is simply assumed that a potential antecedent must occupy an A-position for local binding to be possible. In this section, I will discuss how this restriction might be reformulated, and possibly derived, within the model of anaphoric binding argued for here and without appealing to the notions of A/A'-positions.

7.7.1 Local binding relations

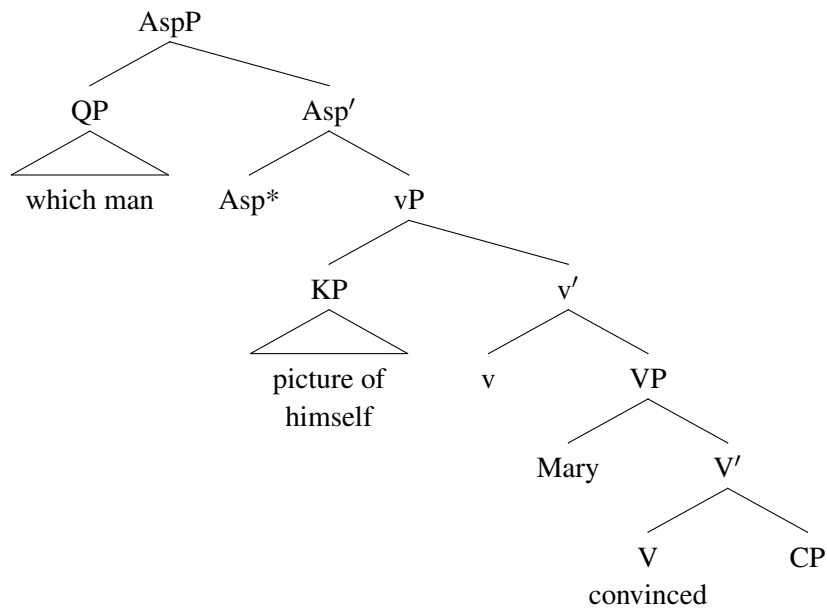
Let us start tackling the ban on A'-binders by focusing on example (102d), repeated below, in a bit more detail.

- (103) *John wondered [which man]_{*i*} pictures of himself_{*i*} convinced Mary that she should investigate <which man>.

The *wh*-phrase undergoes successive-cyclic movement to SpecCP. What (103) seems to illustrate is that the *wh*-phrase cannot bind the reflexive from SpecCP. Note that given our earlier assumptions, we already predict that a *wh*-phrase should fail to bind a reflexive embedded in a picture-NP from the Spec of CP. In such contexts, the picture-NP will already have moved from its base-generated position to Spec of TP. Recall from our discussion in section 7.5 that when a picture-NP occupies Spec of TP, the anaphor embedded inside it is already spelled out and hence can no longer be locally bound. Binding an anaphor in the complement of an NP from the Spec of CP is therefore expected to be impossible.

However, on a phase-based approach, a phrase undergoing successive-cyclic movement must stop off at every intermediate phase edge. On our assumptions (at least) C and Asp are phasal heads. Therefore, to successfully rule out (103), we must in fact focus on an earlier stage in the derivation when the *wh*-phrase lands in Spec of Asp. What we need to rule out therefore is a binding relation between the anaphor and the *wh*-phrase in Spec of AspP. At this point, the anaphor is still not spelled out and occupies the same phase as the potential *wh*-antecedent. Hence, given what we have said so far, it should be possible to establish a local binding relation between the *wh*-phrase and the reflexive in Spec of vP. The relevant configuration is given below:

(104)



In order to rule out anaphoric binding in a configuration such as this, a potential antecedent occupying a position in the phase edge must not be allowed to bind an anaphor embedded within the domain of the phase. Such a result would be achieved if Spell-Out applied before binding. I would therefore like to suggest that upon the merge of a phase head, no operations other than the evacuation of elements with unvalued features can take place prior to Spell-Out. As we have already noted, it is standardly assumed that once a phasal head is merged, the domain of the phase spells out. The relevant question to ask then is at which derivational point precisely the Spell-Out applies, and what operations involving the phase head are allowed to apply before a part of the syntactic structure is sent to the interfaces. We know that the Spell-Out of the relevant chunk of syntactic structure must be slightly delayed to allow elements with unvalued features to move out of the Spell-Out domain, otherwise there could be no movement across phasal boundaries.⁴⁵ I assume here that movement of elements in need of further feature valuing to the phase edge is the only operation allowed to take place before Spell-Out makes part of the syntactic structure, or more precisely all the features/projections lower than Asp on the functional hierarchy, inaccessible to further syntactic computation.⁴⁶ With this assumption in place, binding in configurations such as (103) becomes impossible.⁴⁷ In (104), Spell-Out

⁴⁵I won't go here into the issue of how these intermediate movement steps should be implemented.

⁴⁶See Hiraiwa (2002) for similar ideas. Hiraiwa argues that the complement domain of H_P , H_P a phasal head, becomes inaccessible by the PIC as soon as Merge extends the H_P .

⁴⁷Note that the assumption adopted here does allow other operations involving the phasal head to apply after Spell-Out. Since the phasal head is not spelled-out, it is available to further

will apply as soon as the QP moves to the edge of AspP. The anaphor will then be spelled out and local binding therefore impossible. Hence, on this view, no binding relation can be established between a phrase moved to the phase edge and an element embedded within the Spell-Out domain.

Observe that if the picture-NP in (104) would itself move to the Spec of AspP, i.e. to the phasal edge, it would nevertheless be impossible to establish the relevant binding relation. Recall that, given the Spell-Out Principle proposed in chapter 5, the anaphor in the complement position of the noun will be spelled out when Asp is merged even if it occupies the Spec of AspP, as discussed at length in section 7.5. If Spell-Out must apply as soon as the picture-NP (and the *wh*-phrase) moves to the edge of AspP, as assumed here, then we correctly predict that the anaphor cannot be bound by *which man* in (103). The relevant binding relation is thus ruled out regardless of whether the picture-NP moves to the phase edge or not.

Note that if the assumption regarding the timing of Spell-Out suggested here is adopted, then it becomes crucial to treat Asp, rather than *v* as a phasal head. If *v* had the status of a phase, then we would predict that an external argument merged in the Spec of *v*P cannot locally bind an anaphor within VP, given that phrases in the Spec of a phasal head cannot serve as potential binders. This would rule out an array of grammatical local binding configurations. Having Asp serve as a phasal head circumvents this problem. This can then be seen as an additional argument in favour of treating Asp, rather than *v*, as phasal.⁴⁸

The account of the ban against A²-binders just sketched raises a potential problem for cases involving subject-to-subject raising. Consider the following example (from Reinhart and Reuland (1993)):

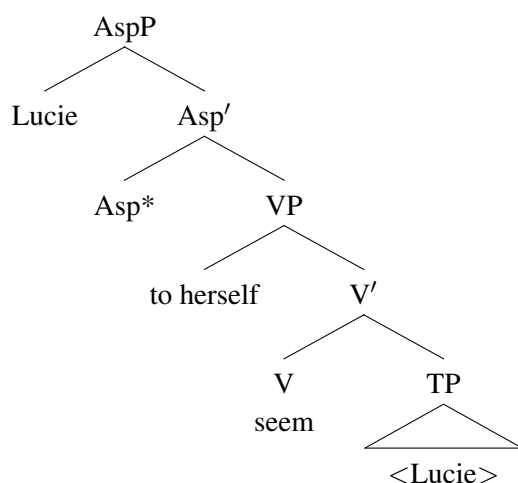
(105) Lucie_{*i*} seems to herself_{*i*} to be beyond suspicion.

Examples such as (105) have been used to argue that in contrast to A²-moved elements, A-moved phrases can act as binders from the derived position. The raised subject is able to bind a reflexive in the matrix clause. To see why this binding relation might be problematic, consider (106).

syntactic manipulation. Thus, the possibility of externally merging specifiers and adjuncts to a phase head is not ruled out.

⁴⁸For further motivation refer back to section 7.2.

(106)



The subject raises from the embedded clause to the matrix TP, making a stopover in Spec of AspP. It cannot bind the reflexive from the Spec of matrix TP, since there is a phase boundary intervening, namely AspP. It cannot however bind the reflexive from the Spec of AspP either since we have just ruled out the possibility of establishing a binding relation between a phrase in the Spec of a phase head and an anaphor embedded within the same phase.⁴⁹ We are thus forced to conclude that there is a landing site for the raised subject below AspP. I will not speculate further on the nature of this landing site. I would just like to add two additional comments regarding (105). One is that the problem which arises in (105) does not necessarily generalize to all examples involving A-movement. In other words, it is not the case that A-moved elements can never bind anaphors from derived positions under our current set of assumptions. We will return to this point shortly. The second thing to note regarding (105) is that constructions involving raising across experiencers are puzzling in other respects as well. One problem raised by such constructions that has attracted a lot of attention in the literature is that the subject moves across the experiencer in apparent violation of Attract Closest. It is a cross-linguistically robust fact that the experiencer blocks raising of the embedded subject. As an illustration, consider the Icelandic pattern below. That it is the presence of experiencer which blocks raising is made clear by contrasting (107b) with (107c), where the experiencer is not lexically realized (examples from Boeckx (1999)).

- (107) a. þeim hafði virst Ólafur vera gáfaður.
 them.DAT has seemed Olaf.NOM be intelligent
 ‘Olaf seemed to them to be intelligent.’

⁴⁹Note that no problem would arise if raising verbs did not constitute phases. However, such an assumption would contradict our conclusions from chapter 4.

- b. *Ólafur hafði virst þeim vera gáfaður.
Olaf.NOM has seemed them.DAT be intelligent
- c. Ólafur hafði virst vera gáfaður.
Olaf.NOM has seemed be intelligent

English therefore appears to be special in allowing subjects to raise across the experiencer. In most languages, configurations such as that in (105) would not even arise. Preminger (2010), for instance, suggests that this locality puzzle observed in English might be handled if we were to assume that the raised subject stops off in the Spec of ApplP, the same projection which hosts the experiencer. It is often assumed that multiple specifiers of the same projection are equidistant with respect to a structurally higher probe. If the experiencer and the raised subject occupy specifiers of the same head at a certain point in the derivation, then no intervention effects are expected. An analysis along these lines thus suggests that there might be a landing site for the raised subject below the matrix AspP, as the account of binding facts adopted here leads us to conclude.

Let us now briefly return to some other instances of A-movement and see how their binding properties can be accounted for in the current model. Consider the following examples of clause-internal scrambling in German (from Müller (1995)):

- (108) a. daß der Arzt_i den Patienten_j sich_{i/j} <den Patienten>
that the doctor.NOM the patient.ACC self.DAT
im Spiegel zeigte.
in-the mirror showed
- b. daß man die Gäste_i einander_i <die Gäste>
that one.NOM the guests.ACC each.other.DAT
vorgestellt hat.
introduced has

As already noted, it is standardly assumed in the literature on German that indirect objects precede and c-command direct objects. What the examples in (108) then show is that movement of the direct object across the indirect object can create new binding possibilities.⁵⁰ The pattern can easily be accommodated within the current model as long as the displacement of the direct object is extremely local, i.e. within the domain of AspP. This is indeed the standard view in the literature, where this type of scrambling in German is generally analyzed as adjunction to VP.

There are also instances of clause-internal scrambling targeting a position in the middle field which seem to be able to feed anaphoric binding, as illustrated by the following example from Japanese (from Saito (1992)):

- (109) a. ?*[otagai_i-no sensei]-ga karera_i-o hihansita (koto).
[each.other teacher].NOM they.ACC criticized fact
'(The fact that) each other's teachers criticized them.'

⁵⁰See chapter 3, section 3.3 and chapter 5, section 5.2.4.

- b. ?Karera_i-o [otagai_i-no sensei]-ga <karera-o> hihansita (koto).
 they.ACC [each.other teacher].NOM criticized fact
 '(The fact that) them, each other's teachers criticized.'

This type of scrambling is commonly analyzed as adjunction to IP. Whether the scrambled phrase adjoins to IP, or moves to the Spec of a functional projection immediately dominating IP is immaterial on the approach pursued here. What is crucial is that this landing site is below the CP phase. Therefore, the binder and the anaphor are within the same phase domain and the observed binding relation can be established.

It is a well-known fact that in contrast to clause-internal scrambling, long-distance scrambling in Japanese never gives rise to new binding possibilities. Contrast (109b) with a case of scrambling across the CP-boundary illustrated below (from Saito (1992)):

- (110) *Karera_i-o [otagai_i-no sensei]-ga [Hanako-ga <karera-o> hihansita
 they.ACC [each.other teacher].NOM Hanako.NOM criticized
 to] itta (koto).
 COMP said fact
 '(The fact that) them, each other's teachers said that Hanako criticized.'

As (110) shows, a long-scrambled phrase cannot bind an anaphor in the matrix clause. The contrast is standardly attributed to the A vs A'-nature of the movements involved. Here the contrast reduces to the fact that long-scrambling targets a position in the matrix C-domain. As a result, a phase boundary intervenes between the scrambled phrase and the anaphor and the binding relation cannot be established. Note that even if the scrambled phrase lands in the Spec of CP it cannot bind the anaphor, given that we have ruled out the possibility of binding by a phrase in the phase edge. Finally, moving the long-scrambled phrase to a position in the middle field is ruled out as an instance of improper movement.

Summarizing this section, we have seen that the unavailability of phrases occupying A'-positions to serve as potential binders reduces to the fact that in such configurations there is a phase boundary intervening. Even in cases where the antecedent occupies the phase edge, a binding relation cannot be established given our assumption that upon merge of a phase head, no operations other than evacuation of elements with unvalued features can take place prior to Spell-Out. With this assumption in hand, all the potential cases of binding between an anaphor and an A'-binder are ruled out.

Note that the proposal just sketched rules out only cases of local binding. A question that arises at this point is why the reflexive in English cannot be nonlocally bound by the antecedent in an A'-position, since we know that reflexives in English need not be locally bound. The discussion to follow outlines a possible way of tackling this issue.

7.7.2 Non-local binding relations

In the previous subsection, I have discussed how a local binding relation between an anaphor and a phrase in an A'-position might be ruled out on the current approach. However, it remains unclear why the anaphor cannot be non-locally bound in the same contexts. Given our conclusion from section 7.6 that nonlocal binding can be established at the interface, the fact that examples like (111) are ungrammatical seems to suggest that the ban on A'-binders holds at this level as well, something that remains unaccounted for under our current set of assumptions. Here I will make some tentative remarks as to why this might be the case, leaving a thorough investigation of this issue for future research.

- (111) *John wondered [which man]_i pictures of himself_i convinced Mary that she should investigate <which man>.

Since our account of the prohibition against A'-binders captures only cases of local binding, let us therefore take it that no such prohibition holds of nonlocal binding. This means that the ungrammaticality of (111) is not due to the fact that the antecedent occupies an A'-position but must be attributed to some other factor. What could this factor be? One option that might be worth exploring relates to the observation that reflexives in English are subject-oriented. It is a well-known fact that long-distance bound anaphors tend to prefer subjects as binders cross-linguistically.⁵¹ The ungrammaticality of (111) might then be due to the fact that the antecedent is not a subject. Chomsky (1986a) provides the following examples in support of the claim that long-distance bound anaphors in English are subject-oriented (from Chomsky (1986a), his (237) and (217ii) respectively):

- (112) a. They_i told us_j that [pictures of each other_{i/*j}] would be on sale.
b. *I told them_j that Bill liked each other_i.

In (112a), the binder of *each other* must be *they*, not *us*. Note that subject-orientation does not hold for cases of local binding.

- (113) They_i told us_j about each other_{i/j}

In (113), either *they* or *us* may be the antecedent of the anaphor. Thus, for cases of local binding we still need an independent account for the ban on A'-binders. In cases of nonlocal binding on the other hand, the ungrammaticality of relevant examples may reduce to the requirement that the anaphor be bound by a subject.⁵²

⁵¹Consult for instance Giorgi (1984), Chomsky (1986b), Reuland and Koster (1991), Cole and Hermon (1998).

⁵²How subject-orientation of long-distance anaphors might be derived is too complex an issue to be addressed within the scope of this section. I will therefore set this important question aside for now. The solution proposed here must therefore remain just a tentative suggestion, until an analysis of subject-orientation compatible with the set of assumptions

Interestingly, that binding across greater distances should in principle allow A'-antecedents is also predicted on Williams's (2003) approach. Let me briefly review Williams's proposal and the evidence he provides in support of this prediction.⁵³ As already noted in chapter 3, Williams (2003) derives a typology of anaphoric elements by associating them to different levels of his Representation theory. Associating different anaphors to different levels interacts with the Level Embedding Conjecture to determine the properties of anaphoric elements, namely their locality restrictions and the choice of antecedents. An anaphor introduced at a particular level will be bindable by elements that are defined at that level. Thus, a CS (Case Structure) anaphor will not be able to take topics as antecedents, since topics are introduced at a later level. That same anaphor will also have to find an antecedent within the same tensed clause, since tensed-clause embedding also takes place at a later level. However, if a particular anaphor can be bound across a tensed clause boundary, then that anaphor must itself be introduced at a later level and thus have at its disposal a broader set of possible antecedents. In other words, if topics appear at the level of SS, and tensed-clause embedding happens at the same level, then an anaphor bound across clausal boundaries should allow topics as antecedents. In this system therefore, the loosening of locality requirements on binding relations predicts a larger set of potential antecedents to be possible.

As a support for this prediction, Williams (2003) points out that in some languages nonlocally bound anaphors can indeed be bound by elements in A'-positions. Korean appears to be among these languages. Consider the behaviour of the Korean anaphor *caki* (from Williams (2003), citing Gill (1999)). Example (114a) shows that *caki* can be nonlocally bound. (114b) and (114c) show that *caki* can be bound by an A'-antecedent, namely a Topic. Similar facts hold for *zibun* in Japanese and *ziji* in Chinese.

- (114) a. John_i-i Bill_j-ekey Mary_k-ka caki_{i/*j/k}-lul cohahanta-ko
 John-NOM Bill-DAT Mary-NOM self-ACC like-COMP
 malhayssta
 told
 'John told Bill that Mary likes self.'
- b. John_i-un ttal-i caki_i-pota ki-kate kuta
 John-TOP daughter-NOM self-than height-NOM more is tall
 'As for John, his daughter is taller than self.'
- c. John_i-un caki_i-ka ka-ss-ta
 John-TOP self-NOM go-PAST-DSE
 'As for John, self went.'

The paradigm in (114) seems to straightforwardly support the prediction made on both Williams's approach and the one pursued here, namely that in configurations

we have adopted is provided.

⁵³I refer the reader back to chapter 3, section 3.3, for a more general presentation of Williams's model.

where locality restrictions on binding are suspended, a broader set of antecedents becomes possible. However, as far as I can see, there are (at least) two other ways of interpreting the paradigm in (114) under the current approach.

One option would be to assume that when it is nonlocally bound, *caki* can either be bound at the interface or at the discourse level. In the former case, *caki* must be bound by a subject. Note that *caki* cannot take the matrix clause object in (114a) as its antecedent, although it can take the more remote subject. In (114b) and (114c), *caki* would be bound at the level of discourse. The condition imposed on discourse binding of *caki* would then be that the antecedent be a topic, which is well in line with the notion of *point of view*, commonly employed in the discourse based accounts of long-distance anaphora.

Another option would be to treat all cases in (114) as cases of nonlocal binding at the interface, while keeping the assumption that *caki* must be subject-oriented. The examples in (114b) and (114c) would seem at first to pose a problem for this view since in these sentences the binder is clearly not a subject. A possibility of bringing these in accord with the requirement on subject-orientation arises on the analysis of these constructions proposed by Gill (1999). Gill (1999) takes as a starting point the observation that the East-Asian languages noted here which apparently allow A'-binders are all languages which have double nominative constructions. In (115), two nominative marked NPs occur with a one-place predicate. Note also that the first nominative NP can readily bear the topic marker *-nun*.

- (115) a. John-i/un ton-i issta.
 John-NOM/TOP money-NOM exist
 'John has money.'
- b. Mary-ka/nun meri-ka norahta
 Mary-NOM/TOP hair-NOM yellow.is
 'Mary's hair is yellow.'

Gill (1999) then argues that cases like (114b) and (114c) where the anaphor is apparently bound by an A'-element are underlyingly double nominative constructions. In other words, the topic that binds *caki* in these instances is always in a legitimate argument position.

We therefore have three analyses of the Korean binding facts presented here at our disposal that are in accord with the model of anaphoric binding argued for here. The first one would treat all cases in (114) as instances of binding at the interface, the availability of A'-elements as possible antecedents resulting from the fact that no ban against A'-binders holds at this level. The second approach would analyze (114a) as involving binding at the semantic interface, while anaphors in (114b) and (114c) would be discourse bound. Finally, the third approach would share with the first one the assumption that all binding relations in (114) are established at the interface, but would treat cases in (114b) and (114c) as instances of A-binding. This is made possible by reanalyzing these examples as underlyingly double nominative constructions. There might be a reason however to favour the latter two analyses over the first one.

Note that only on the last two approaches would the anaphor be subject-oriented. This assumption cannot be adopted on the first approach since we need to rule in (114b) and (114c). It therefore becomes puzzling on the first account why the matrix clause object is not a potential antecedent in (114a).⁵⁴ Observe that an object *can* bind *caki* in a local binding configuration (from O'Grady (1987)):

- (116) Nay-ka John-eykey caki-lil kewul-lo pichwuepoyecwu-ess-ta.
 I-NOM John-DAT self-ACC mirror-in showed
 'I showed John himself in the mirror.'

In light of these considerations, it seems that the first approach must be abandoned in favour of one of the latter two. I won't discuss here which of these two remaining analyses might be preferred, as either option is compatible with our view of anaphoric binding. The Korean facts thus seem to mirror the English binding patterns, modulo the fact that English does not allow double nominative constructions. In both languages, there is no A'-binding either in local or in nonlocal configurations, though for different reasons. In local contexts, A'-elements are not possible antecedents due to the nature and timing of Spell-out. In nonlocal configurations, it is the subject-orientedness that makes A'-elements ineligible as potential antecedents.

The Korean/Japanese facts thus cannot be used to convincingly argue that loosening of locality restrictions on binding makes available a broader range of antecedents, as predicted on both Williams's (2003) model and the one argued for here. For English it is even clearer that although anaphors can find antecedents outside the local binding domain, they do not allow A'-elements as potential binders. What is more, both Korean and English seem to show that, rather than being expanded, the set of possible antecedents in fact becomes more limited. Note that this state of affairs is problematic on Williams's view. In Williams's system, an anaphor which is introduced at a level *X* will take as possible antecedents elements that are defined at that level. However, if the same anaphor can be also bound at the next higher level, *X+1*, it should take as antecedents all elements available at the level *X*, plus all the new elements defined at the level *X+1*. Crucially, an anaphor which is bindable at several levels is not expected to lose potential antecedents along the way. However, this precisely seems to be the case in English and Korean. It remains to be shown therefore how the observed binding patterns might be captured on Williams's model. On the other hand, I have suggested that, on the approach pursued here, the inaccessibility of objects as potential binders might be attributed to the fact that the anaphor in nonlocal binding configurations is subject-oriented. This same restriction rules out A'-binders in the cases under consideration.

The restrictiveness of English anaphors raises further concerns on Williams's view. Namely, the question arises why English anaphors do not behave like Korean ones in allowing A'-antecedents since both languages allow binding relations to be established across relatively large distances. The account of the Korean pattern that

⁵⁴The object is erroneously marked as a possible antecedent in Williams (2003).

Williams (2003) puts forth goes as follows. *Caki* can be bound across a tensed-clause boundary. Therefore, “*caki* must be an anaphor introduced at a late level, perhaps SS or FS, the levels at which tensed clause embedding takes place. As a SS or FS anaphor, it will take as its antecedents the elements that are developed at that level, among them Topic and the Focus of the utterance.” (Williams, 2003, p. 86). It is thus not surprising that *caki* allows A'-antecedents.⁵⁵ It is surprising though that English anaphors do not, considering that they can also be bound across a tensed-clause boundary. In fact, the English pattern is even more puzzling in light of the fact that in Williams's system, *John and Mary* in (117) cannot bind the anaphor from the canonical subject position (SpecTP on standard approaches). Since the binding relation crosses the clausal boundary, it must be established at a later level, with *John and Mary* mapped onto a position made available at that level (possibly a Topic position). The same goes for the binding relation between *John* and *caki* in (114a), though this seems harmless in the case of Korean since *caki* does apparently allow Topics as antecedents.

- (117) John and Mary think pictures of each other are in the post office. (from Williams (2003))

We might suggest that binding in (117) is done at a level between the one where tensed-clause embedding takes place and the one where topics/*wh*-phrases (i.e. the standard A'-elements) are introduced, but this kind of view would in my opinion raise issues regarding the explanatory adequacy of Williams's approach, and would certainly undermine his correlation regarding the distance crossed and the choice of antecedent.

A closer investigation of the Korean(/Japanese) binding patterns has thus revealed that the behaviour of anaphors in these languages cannot be used to convincingly argue for a correlation between the locality of the binding relation and the available set of antecedents. We have yet to find a language where anaphors bound across clausal boundaries would allow A'-binders as antecedents, as predicted to be possible on both Williams's approach and the one pursued here. A possible candidate might be Latin, which is also discussed in Williams (2003). Citing the facts and analysis reported in Benedicto (1991), Williams notes that the Latin *se* anaphor has both a greater scope and a greater class of possible antecedents than standard anaphors. As illustrated in (118), reflexive binding of *se* (which is in its dative form here) can penetrate into finite clauses. That reflexive binding can target A'-antecedents is argued on the basis of examples like (119) (all examples from Williams (2003), citing Benedicto (1991)).⁵⁶

⁵⁵Recall that Williams (2003) dispenses with the notion A and A'-position, which have no status in his system. I will however continue to use these terms for ease of exposition, as Williams himself does.

⁵⁶According to Williams, Benedicto notes that passive *by*-phrases cannot normally bind reflexives. If so, then the fact that the *by*-phrase in (119b) can serve as an antecedent suggests that it can do so solely by virtue of its role as a topicalized NP, which is also consistent with its surface position.

- (118) Cicero_i effecerat [ut Quintus Curius consilia
Cicero.NOM achieved COMP Quintus Curious.NOM designs.ACC
Catinilina sibi proderet]
Catalina.GEN REFL-DAT reveal.SUBJ
'Cicero had induced Quintus Curius to reveal Cataline's designs to him.'
- (119) a. Canum_i tam fida custodia quid significat aliud
dogs.GEN such trusty watchfulness.NOM what mean else
nisi [se_i ad hominem commoditates esse generatos]?
except REFL-ACC for men.GEN comfort.ACC be created.INF
'The trusty watchfulness of the dogs, ... what else does it mean, ex-
cept that they were created for human comfort?'
- b. A Caesare_i ulade liberaliter inuitor [sibi_i ut sim
by Caesar.ABL very generously invited REFL-DAT COMP be.SUBJ
legatus]
legate.NOM
'Caesar most liberally invites me to take a place on his personal staff.'

If Williams's interpretation of the Latin facts is correct, then we have evidence that A'-binders should not categorically be ruled out in nonlocal binding configurations.

7.7.3 Long-distance scrambling

Recall again the contrast between clause-internal and long-distance scrambling in Japanese. Only clause-internal scrambling can feed anaphoric binding. The relevant examples are repeated below.

- (120) a. ?Karera_i-o [otagai_i-no sensei]-ga <karera-o> hihansita (koto).
they.ACC [each.other teacher].NOM criticized fact
'(The fact that) them, each other's teachers criticized.'
- b. *Karera_i-o [otagai_i-no sensei]-ga [Hanako-ga <karera-o>
they.ACC [each.other teacher].NOM Hanako.NOM
hihansita to] itta (koto).
criticized COMP said fact
'(The fact that) them, each other's teachers said that Hanako criti-
cized.'

I have argued that the contrast observed in (120) reduces to the fact that long-scrambling targets a position in the matrix C-domain. As a result, a phase boundary intervenes between the scrambled phrase and the anaphor and the binding relation cannot be established. Note again that this rules out only the local binding relation. We also want to know why the long-scrambled phrase cannot nonlocally bind the anaphor.

Recall now our discussion of the binding patterns in Korean. Since Japanese has been argued to pattern with Korean in the relevant respects, I will assume that our conclusions regarding Korean carry over to Japanese. Now I noted that there were three possible analyses of the Korean/Japanese nonlocal binding data. On the first approach, nonlocally bound anaphors in Korean/Japanese would in principle allow A'-binders. On the latter two approaches, anaphors in these contexts would be subject-oriented. We have already seen some evidence speaking against the first option. Cases involving long-distance scrambling offer further evidence to this effect. If A'-binders were in general available in Japanese as potential antecedents in nonlocal binding configurations, then the ungrammaticality of (120b) becomes completely mysterious. We would expect that the A'-moved phrase should be able to bind the anaphor in (120b). Since it cannot, then it must be the case that A'-binding is not generally allowed in Japanese even in nonlocal binding configurations, and thus the first potential analysis of apparent cases of A'-binding by topics must be discarded. If, on the other hand, long-distance bound anaphors in Japanese are subject-oriented, then the grammaticality status of (120b) is expected since the scrambled phrase does not occupy the canonical subject position. Note further that (120b) cannot instantiate a concealed double nominative construction, as argued for apparent cases of A'-binding in Korean. Such an option is categorically ruled out on the approach pursued here since it would require that the landing site of a long scrambled phrase be below the C domain, a clear instance of improper movement.

The full account of the ungrammaticality of (120b) under the current approach would then go as follows. The long scrambled phrase lands in a position in the matrix C domain. From this position it cannot locally bind the anaphor since there is a phase boundary intervening, namely C. Recall that even if the long-scrambled phrase lands in SpecCP, (120b) is still predicted to be illicit since no binding relation can be established with a phrase sitting in the edge of a phase. For the same reason, the scrambled phrase cannot bind the anaphor from the intermediate SpecAspP, where the scrambled phrase is forced to stop off on its way to the C domain. As a result, all instances of local binding are ruled out. A nonlocal binding relation between the scrambled phrase and an anaphor in the matrix clause cannot be established because nonlocally bound anaphors are subject-oriented and the scrambled phrase is not, in fact cannot on our view, land in the canonical subject position.

7.8 Binding into PPs

This section focuses on anaphoric relations in configurations involving prepositional phrases. So far I have set PP-environments aside. However, it is virtually impossible to discuss Condition A effects without looking at PPs, and in fact the attentive reader might have observed that several examples discussed in previous sections, such as those in (121), featured anaphors embedded in PPs.

- (121) a. John_i sent a letter to himself_i.

- b. John_i seems to himself to be smart_i.

Thus, an account of these examples and Condition A effects more generally cannot be complete without discussing PP-anaphora. This section therefore investigates the nature of the binding relation between anaphors embedded in PPs and their antecedents, with the goal of determining whether anaphors in these environments are locally or non-locally bound.

In the literature on binding phenomena, typically two classes of prepositions are distinguished, the main distinguishing criterion concerning the distribution of anaphors and bound pronominals. When the preposition belongs to the class of what I will call *functional* Ps, only anaphors are allowed, as shown in (122). In this respect, environments featuring functional Ps pattern with canonical local binding contexts, such as those in (123).

- (122) a. John_i always talks about himself_i/*him_i.
b. John_i seems to himself_i/*him_i to be smart.

- (123) a. John_i praises himself_i/*him_i.
b. John_i considers himself_i/*him_i to be smart.

However, both anaphors and bound pronominals can appear as complements to spatial PPs in English. Some illustrative examples are provided in (124).⁵⁷

- (124) a. John_i saw a snake near himself_i/him_i.
b. John_i looked around himself_i/him_i.

I will discuss each of these classes in turn. The conclusion we will reach is that anaphors embedded in PPs are always locally bound, regardless of the type of preposition chosen. This means that PPs do not define local binding domains, i.e. they do not constitute phases on our view. Since it is generally agreed in the literature that anaphors which are complements of functional Ps are instances of local anaphora, I will discuss functional Ps only briefly here, and focus rather on the more controversial cases involving spatial Ps.

7.8.1 Functional Ps

As the examples below illustrate, bound pronominals are illicit as complements to functional Ps. In this respect, PPs headed by functional prepositions pattern with other local binding configurations.

- (125) a. John_i sent a letter to himself_i/*him_i.
b. John_i seems to himself_i/*him_i to be smart.

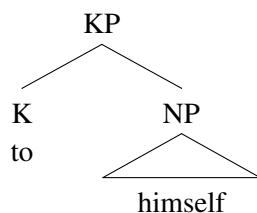
⁵⁷It should be pointed out that the judgements are not always clear and consistent across speakers. This raises many issues which I cannot hope to address within the scope of this study.

- c. John_i talked about himself_i/*him_i.

It is also often noted that prepositions of this type differ from spatial Ps in their theta-marking capacities. Marantz (1984) has argued that when a verb selects for a spatial PP, the whole PP, rather than the NP in it, carries the thematic role of the verb. The NP embedded inside the PP is assigned a θ -role by the preposition. One indication that there is no (direct) semantic relation between the NP-complement of a spatial preposition and the verb is that verbs selecting the spatial role allow all locative prepositions. On the other hand, in cases like (125), the verb selects its argument only via one specific preposition, suggesting a tighter semantic relation. Another way of stating Marantz's generalization, according to Reinhart and Reuland (1993) (see also Hestvik (1991)), would be to say that in (125) the preposition and the verb necessarily form a complex thematic unit selecting the NP. In general, such prepositions seem to have limited semantic import and appear to act simply as Case-markers.

These observations might be taken to suggest that functional prepositions simply spell out case layer(s), and are not in fact instantiations of category P (see Ross (1967) for the origins of this idea). I have already been assuming that there is a K(ase)P within the nominal functional sequence. A preposition like *to* in English might then be treated as the Spell-Out of this functional layer. On this view, the structure of the phrase *to himself* as in (126) would look as follows:

- (126) John talks to himself.



The K-part of this phrase would be spelled out by the preposition *to*, while the NP would be spelled out by *himself*. Arguably, there could be many more functional layers present in the nominal sequence, and the KP might itself be decomposable into a sequence of case layers as suggested by Caha (2009). The logic of the solution would however remain unaltered and I will stick to the maximally simplified representation for the sake of clarity. Note that this type of analysis presupposes that lexical insertion takes place post-syntactically, with lexical items spelling out chunks of syntactic structure. The present dissertation adopts such a view of lexical insertion. The details of the Spell-Out procedure will be discussed in the following chapter.

On the analysis of functional Ps just sketched, there would be no category P in the structural representation of (126). The lexical item *to* would in fact spell out a projection in the nominal functional sequence. The lack of the category P would then straightforwardly account for Marantz's (1984) observation that there is a more direct

semantic relation between the verb and the *to*-complement than between a verb and a complement of spatial PPs. Since there is no category P in the former case, P cannot act as a θ -role assigner.

The suggested analysis might also allow us to deal with the long-standing c-command problem related to examples featuring functional Ps.⁵⁸ Recall that the antecedent must c-command the anaphor in order to bind it. Consequently, under the assumption that *to* heads a prepositional phrase, the fact that *Mary* can bind the reflexive in cases like (127) is surprising since *Mary* appears not to c-command the reflexive.

(127) John talked to $Mary_i$ about herself $_i$.

No problem regarding c-command would arise in cases like (127) if the PP-layer were in fact missing, as suggested here. On this view, *Mary* would not “c-command out of the PP” as it is usually phrased. There would simply be no PP in the first place.

If there is no PP present in examples such as (125) and the like, the issue regarding the phasehood of PP does not arise and we would straightforwardly predict that the binding relation between the antecedent and the anaphor in these configurations should be a local one. However, it will soon become clear that whether or not an anaphor embedded in a PP is available for purposes of local binding cannot be dependent on the possibility of reanalyzing the head of such a PP as a case marker. In the following section we will turn our attention to spatial Ps. There are no obvious reasons to treat spatial Ps as case markers rather than true members of the category P. Yet, the following section will argue that anaphors embedded in spatial PPs can be locally bound.

7.8.2 Spatial Ps

Let us now turn to binding relations between anaphors embedded in spatial PPs and their antecedents. Some representative examples are repeated below:

- (128) a. John $_i$ saw a snake near himself $_i$.
 b. John $_i$ looked around himself $_i$.
 c. John $_i$ pulled the blanket over himself $_i$.

In contrast to functional Ps, these prepositions are not semantically bleached and express various spatial relations. As already noted, they also differ from functional Ps in their theta-marking capacities. I will therefore assume that these prepositions are true exponents of the category P.⁵⁹ Given that anaphors in cases like (128) are

⁵⁸See for instance Reinhart (1983a), Jackendoff (1990), Larson (1990).

⁵⁹I find that using the term *locative-PP* when referring to this class of PPs, as commonly done in the literature, leaves it somewhat unclear whether directional Ps should be included in the same class or not. I opt thus for the notion *spatial* prepositions, as a cover term for locative and directional Ps.

embedded inside PPs, the question arises what the nature of the binding relation between anaphors and their antecedents in (128) is. In other words, do PPs count as local binding domains (i.e. phases on our view), or not? If they do, then the binding relations in (128) are nonlocal in nature. If they do not, then we are dealing with local anaphora also in these environments.

There is no consensus in the literature regarding this issue. Reinhart and Reuland's (1993) influential theory of binding predicts that anaphors embedded in spatial PPs are instances of nonlocal anaphora. Their main piece of evidence in support of this prediction is the availability of both anaphors and bound pronominals in these environments, as illustrated below. Recall that noncomplementarity in the distribution between anaphors and bound pronominals is typically seen as one of the hallmarks of nonlocal binding relations, following the work of Lebeaux (1984).

- (129) a. John_i saw a snake near himself_i/him_i.
 b. John_i looked around himself_i/him_i.
 c. John_i pulled the blanket over himself_i/him_i.

In contrast to Reinhart and Reuland (1993), Hicks (2009) argues that anaphors embedded in spatial PPs are locally bound. This view is shared by the canonical binding theory as well. The canonical binding theory, as already noted, does not recognize the distinction between local and nonlocal anaphora, but subsumes all cases of bound anaphora, including those in (128) under local binding. On these approaches, the observed non-complementarity in distribution between anaphors and pronominals cannot be attributed to the nonlocal nature of the pertinent binding relations, and alternative solutions must be sought. A common strategy in dealing with this issue is to assume that local binding domains for anaphors and pronominals are not the same (see for instance Hestvik (1991), Hicks (2009)). Non-complementarity between anaphors and bound pronominals cannot therefore be treated as a reliable diagnostic in determining the nature of binding relations.

Can some other evidence be adduced in support of either a local or nonlocal nature for the binding relations in question? Recall that Lebeaux (1991) has identified a number of properties that distinguish local from nonlocal binding, noncomplementarity between anaphors and pronominals being just one of these. Other relevant properties are repeated below.

1. nonlocal binding gives rise to strict/sloppy ambiguity, while local binding allows only a sloppy reading
2. local binding obeys the c-command requirement, nonlocal binding does not
3. nonlocal binding allows split antecedents, local binding does not

Applying these tests to PP-anaphora might then help us establish whether we are dealing with a local or nonlocal binding context. According to Hicks (2009), the diagnostics listed indicate that binding in PP-environments is local in nature. Exam-

ple (130a) shows that the anaphor cannot have split antecedents, and (130b) illustrates that the antecedent must c-command the anaphor. Finally, under VP-ellipsis in (130c), only the sloppy reading is available.

- (130) a. *John_i showed Mary_j a snake near themselves_{i+j}.
 b. *John_i's mother found a snake near himself_i.
 c. *John_i found a snake near himself_i, and Bill did too.
 =found a snake near Bill, ≠ found a snake near John

I will accept Hicks's conclusion that the pertinent binding relations are local in nature. However, since the persuasiveness of this argument depends on our understanding of Lebeaux's tests, let me add a few comments regarding the status of these diagnostics and acknowledge some issues that they raise. Why should nonlocal binding show the properties that it does? What is more, given that I have distinguished two types of nonlocal binding relations in the previous section, we might ask whether Lebeaux's tests truly distinguish local from nonlocal binding contexts, or do they rather distinguish two types of nonlocal binding relations. If the latter turns out to be the case, then we don't have an argument here for treating binding into PPs in terms of local binding.

In the preceding sections I have argued that binding relations can be established in the course of the syntactic derivation, at the semantic interface, or at the level of discourse. The local domain, namely the phase, is relevant only for the first type of binding relations. We still however have a two-way distinction between binding relations that do not respect this locality restriction. I have also argued that, from the semantic point of view, discourse binding involves coreference, while the other two types of binding relations involve variable binding. Consequently, a relation between an antecedent and a quantified NP cannot be established in the discourse component. The relevant properties are summarized in the table below:

	domain	semantic relation	c-command required	beyond principles of grammar proper
binding established in syntax	phase	variable binding	yes	no
binding established in semantics	sentence	variable binding	yes	no
binding at the level of discourse	discourse/text	coreference	no	yes

Let us now reconsider Lebeaux's tests in light of this typology of binding relations. According to Lebeaux, local binding obeys the c-command requirement while nonlocal binding does not. However, I have argued that the c-command requirement holds not only of local binding relations but also of binding at the interface (nonlocal), and does not hold of binding at the discourse level (nonlocal). Therefore, while the fact that anaphors must be c-commanded by their antecedents is compatible with the view that they are locally bound, this is in fact not a necessary conclusion.

What about the remaining two tests? Let us take cases of VP-ellipsis first. It is standardly assumed that the sloppy reading involves variable binding, while coreference gives rise to the strict interpretation (see for instance Reinhart (1983b), Heim and Kratzer (1998)). Consider the following example (from Reinhart (1983b)):

- (131) a. Felix hates his neighbours and so does Max.
 b. Max hates Felix's neighbours.
 c. Max hates Max's neighbours.
 d. Felix (λx (x hates x's neighbours)) and Max (λx (x hates x's neighbours))

In the first conjunct, the pronominal is assigned the same referent as *Felix*, regardless of whether it is bound by *Felix* or coreferential with it. The interpretation of the second conjunct is crucial and in fact (131a) is ambiguous between the reading in (131b) and (131c). In addition, the sentence of course has a third reading on which both Max and Felix hate a third person, identifiable from the context, but that reading need not concern us here. In order to license ellipsis, a 'parallelism' requirement must be satisfied, i.e. the elided element must be identical (in certain relevant respects) to the antecedent. On the interpretation in (131b), the pronominal is referential and simply corefers with *Felix*. The parallelism is then satisfied if the pronominal retains the same reference in the second conjunct. This gives rise to the strict interpretation. To derive the reading in (131b), we need to assume that the first conjunct of (131a) contains an open formula *x hates x's neighbours* which is satisfied by Felix in the first conjunct and by Max in the second. This reading is represented in (131d) and is referred to as the sloppy reading.

It is well-known that locally bound anaphors allow only sloppy readings (e.g. Sag (1976), Williams (1977), etc):

- (132) John_i hit himself_i and Bill did too.
 (=Bill hit Bill, *but* ≠ Bill hit John)

On the other hand, Lebeaux (1984) notes that nonlocally bound anaphors give rise to strict/sloppy ambiguity and thus pattern more like pronominals in this respect:

- (133) John_i thought that there were some pictures of himself_i inside, and Bill did too.
 (=pictures of John, *or* =pictures of Bill)

The availability of sloppy reading thus sheds no light on the locality of the binding relations, since it can arise both in local and in nonlocal binding environments. However, the unavailability of strict interpretation in cases like (132) shows that the coreferential strategy is not available for local anaphors.⁶⁰ The observation that the strict reading does not arise in PP-binding contexts thus suggests that these environments constitute local binding domains.

Testing whether split antecedents are possible reveals that PP-binding contexts again pattern with local binding configurations (134b) and contrast with examples involving nonlocal binding (134c):

- (134) a. *John_i showed Mary_j a snake near themselves_{i+j}.
 b. *John_i told Mary_j about themselves_{i+j}.
 c. John_i told Mary_j that there were some pictures of themselves_{i+j} inside.

Why local anaphors disallow split antecedents is an open question. Let me just note that this cannot follow from interpretative facts about anaphors, such as the fact that they must function as bound variables. Pronominals interpreted as bound variables can have split antecedents, as the following examples demonstrate:

- (135) a. Everyone told someone that they should get married. (from Sportiche (1988))
 b. Someone persuaded every kid that they should tell each other a story. (from Hornstein (2001))

In light of these considerations, let us now reconsider the conclusion that was reached regarding the nature of binding into PPs. The c-command requirement, I have argued, distinguishes two types of nonlocal binding relations, rather than singling out local anaphors. Therefore, we can't use c-command as a diagnostic of local binding. The remaining two tests, though not fully conclusive, indicate that anaphors contained in PPs are locally bound.

This conclusion is further corroborated by cross-linguistic data. If the configurations in (128) involved nonlocal binding, then the grammaticality of anaphors in these contexts would hinge on whether or not a particular language allows nonlocally bound reflexives. We already know that English allows reflexives to be nonlocally bound. Thus, whether PPs are opaque for purposes of local binding cannot be de-

⁶⁰This claim has been challenged by Pollard and Sag (1992) who provide the following example to show that the strict reading is licit in a local binding configuration:

- (i) If John doesn't prove himself to be innocent, I'm sure that the new lawyer he hired will.

Until these data are fully understood, any conclusions based on the (un)availability of strict/sloppy readings must be taken with a grain of salt, and should not be used as a sole piece of evidence for treating any binding configurations as either local or nonlocal in nature.

terminated by simply checking the grammaticality status of the sentences in question, but we must resort to diagnostics such as those identified by Lebeaux. The situation is however different in languages that require their anaphors to be locally bound. One case in point is Serbian. As already noted, the Serbian reflexive *sebe* cannot be nonlocally bound (see (136)). Nevertheless, counterparts of English (128) are grammatical in Serbian, with *sebe* in place of the reflexive *himself*, suggesting that PP is transparent for purposes of local binding.

- (136) *David_i je mislio da će članci o sebi_i biti interesantniji.
 David aux thought that will articles about self be more.interesting
 'David thought that articles about himself would be more interesting.'
- (137) a. Džon_i je video zmiju pored sebe_i.
 John aux seen snake next.to self
 b. Džon_i je pogledao oko sebe_i.
 John aux looked around self
 c. Džon_i je stavio knjigu ispred sebe_i.
 John aux put book in.front.of self

The same conclusion can be reached by inspecting the data from German.⁶¹ I have already noted that the German anaphor *sich* cannot be nonlocally bound. A relevant example is repeated in (138). However, it is the reflexive *sich* which is used in PP-binding contexts.⁶²

⁶¹See also Rooryck and Vanden Wyngaerd (2007) for the same argument based on Dutch facts.

⁶²Faltz (1985) and Reuland (2007) also note that pronominals are illicit in these contexts in German:

- (i) a. *John_i sah eine Schlange neben ihm_i.
 John saw a snake near him
 b. *Claudia_i setzte die Pflanze hinter ihr_i/sie_i.
 Claudia put the plant behind her

In Serbian, many speakers find bound pronominals relatively acceptable in PP-contexts. However, the examples become unacceptable if a quantified antecedent is used instead of a referential one, showing that only coreferential strategy, which is outside the purview of grammar proper, is available in these configurations. Recall that this is not true of English, where pronominals contained in PPs can be bound by quantified antecedents.

- (ii) a. ?Džon_i je stavio knjigu ispred njega_i.
 John aux put book in.front.of self
 b. *[Svaki učenik]_i je stavio knjigu ispred njega_i.
 every student aux put book in.front.of him

- (138) Martell_i hofft, dass eine Reportage über ihm_i/*sich_i im Radio
 Martell hopes that a report about him/self in-the radio
 gespielt wird. (from Büring (2005a))
 played becomes
 ‘Martell is hoping that a report about himself is going to be aired.’
- (139) a. John_i sah eine Schlange neben sich_i. (from Faltz (1985))
 John saw a snake near self
 b. Claudia_i setzte die Pflanze hinter sich_i. (from Reuland (2007))
 Claudia put the plant behind self

Putting the observations from cross-linguistic data together with the results we got by applying Lebeaux’s diagnostics, we are led to the conclusion that binding into spatial PPs is local in nature. On the model of binding developed here, this means that PPs are not phases, given that phases constitute local binding domains.

7.8.3 Interim summary

In this section, I have examined the nature of binding relations in configurations where anaphors are embedded in PPs. Two contexts have been distinguished: those involving functional Ps, which allow only anaphors, and those involving spatial Ps, which take both anaphors and bound pronominals as their complements. The conclusion we have reached is that anaphors embedded in PPs can always be locally bound, regardless of the type of the P. I have focused in particular on cases involving spatial Ps since the nature of binding into this type of PPs is subject to debate. The conclusion that anaphors embedded inside spatial PPs are instances of local anaphora has been reached on the basis of two kinds of considerations. First of all, the results obtained by applying Lebeaux’s diagnostics for distinguishing local from nonlocal anaphora speak in favour of such a conclusion. The crosslinguistic data provide further support. Namely, in contrast to English, in many languages anaphors must be locally bound and cannot be coreferential with a nonlocal antecedent. The fact that in such languages anaphors can nevertheless surface as PP complements strongly suggests that binding into PPs is local in nature.

The conclusion reached in this section bears significantly on the issue of apparent noncomplementarity between pronominals and anaphors in PP-binding contexts in languages like English. An analysis of the noncomplementarity issue based on the assumption that binding into PPs is nonlocal in nature, such as the one advocated by Reinhart and Reuland (1993), cannot be adopted and an alternative account must be sought. Addressing this question requires making explicit claims about the status of Condition B. So far I have set Condition B effects aside and focused exclusively on the distribution of anaphors. The following chapter will offer some discussion of Condition B effects, including the apparent non-complementarity issue in spatial PPs. The main focus however will remain on Condition A effects, and a more extensive study of the distribution of pronominals will be left for future research.

7.9 Hicks (2009)

In this chapter I have argued that the local domain for anaphoric binding reduces to the syntactic notion of the phase, with the nature and timing of Spell-Out again playing a crucial role, as in our account of improper movement. That the phase is the relevant local binding domain is not a novel idea though. It has been argued for most systematically by Hicks (2009). The precise implementation of this idea developed here differs however in certain important respects from Hicks's. I will end this chapter by briefly reviewing Hicks's proposal.⁶³

Hicks (2009) argues that features which drive the narrow-syntactic derivation are of two types: morphosyntactic and semanticosyntactic features. Each of the two interfaces interprets only a single type of feature: morphosyntactic features are interpreted at PF, and semanticosyntactic features at LF. This is an important assumption around which Hicks builds his analysis of binding relations. Hicks then further assumes that the points in the derivation where the interfaces read off the features interpretable to them need not be the same for both PF and LF, introducing thereby a distinction between LF- and PF-phases. At PF-phases, the PF reads off morphosyntactic material, while at LF-phases, it is the semanticosyntactic features that are interpreted. The distinction between LF- and PF-phases is crucial for Condition A effects, and even more so in accounting for the (non)complementarity in distribution between anaphors and pronominals, the issue we will return to in more detail in the following chapter.

Hicks then argues that Condition A effects are entirely reducible to feature-agreement. Anaphoric binding is assumed to involve an agreement relation between an anaphor bearing an unvalued feature, and an antecedent capable of valuing it.⁶⁴ Binding viewed in this way becomes thus just another instance of feature valuation under Agree in syntax. If anaphor binding is determined by Agree, then the locality restrictions on binding relations are simply a reflex of locality restrictions imposed on the operation Agree, which in the framework of Chomsky (2000, 2001) is constrained by phases. It then follows that the local domain for anaphoric binding is the phase. Since the relevant feature belongs to the type of semanticosyntactic features, it is sensitive to LF-, but not PF-phases. In other words, the anaphor and its antecedent must occupy the same LF-phase so that an Agree relation can be established. *C* and *v* are assumed to be LF-phases (and PF-phases), thus a local binding relation cannot be established across a CP or a vP.

In many cases, Hicks's account makes the same predictions as the model developed here. I won't therefore step through all the relevant contexts in order to show how these are handled on Hicks's view, rather I will limit my attention to those con-

⁶³See also Hicks (2009) for a brief presentation and critical review of some other related proposals I will not discuss here.

⁶⁴The identity of the assumed feature and the precise mechanisms of feature valuation will be discussed in the following chapter.

figurations where Hicks's proposal makes different predictions from mine.⁶⁵

Consider first Hicks's treatment of binding into picture-DPs. The proposal assumes that the internal structure of DPs is as given in (140). A DP-internal subject, when present, is base-generated in Spec of *nP*, and moves to the Spec of NumP, where it receives genitive case.

(140) [DP [NumP Bill's [Num] [nP <Bill> [NP picture]]]]

Hicks then argues that DPs are not LF-phases (only PF-phases). In addition, by analogy with the light *v*, he assumes that *nP* is an LF phase when it introduces a subject, but not otherwise. These assumptions lead to the following set of predictions. Consider first cases where a DP-internal subject is present.

(141) John [_{vP} <John> likes [DP [NumP Bill's [nP <Bill> picture of himself]]

The anaphor must find an antecedent within its minimal LF-phase, *nP*. Consequently, *Bill* is the only possible antecedent, since *John* does not enter the derivation until the next LF-phase, the matrix *vP*. On the other hand, in subjectless picture-DPs, such as (142), the local binding domain for the anaphor expands to *vP*, since *nP*, lacking a subject, fails to constitute an LF-phase. In (142) therefore, *John* is a possible local antecedent for the anaphor.⁶⁶

(142) John [_{vP} <John> likes [DP [NumP [nP picture of himself]]

In contrast to Hicks, I have argued that there are no phasal heads in the nominal sequence, a conclusion we have reached on the basis of facts relating to improper movement. Rather, the clausal phasal heads, C and Asp on our view, can make parts of the nominal structure opaque for further syntactic operations. On the analysis pursued here, neither DP, nor any DP-internal projection constitutes a phase. Thus, any noun phrase occupying a position in the same AspP where the picture-DP is merged is a potential local antecedent for an anaphor embedded inside the picture-DP. For the cases at hand, this leads to different predictions only regarding the possibility of establishing a binding relation between the clausal subject and the anaphor in the presence of a possessor. As it was already pointed out, given our current set of assumptions, the analysis pursued here, in contrast to Hicks's, predicts that a binding relation of that kind should be licit. However, once I have outlined my view regarding how binding relations are encoded, this discrepancy between Hicks's proposal and mine will be removed, and both approaches will end up predicting the pertinent

⁶⁵The following chapter investigates mechanisms responsible for encoding binding relations and addresses the issue of (non)complementary distribution between anaphors and pronominals. There will be considerable differences between Hicks's and my way of handling these issues.

⁶⁶Hicks argues that in the same configuration the anaphor can also be logophoric. The factors regulating the distribution of logophoric uses of reflexives remain however largely unclear. I will set this point aside for now.

binding relation to be ungrammatical.

A more significant difference between Hicks's approach and my own concerns cases where the picture-DP is displaced from its base-generated position. Recall that on the analysis outlined here, parts of the DP are spelled out when a clausal phasal head is merged. Thus, in cases like (141) and (142), the anaphor will become inaccessible for further syntactic operations as soon as AspP is merged on top of the ν P. Thus, an anaphor in a complement position of a noun cannot be bound by an antecedent outside the AspP, even if the entire DP moves to a position within the same phase domain as the potential antecedent. On Hicks's view however, the binding possibilities in picture-DPs are not altered by displacement of that DP, as long as both the antecedent and the picture-DP are within the same phase. One relevant set of cases to look at in this regard are configurations discussed by Barss (1986), repeated below.

- (143) a. John_i wondered which picture of himself_i Mary saw <which picture of himself>.
 b. John_i wondered which picture of himself_{i/j} Bill_j saw <which picture of himself>.

The approach pursued here predicts that the binding relation between *Bill* and the anaphor, and the one between *John* and the anaphor in (143b) are not of the same nature. *Bill* can locally bind the anaphor from its base-generated position, ν P, given that it occupies the same phase domain and no part of the picture-DP is spelled out at this point in the derivation. However, binding between *John* and the anaphor can only be nonlocal, since by the time the picture-DP reaches the Spec of the embedded CP, the anaphor will already have been spelled out and be inaccessible to further syntactic operations. This has been discussed in detail in section 7.5. There I have argued that this prediction of the current approach is supported by comparative data. While in English, where anaphors can be both locally and nonlocally bound, it is difficult to determine the nature of a particular binding relation, this is not the case in languages where anaphors have to be locally bound. Thus, German and Serbian for instance distinguish the two contexts in question. In cases like (143b), an anaphor in these languages can only be bound by the subject of the embedded clause, but crucially not by the subject of the matrix clause, suggesting that the matrix clause subject and the anaphor are not in a local binding configuration.

- (144) a. Hans_i fragte [welche Bilder von ihm_i/*sich_i] ich gesehen hatte.
 Hans asked which pictures of him/self I seen had
 (from Buring (2005a))
 b. [Das Buch über sich_i/*sich_j] glaubt der Gernot_j mag der Ulrich_i.
 the book about self believes the Gernot likes the Ulrich
 (from Kiss (2001))

- (145) a. David_i je pitao [koju sliku sebe_{*i/j}] je Petar_j izgubio.
 David aux asked which picture self aux Peter lost
- b. *[Koji članak o sebi_i] je David_i rekao da su novine
 which article about self aux David said that aux papers
 objavile?
 published

On Hicks's view however, both *John* and *Bill* can serve as local antecedents in (143b). Recall that there are no LF-phasal boundaries in subjectless picture-DPs. Thus, the anaphor can be locally bound by *Bill* within the *vP* phase of the embedded clause. If this option is not chosen, the picture-DP, by virtue of its *wh*-feature requirements, moves to the Spec of the embedded CP. At this point, the anaphor, within the CP-edge, is once again in a local binding configuration, this time with *John*. The binding patterns in languages like German and Serbian thus remain unaccounted for on this approach.

The present chapter also addresses in some detail two issues which receive only scant attention in Hicks (2009). The first issue concerns the nature of nonlocal binding. Hicks merely notes that, on their logophoric use, reflexives share the feature composition of pronominals, i.e. come with the valued version of the pertinent feature, and thus behave essentially like pronominals in relevant respects. The second issue that is only briefly discussed in Hicks's work regards the ban on A'-binders. Recall from section 7.7 that an A'-moved antecedent cannot bind an anaphor from the derived position. This is shown for *wh*-movement in (146) (the second example is from Hicks (2009)).

- (146) a. *John wondered [which man]_i pictures of himself_i convinced Mary
 that she should investigate <which man>.
- b. *Who_i does each other_i's mother love <who>?

The would-be antecedents, *which man* and *who* in the examples above, c-command the anaphor in a sufficiently local configuration, since there are no intervening LF-phase boundaries between the anaphor and the surface position of these *wh*-phrases. Yet, the binding relation is not successful. Hicks speculates that the ungrammaticality of examples like (146) has nothing to do with constraints on anaphor binding. Rather, he draws a parallel between these cases and configurations involving a weak crossover violation, as in (147), where a variable is bound by an element which moves across it, but whose trace the variable doesn't c-command.

- (147) ??Who_i does their_i mother love <who_i>?

He then suggests that whatever explanation turns out to be correct for weak crossover effects will carry over to examples like (146). However, the ungrammaticality associated with weak crossover effects is typically weak, and contrasts with

sharper ungrammaticality of cases like (146).⁶⁷ This suggests that the status of examples like (146) cannot be attributed solely to a weak crossover violation.

Facts from German cast further doubt on the suggestion that the ban on A'-binders might be reduced to a weak crossover violation. Recall that short *wh*-movement and topicalization in German do not display weak crossover effects, as illustrated in the following examples (from Grewendorf (2005)).

- (148) a. Wem_i hat seine_i Tante einen US-Aufenthalt finanziert?
 who.DAT has his aunt a US-stay paid
 b. Den Studenten_i hat seine_i Mutter finanziell unterstützt.
 the student.ACC has his mother.NOM financially supported

If unavailability of A'-binders is due to a weak crossover violation, then we expect *wh*-moved and topicalized phrases to be able to act as potential antecedents in German since these movement operations do not display weak crossover effects in this language. However, this prediction is not borne out, as illustrated for *wh*-movement in (149) and topicalization in (150) (from Grewendorf (2003, 2005)).

- (149) a. *Welche Studenten_i hat [der Professor von sich_i]
 which students.ACC has the professor.NOM of self
 <welche Studenten> unterstützt?
 supported
 b. *Welchen Studenten_i hätte [eine Wohnung für sich_i] ein
 which student.ACC would.have an apartment for self a
 Vermögen gekostet?
 fortune cost
- (150) a. *Den Studenten_i hat [der Professor von sich_i]
 the students.ACC has the professor.NOM of self
 <welche Studenten> unterstützt.
 supported
 b. *Den Studenten_i hätte [eine Wohnung für sich_i] ein
 the student.ACC would.have an apartment for self a
 Vermögen gekostet
 fortune cost

I conclude therefore that a potential weak crossover violation is not enough to exclude elements in A'-positions from the class of possible antecedents. In section 7.7, I have suggested how the ban on A'-binders might be restated on a phase-based approach to anaphoric binding.

Overall, although Hicks's proposal and the one argued for here share some important assumptions, they nevertheless differ in details of implementation, which at

⁶⁷Hicks acknowledges this in a footnote, though he expresses some uncertainty regarding the precise strength of ungrammaticality in cases like (146).

times lead to distinct empirical predictions. Further similarities and differences between Hicks's approach and my own will be discussed in the following chapter.

7.10 Summary

In this chapter, I have argued that the phase constitutes the relevant local domain for anaphoric binding. Theoretical benefits of such an approach are easily discernible. Locality domains for movement and binding are thereby unified, and no binding-specific locality domain needs to be postulated. This yields a more parsimonious theory, with fewer technical devices necessary, which is in line with the methodology and aims of the Minimalist Program. It has been shown that such an approach is not only theoretically appealing, but it also goes a long way in capturing the empirical facts regarding anaphoric binding. The conception of phases employed throughout this chapter has been the one motivated in the preceding chapters and used there to account for improper movement phenomena. One significant departure from the standard view of phases concerns the way Spell-Out ships off parts of syntactic structure to the interfaces. In particular, I have argued that once a phasal head is merged, all features/projections lower on the functional hierarchy than the merged phasal head undergo Spell-Out, rather than the complement of the phase, as generally assumed. In the realm of binding, this leads to some novel empirical predictions, as we have seen. Thus, parts of the DP will be spelled out incrementally in the course of the syntactic derivation, making an anaphor embedded inside it accessible to antecedents outside the DP when that DP is in its base position, but not after it has been displaced (above AspP). I have argued that anaphors in languages such as German and Serbian show precisely the predicted pattern of distribution, supporting thereby the current proposal. In English, the picture is often blurred by the fact that English reflexives can be both locally and nonlocally bound.

As we have seen, the distinction between local and nonlocal binding relations is important to maintain. As pointed out by Reinhart and Reuland (1993), the failure to do so in the canonical binding theory has led to many misconceptions concerning the syntactic distribution of anaphors. While I do share this view, I disagree with Reinhart and Reuland regarding both the identity and the nature of nonlocal binding relations. In particular, I argue that treating the phase as the local domain makes the correct split between local and nonlocal binding contexts. In addition, in contrast to Reinhart and Reuland, I don't assume that all nonlocal binding relations are subject solely to the principles of the discourse component. Rather, I make a distinction between two types of nonlocal relations: those established at the interface, and those involving discourse binding. Semantically, these two binding relations differ significantly; the first one involving variable binding and the latter coreference. Further details concerning the nature of various binding relations will be clarified in the following chapter.

In this chapter, I have also discussed how the ban on A'-binders might be restated

within the current approach. On the view pursued here the relevant examples are ruled out by prohibiting that a binding relation be established between an anaphor and a phrase occupying a phase edge position. Consequently, a phrase in the Spec of CP, the canonical A'-position, cannot bind an anaphor within the same clause.

This chapter has argued that treating the phase as the relevant locality domain for binding relations is both theoretically appealing and empirically plausible. The following chapter will attempt to provide an answer as to why phases should play a role in regulating the distribution of anaphoric expression. As we will see, the relevance of phases reduces to the fact that a syntactic dependency must be established between an anaphor and its antecedent, which then translates into a particular interpretative dependency at the interface.

Chapter 8

Encoding binding relations

In the previous chapter, I have focused on showing that we do not need a notion of local domain specific to the binding theory. Rather the locality domain relevant for movement operations is also the domain within which local binding relations are established, namely the phase. If so, then Binding Condition A should be reformulated in the following way:

- (1) *Binding Condition A*
An anaphor must be bound within the phase.

So far I have offered no explanation as to why the phase should be the relevant locality domain for anaphoric binding. The answer to this question can be found in this chapter, where I will investigate mechanisms responsible for encoding binding relations. In doing so, it will be shown that Condition A need not be stated, but that its effects can be reduced to more general principles of grammar. In particular, I will argue that an Agree-relation is established between the anaphor and its antecedent, following some earlier accounts that attempt to do away with the binding conditions. Locality restrictions on anaphoric binding thus follow from locality restrictions imposed on the operation Agree, which on the current view is constrained by the PIC. The syntactic dependency between the anaphor and its antecedent then translates into a bound variable dependency at the interface.

The proposal will also take us into discussion of Condition B effects, though a full and detailed account of the distributional properties of pronominals will not be attempted here. We will focus on those distributional patterns that fall out as a consequence of our analysis of Condition A effects. Ways of handling other aspects of the distribution of pronominals will only be sketched out and details of implementation left for further research.

The chapter is organized as follows. I first discuss some earlier proposals regarding possible ways of encoding binding relations. An analysis of anaphoric binding is then developed in section 8.2. I start out by spelling out my assumptions regarding how lexical insertion proceeds, which will be shown to have significant implications

for the shape of the solution adopted. I then argue that a binding relation between an anaphor and its antecedent involves an Agree-operation established in the course of the syntactic derivation and discuss the nature of the features involved in some detail. The proposed model is then further refined in order to capture the distributional properties of anaphoric expressions in languages that make a distinction between two kinds of anaphors. A language of this type that we will focus on is Dutch. In section 8.3, I then turn to the issue of (non)complementarity between anaphors and bound pronominals in various kinds of binding environments, investigating what effects the proposed analysis of Condition A has on the distribution of pronominals. The chapter ends with a summary of the conclusions reached.

8.1 Previous proposals

There have already been a number of proposals as to how binding relations could be encoded, consistent with the view that Condition A effects are determined in the course of the syntactic derivation. Two general approaches have emerged: encoding binding relations via Move, or via Agree. Some representative examples of these two types of approaches will be briefly reviewed in this section. The rest of this chapter will then argue for a particular implementation of an Agree-based approach.

8.1.1 Binding relations derived by Merge/Move

8.1.1.1 Hornstein (2001)

Hornstein (2001) aims to eliminate the binding theory by recruiting the theory of movement. He suggests that the derivation of a sentence like *John likes himself* starts by merging *John* with *self* in the base-position. *John* receives the internal theta role in the object position of the verb. It then raises to Spec VP where it gets the external theta role of *like*. Note that the analysis abandons the theta criterion since a DP is permitted to receive more than one theta-role. Finally, *John* raises to SpecIP where it checks case and EPP features. The lower copies of *John* are deleted for usual reasons of linearization at PF. As a result, the bound morpheme *self* is left unsupported, triggering the insertion of a pronominal. The pronominal part of the reflexive is effectively the ϕ -features of the moved DP spelled-out. On Hornstein's view, bound pronominals are not lexical elements and therefore not present in the numeration, but are grammatical formatives that are added in the course of the derivation.¹ *Self* is

¹This appears to violate Chomsky's (1995) Inclusiveness Condition, which states that objects with semantic import that are not present in the numeration cannot be introduced in the course of the derivation. Hornstein (2001) however points out that no violation of the Inclusiveness Condition is incurred if the proposed grammatical formatives are semantically inert. Since bound pronominals are essentially bundles of ϕ -features, he concludes that these ϕ -features have no semantic import. Bound pronominals are then assumed to be crucially different from referential pronominals in that the latter ones are not grammatical formatives

present for Case-checking reasons, allowing a nominal to bear more than one Case. Since anaphors are simply residues of A-movement, it is clear why the same locality domain should be relevant for both movement operations and binding relations: the local domain for the reflexive reduces to the domain in which the antecedent can A-move.

Hornstein derives Condition B effects by assuming that bound pronominals are ‘costly’ and appear only when reflexives are unavailable. The competition is not simply between different morphological forms, Hornstein argues, but rather between derivational alternatives: operations that introduce bound pronominals are less economical than operations which license reflexives. Since reflexives surface in configurations involving movement, this means that the preferred way of encoding referential dependencies is via movement. Hicks (2009) however points out that it is difficult to see what exactly makes the pronominal so costly. The derivation of reflexives also requires insertion of a bound pronominal as the spellout of the movement trace, so it is not entirely clear why this derivation is less costly than the one where only the pronominal is used.

8.1.1.2 Kayne (2002)

Adopting Hornstein’s view that binding relations should be derived via movement and focusing particularly on bound readings of pronominals, Kayne (2002) argues that the reading of (2a) on which *he* takes *John* as its antecedent involves a derivation in which *John* and *he* start out together as part of one ‘doubling’ constituent. The antecedent, *John*, then extracts out of this complex DP and moves into the matrix clause. The derivation is sketched in (2b). The analysis complies with the theta criterion since *John* is not assigned two theta roles. It is the entire doubling constituent that receives a theta role in the embedded clause, while *John* gets a theta role by moving into the subject theta position of *think*. Note also that the pronominal is not inserted in the course of the derivation, as on Hornstein’s view, but is already present in the numeration.

- (2) a. John thinks he is smart.
 b. thinks [John he] is smart → John_i thinks [t_i he] is smart

As Kayne rightly observes, if this were the whole story, then we would expect cases like (3) to be derivable in similar fashion, making it possible for *John* and *him* to corefer.

- (3) *John_i likes him_i.

To rule out Condition B violations such as (3), Kayne therefore introduces a number of additional stipulations. In particular, he assumes that the doubling constituent must move to a licensing position, which is crucially higher than SpecvP, before the

and their ϕ -features are semantically active.

antecedent extracts out of it. Once this movement has applied, *John* in (3) can no longer move to SpecvP to pick up a theta role, given that downward movement is disallowed. (3) therefore reduces to a violation of the theta criterion.

Pursuing this line of reasoning further, Kayne argues that reflexives are licit in configurations such as (3) because the addition of *self* makes the required licensing position available. The presence of the noun *self* yields a possessive-type DP structure, in whose Spec the doubling constituent can be licensed. Since the needs of the doubling constituent are met DP-internally, the doubling constituent can remain inside the VP, making it possible for the antecedent to move into the Spec of vP to receive a theta-role.

Without going into further details, note that this type of analysis offers little hope of capturing locality restrictions on anaphoric relations. Since the movement of the antecedent from within the doubling constituent must be extremely free to allow pronominals to be long-distance bound, it is difficult to see how this movement step can be restricted in cases involving reflexives. The account therefore loses one appealing aspect of Hornstein's proposal, namely the unification of locality restrictions on movement and binding. The issue of locality is further complicated by Kayne's assumption that there are no cases of 'accidental coreference' so that *all* cases of coreference between a pronominal and its antecedent must be derivable in the manner described above. Thus, even examples like (4a) would, on Kayne's assumptions, involve movement of the antecedent out of the doubling constituent. Kayne suggests that in cases such these, the two sentences initially form a single syntactic entity, akin to coordination as in (4b). Nevertheless, locality restrictions on movement would have to be considerably loosened in order to allow movement of the antecedent from inside the doubling constituent in configurations of this type.

- (4) a. John is famous. He's smart, too.
 b. John is famous, and he's smart, too.

8.1.1.3 Zwart (2002)

In an attempt to avoid the problems that Kayne's analysis faces, Zwart (2002) suggests that only binding relations involving anaphors are encoded by merger. Pronominals, on the other hand, are not merged with their antecedents. On his view, (nonaccidental) coreference of α and β arises if and only if α has been merged with β in a structure like (5), where α is the antecedent of the PRONOUN β , and β is the head of the noun phrase XP containing α . PRONOUN is the label for a generic variable referential element, encompassing both pronominals and anaphors.

- (5) [XP [antecedent] [PRONOUN]]

Merging an antecedent with a PRONOUN is what *makes* the PRONOUN reflexive. As an automatic consequence of the merger operation, the PRONOUN acquires the feature [+coreferential]. A coreferential PRONOUN is invariably spelled out at PF

and interpreted at LF as an anaphor. Anaphoricity is thus treated as a property acquired in the course of the syntactic derivation, rather than a lexical feature which is present from the outset. As on Kayne's view, the antecedent then moves out of the complex phrase into a position in which it can get a theta role and Case. Since the target for the noun phrase movement is an A-position (by definition), locality restrictions on anaphor binding are reduced to locality restrictions on A-movement, as on Hornstein's account.

Interestingly, though their analyses are similar in many respects, Zwart and Kayne reach entirely opposite conclusions regarding the availability of accidental coreference. Recall that for Kayne, there is no accidental coreference and any instance of referential dependency between a pronominal and another DP, including cases that span sentential boundaries, involves the formation of a doubling constituent. On the other hand, Zwart assumes that coreference as a function of sisterhood of the antecedent and the dependent applies to anaphor binding only. Hence, he is led to the conclusion that all cases of pronominal binding are instances of accidental coreference. I have already noted some problematic aspects of Kayne's view. Zwart's approach also raises a number of concerns. For one, it is not obvious how examples such as (6), involving quantified DPs as antecedents, should be handled. As already pointed out, binding relations in examples of this type cannot be due to coreference, since quantified DPs are not referential expressions.

(6) [Every participant]_{*i*} thinks that he_{*i*} is a genius.

What is more, even in examples featuring referential expressions, a distinction must be made between coreferential and bound variable interpretation of pronominals, as already discussed in the previous chapter.² It is unclear how this can be achieved on either Zwart's or Kayne's approach, since both proposals treat all cases of pronominal binding in a uniform way. Finally, given how freely 'accidental' coreference is allowed on Zwart's approach, one wonders why this option is not available in cases of local binding such as (7).

(7) *John_{*i*} likes him_{*i*}.

8.1.2 Binding relations derived by Agree

A different line of approach suggested in the literature is to encode binding relations via the operation Agree (Reuland (2001, 2006), Hicks (2009), among others). In the framework of Chomsky (2000, 2001), the application of Agree is a prerequisite for movement. Locality constraints on movement are thus reinterpreted as constraints on Agree. If binding operations are encoded via Agree, then it is again unsurprising that movement and binding should share the same locality restrictions. Intuitively, Agree seems to be an appealing way of encoding anaphor binding relations since an

²See also section 8.1.2.2 of this chapter.

anaphor must share the ϕ -features and reference of its antecedent.

Agree operates between features, typically assumed to be attribute-value pairs. Features can enter the derivation already bearing a particular value, or they can be unvalued. On standard assumptions, an unvalued feature is not tolerated at the interfaces so it must enter into an Agree relation with another feature capable of valuing it. If local binding is reducible to Agree, then anaphors are obvious candidates for bearing the unvalued version of the relevant feature, given that anaphors are dependent on the presence of an antecedent. A syntactic dependency which would value an anaphor's unvalued feature must thus be established between the anaphor and its antecedent before the relevant chunk of syntactic derivation is sent off to the interfaces for interpretation. This is the common intuition behind Agree-based approaches. Though the logic of the proposal is clear and appealing, determining which feature(s) is involved in encoding binding relations is far from a trivial matter. Let me discuss some options that have been suggested in the literature.

8.1.2.1 Referential features

Given that an anaphor and its antecedent share the same referent, a plausible approach would be to posit a referential feature, call it [REF], unvalued on anaphors and valued on pronominals and other DPs. Regarding the values of [REF], these might be thought of as simply integers, corresponding to the indices of pre-Minimalist approaches to binding. This possibility has been explored by Hicks (2005).

The most serious drawback of the proposal relying on the [REF]-feature from an empirical perspective is that it does not extend to cases where the antecedent is a quantified DP. As already noted, quantified DPs are not referential, and thus cannot bear a valued [REF]-feature, yet they can serve as antecedent for anaphors, as (8) shows.

(8) [Every boy]_{*i*} likes himself_{*i*}.

8.1.2.2 Operator and variable features

Hicks (2009) argues that the empirical challenge noted above can be overcome if we were to assume that features involved in binding relations are more closely related to the semantics of the DPs in question. In the spirit of Adger and Ramchand's (2005) proposal regarding the features that give rise to operator-variable dependencies in *wh*-movement constructions, he proposes that operators and variables are encoded syntactically as distinct semantico-syntactic features, [OP] and [VAR] respectively. An element bearing an unvalued [VAR]-feature must enter into an Agree-relation with an element bearing an [OP]-feature.³ The values for [OP] might be \forall and \exists . The values of [VAR] must be able to identify it with respect to other variables, and

³This requirement seems to be relaxed in cases involving referential DP antecedents. Refer to the discussion regarding example (11) for clarification.

to this end Hicks assumes it is simply an integer.⁴ Pronominals and anaphors, as variables, are lexically specified with a [VAR]-feature. Anaphors enter the derivation with an unvalued [VAR]-feature, which is the trigger for an Agree operation. During the derivation of (9), we get the following representations pre- and post-Agree respectively:

- (9) Every boy loves himself.
- a. Every_{[Op:∀] [Var:x]} boy loves himself_[Var:⋯]
 - b. Every_{[Op:∀] [Var:x]} boy loves himself_[Var:x]

Various problems however arise on this implementation of the Agree-based analysis. By assumption, pronominals (and R-expressions) differ from anaphors in having a valued [VAR]-feature. Given that the value of [VAR] is an integer, it cannot be listed in the lexical entry. Hicks thus concludes the following: “What appears to be required is that upon selection of the numeration, a value is selected for [VAR]. In some sense, then, the value of [VAR] on R-expressions and pronouns is simply an instruction to assign an integer upon selection for the numeration.” (Hicks (2009), p. 115). It seems however that we might equally well interpret the unvalued [VAR]-feature of the anaphor as an “instruction to assign an integer”. As far as I can see, the crucial distinction between pronominals and anaphors on Hicks’s view is that with pronominals the integer must be assigned upon selection for numeration, while in case of anaphors this happens in the course of the syntactic derivation. Hence, the approach appears to allow feature-valuation upon selection of the numeration. This is a serious concern since the numeration is not a level of representation, thus as little as possible should happen there. In addition, if pronominals always have the value of [VAR] already assigned upon entering the syntactic derivation, we must somehow ensure that the value of [VAR]-feature covaries with the one of the antecedent in cases where pronominals function as bound variables.

Interestingly, although the approach Hicks proposes accounts for configurations involving quantified-DP antecedents, modulo the worries noted above, it runs into difficulties with referential DPs. The problem is thus converse to the one that the previous solution in terms of the [REF]-feature faced. The question is whether referential DPs bear the Operator feature or not. Recall that pronominals can also function as referential expressions and act as potential antecedents for anaphors.

- (10) He_{*i*} loves himself_{*i*}.

If pronominals, and other referential DPs, bear the [OP]-feature, then the parallel between anaphors and bound pronominals would be lost since pronominals would

⁴It is obvious that the value of [VAR]-feature does not strictly match the value of [OP]-feature. To make this mismatch more plausible, Hicks draws a parallel with the [CASE]-feature. The standard view is that a [CASE]-feature value is assigned as a reflex of ϕ -feature agreement between a head (v or T) and the DP it probes. The relevant head in that Agree operation does not in fact bear a matching valued [CASE]-feature.

have the same feature composition as referential DPs. In order to maintain this parallel, Hicks opts for a different solution and assumes that referential expressions do not bear the [OP]-feature. Since there is no [OP]-feature capable of binding the anaphor's [VAR]-feature, one wonders how the anaphor gets bound by a referential expression. In configurations featuring referential phrases as antecedents, the following features would be involved:

- (11) John_[Var:x]/He_[Var:x] likes himself_[Var:_]

According to Hicks, in (11), the anaphor's variable feature will receive its value from another valued [VAR]-feature. As a result, the anaphor and its antecedent will end up sharing the value for their variable (x), but there will be no operator capable of binding the variable since there is no element bearing the [Op]-feature. Following Heim (1998), Hicks then assumes that in cases such as these, each free index in an LF-representation is assigned a distinct referent by the context. In (11), there is a single free variable index, namely x , and upon interpretation the context assigns an appropriate individual to x , at once for the antecedent and for the reflexive.

Note that this ensures that the anaphor corefers with its antecedent, but not that it is semantically bound by it. In other words, it gives rise only to the coreferential interpretation. Recall that even with referential antecedents, it is possible to find syntactic environments where the coreferential and the bound variable reading can be distinguished. As we have already seen, an example of such an environment is a VP-ellipsis construction. On the basis of such constructions, it can be shown that binding by referential antecedents can involve variable binding. In fact, this is the only interpretation available with locally bound anaphors. Examples such as (12) for instance allow only the sloppy reading (which involves variable binding).

- (12) John hit himself, and Bill did too.
 (=Bill hit Bill, *but* \neq Bill hit John)

This is unexpected on Hicks's approach if the relation between the anaphor and a referential antecedent can only be one of coreference. As a way of dealing with this problem, Hicks suggests that at LF, if the configurational requirements of binding are met (shared [VAR]-feature and c-command), a λ -operator is inserted on the c-commanding DP, resulting in a semantic binding relation. This would be an obligatory interpretative process at LF under the appropriate configurational requirements. Though the suggestion appears to resolve the problem at hand, the proposal loses some of its initial simplicity. Bound variable dependencies are no longer derived simply by mapping [OP]- and [VAR]-features to operators and variables at LF, given that they can be established via the proposed LF-interpretative mechanism also in cases where no [OP]-feature is present in the course of the syntactic derivation. Hence, elements bearing [OP]-features are translated as operators at LF, but not all elements functioning as operators at LF need to bear the [Op]-feature. Note also that in cases like (11), the referential DPs bear the [VAR]-feature but can in fact be interpreted as

operators at LF. As far as I can see, there is therefore no neat correlation between the syntactic [OP] and [VAR] features and operators and variables at LF. One also wonders how the correct interpretation is ensured in cases where an element bears both an [OP] and a [VAR]-feature, as in (9) for instance. Given Hicks's assumption that [OP] and [VAR] are semanticosyntactic features, and that all features of this type are interpreted at LF, it seems that the quantified DP in (9) should be interpreted both as a variable and as an operator. A potential solution to this problem would be to assume a condition that would regulate how features in an Agree-chain are interpreted. A suggestion along these lines has been put forth by Adger and Ramchand (2005) as a way of dealing with the same kind of problem in their analysis of A'-dependencies. They then propose an interface condition which states that interpretable features in an Agree chain are interpreted only once. Such a condition would then ensure that the [VAR]-feature of the quantified DP, though in principle interpretable, is not interpreted when it is a part of an Agree-chain, as in cases like (9). For binding facts that Hicks considers, there might still be a problem though in cases where there is a quantified DP but no anaphor present since no Agree-chain involving [VAR]-features would then be established.

8.1.2.3 ϕ -features

Another suggestion that can be found in the literature is to encode anaphor binding via ϕ -agreement (e.g. Reuland (2001), Rooryck and Vanden Wyngaerd (2010)). For instance, Reuland (2001) argues that the relevant feature distinction between anaphors and pronominals is their ϕ -feature specification, anaphors being underspecified, and pronominals being fully specified for ϕ -features. Without providing the technical details, the core idea of Reuland's account is that an agreement relation in ϕ -features is then establishing between the anaphor and its antecedent. A particular translation procedure mapping the syntactic onto the semantic representation ensures that the ϕ -feature dependency between the two elements results in a binding relation at LF. Specifically, Reuland assumes that an anaphor is translated as a variable.⁵ A pronominal is translated as a variable, or undergoes QR and is treated as a DP. Finally, if XP establishes a CHAIN relation with a DP, their ϕ -features are copies, and $\text{vbl}_{XP} = \text{vbl}_{DP}$.⁶

English poses a challenge for this kind of analysis since English reflexives are morphologically marked for person, number, and gender, i.e. they are not obviously underspecified for ϕ -features.⁷ Reuland focuses mostly on Dutch data, and offers just a brief discussion of English facts. For English, he essentially adopts the analy-

⁵More precisely, a SE-anaphor, such as *zich* in Dutch. See section 8.2.5 for clarification regarding the distinction between SE and SELF-anaphors.

⁶See Reuland (2001) for the characterization of the chain relation involved.

⁷This has been noted already by Burzio (1991), who thus concludes that only a one-way implication can be established: featural defectiveness implies referential defectiveness, but not the other way round.

sis of SELF-anaphors put forth in Reinhart and Reuland (1991, 1993). On this view binding conditions are seen as conditions on linguistic marking of reflexivity. SELF is interpreted as an identity predicate that adjoins to the predicate head, restricting its interpretation. It reflexive-marks a syntactic predicate precisely in those environments where it is a syntactic argument of that predicate and where it is allowed to (covertly) adjoin to the head of the predicate. Further aspects of Reinhart and Reuland's (1991; 1993) proposal and some problems it raises will be discussed in section 8.2.5.5. For now, note only that Reuland (2001) does not extend the ϕ -based analysis to English facts, but rather adopts a significantly different strategy to capture the distribution of English anaphors.

In the following sections, I will follow Reuland's intuition that ϕ -features play a role in determining binding relations, but I will show that English facts can also be analyzed along the same lines. In other words, I will argue that the establishment of a ϕ -feature dependency is crucial in regulating the distribution of anaphoric expressions in both English and Dutch. The two languages will differ not in the strategies they employ for encoding binding relations, but rather in morphological resources they have at their disposal for lexicalizing syntactic structure.

I have here offered just a very brief sketch of Reuland's proposal. However, since I will adopt various aspects of Reuland's account in the course of this chapter, further details of his analysis will be discussed at the points where they become relevant.

8.2 Developing an analysis

In this section, I will develop an Agree-based analysis of binding relations. In other words, I will argue that a feature-dependency is established between an anaphor and its antecedent in the course of the syntactic derivation. This syntactic feature dependency is then translated into a bound variable dependency at the semantic interface. The account will draw on earlier proposals that encode binding relations via Agree, but the precise implementation will differ in a number of important respects. The first difference that will be of significance concerns the assumptions regarding how lexical insertion proceeds. I will therefore start off by outlining my assumptions regarding the nature of lexical insertion.

8.2.1 On lexical insertion

Unlike the Agree-based approaches discussed in the previous section, I assume that insertion of lexical material occurs post-syntactically. Lexical insertion is here seen as a process that provides features which are manipulated by syntax with a phonological expression. Rather than there being one-to-one correspondence between morphemes and syntactic terminals, I assume that a single morpheme can spell out more than one terminal (Starke (2006), Ramchand (2008), etc.). The insertion of lexical

material is regulated by the following interface Spell-Out condition:⁸

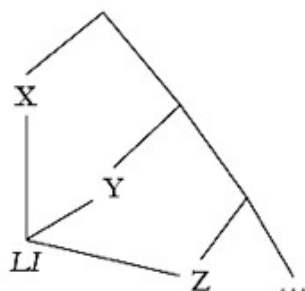
(13) *The Superset Principle*

A phonological exponent of a Vocabulary item can spell out a fragment of a syntactic tree if its lexical entry contains all or a *superset* of features present in the structure that is to be spelled out.

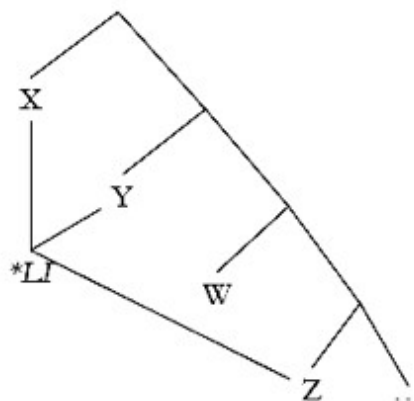
There are also structural conditions on lexical insertion. For the purposes of the current discussion, I will adopt the view that a morpheme can lexicalize a sequence of syntactic heads, where the lexicalized syntactic heads must be contiguous (see Abels and Muriungi (2008), Ramchand (2008), etc). For instance, a lexical item *LI* whose lexical entry contains the features X, Y and Z can spell out these features in (14a), but not in (14b), as in the latter case X, Y and Z are not adjacent.⁹

(14)

(a)



(b)



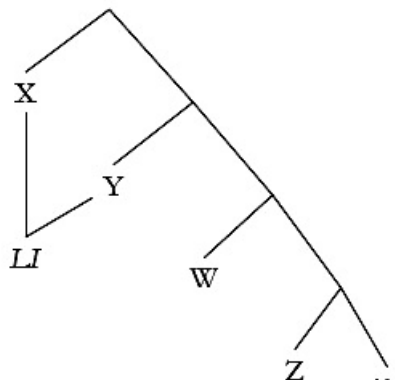
Another licit lexicalization pattern is given in (15). In this case, *LI* spells out only X and Y, W and Z being lexicalized by some other lexical item(s). This is in conformity with the Superset Principle since the spelled-out features are a subset of the features present in the lexical entry of *LI*. On this view, the features of lexical

⁸The Superset Principle was first proposed by Michal Starke in a series of class lectures. For an alternative, widely adopted condition on realization of syntactic structure, namely the Subset Principle of Distributed Morphology, consult for instance Halle and Marantz (1993), Halle (1997). For a detailed discussion of the Superset Principle and some of the benefits of adopting it over the Subset Principle see Caha (2009).

⁹An alternative (see Starke (2006), Caha (2009)) is to require that the fragment of a syntactic tree spelled out by a lexical item be a constituent, which is consistent with (14) only if “...” is empty. Since this view often requires extra movements to create the right context for lexicalization, I resort to the former alternative purely for the sake of convenience, and won’t explore at this point whether there are any empirical benefits in choosing one alternative over the other.

items can be ignored during insertion, but all the features present in the syntactic structure must be lexicalized.

(15)



With this background on lexical insertion in place, let me now return to the discussion regarding the mechanisms involved in encoding binding relations.

8.2.2 On the nature of the feature(s) involved

In the canonical binding theory, binding relations were encoded by means of indexation. Indices are integers that are appended to DPs, though in practice alphabetic subscripts such as *i, j, k* etc. are often used as a notational variant. The Binding Theory controlled how these indices are distributed on DPs and in this way regulated when DPs can corefer and when they cannot. For instance, Condition A stated that anaphors must be *bound*, i.e. coindexed with another c-commanding DP, within a local domain. With the advent of Minimalism, indices have been declared illegitimate syntactic objects, chiefly due to the fact that they violate the Inclusiveness Condition. The Inclusiveness Condition, proposed in Chomsky (1995), requires that any structure formed by computation be constituted of elements already present in the lexical items selected for the numeration, hence “no new objects are added in the course of computation apart from rearrangements of lexical properties (in particular, no indices, bars, traces, lambdas)” (Chomsky (1995), p. 228).¹⁰ Note that indices cannot be present in the featural specification of lexical items since an index borne by an anaphor or a bound pronominal must covary with that of an antecedent, while other noun phrases must be distinguishable on the basis of the integer value they bear. In light of these considerations, binding conditions, if they are to be maintained, have to be reformulated, given that they make reference to the concepts ‘bound’ and ‘free’, whose definitions involved indexation. In Chomsky (1993), binding conditions are restated in the following manner (with D an undefined local domain):

¹⁰Chomsky reiterates this position in his later work. See Chomsky (2000, p. 114) and Chomsky (2001, p. 3).

- (16) *Condition A*
If α is an anaphor, interpret it as coreferential with a c-commanding phrase in D.
- (17) *Condition B*
If α is a pronominal, interpret it as disjoint from every c-commanding phrase in D.

However, as Hicks (2009) notes, it is questionable whether binding conditions conceived of in this way indeed need no recourse to indexation. Unless ‘coreference’ and ‘disjoint reference’ are taken as grammatical primitives, they will need definitions too, which presumably brings us back to indices. Note also that the definition of binding based on indexation allowed two types of relationships to be dealt with in the same manner: binding of anaphors and pronominals by quantified DPs, and binding by referential DPs. The restated binding conditions capture only the latter relationship, leaving binding relations involving quantified DPs without an explanation.

Setting the concerns that the use of indices raises aside for a moment, I would like to follow the intuition that indexation plays a role in encoding binding relations, but implement it differently than was done within the canonical binding theory. In particular, I assume that every noun phrase contains what I label here the In(dex)-feature. At the semantic interface, this feature is interpreted as an instruction to assign an index, i.e. an integer, to the phrase bearing it. Syntax manipulates the [In]-feature, but no indices (in the sense commonly conceived) are present in the course of the syntactic derivation. Integers are only inserted at the point when the syntactic structure is mapped onto a semantic representation.¹¹

This view of indexation avoids, I believe, the problem regarding the violation of the Inclusiveness Condition. Observe that the Inclusiveness Condition prohibits the use of indices only insofar as these are inserted in the course of the syntactic derivation. As far as the Inclusiveness Condition is concerned, there is in principle nothing wrong with employing indices to establish anaphoric dependencies, as long as these are not introduced in the narrow syntactic component. Reuland (2001) also notes that certain linguistic objects, such as lambdas, must be allowed to enter the structure at some level of representation without incurring a violation. He thus concludes that the mapping to the semantic interface does not obey the Inclusiveness Condition.¹² The analysis proposed here leads to the same conclusion.

Let me now sketch how the presence of the [In]-feature can be used to regulate the distribution of anaphoric expressions. Note that SELF anaphors are morphologically

¹¹I will focus here on indexation as a means of encoding anaphoric relations. How other uses of indices, for instance as markers of movement dependencies, should be handled is beyond the scope of the current chapter.

¹²Cf. also Adger and Ramchand’s (2005) proposal, where operator-variable relations in A’-dependencies are encoded via syntactic features Λ and ID, which map at the interface onto a lambda-operator and a variable, respectively.

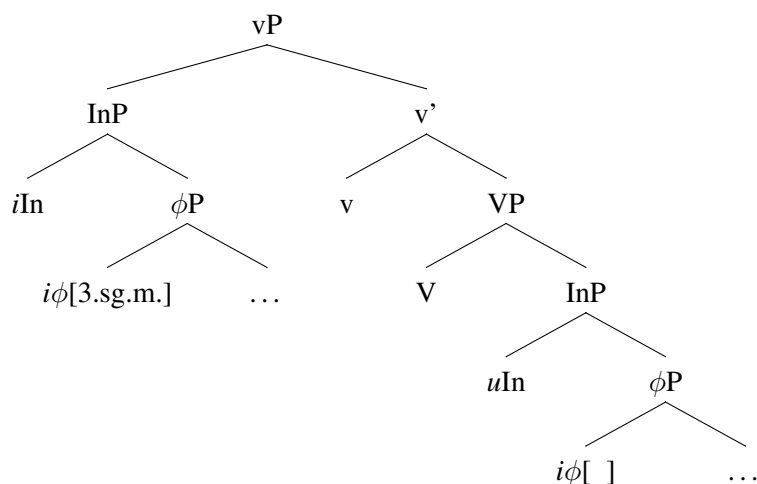
complex, and can be decomposed into a pronominal part and a SELF-morpheme. I assume that the morpheme *self* in English bears the uninterpretable Index-feature, [uIn], in its featural specification. In contrast to Chomsky (1995), I pursue here the view that uninterpretable features need not be checked off and deleted before they are sent off to the interface; they are simply not visible at the interface.¹³ A partial derivation of a simple transitive sentence such as (18) is given in (19).¹⁴

(18) John_{*i*} likes himself_{*i*}.

In (18), the internal argument will start out with [uIn] and interpretable, though unvalued ϕ -features, as illustrated in (19a). Note that nothing is said at this stage about the phonological realization of the syntactic structure. The ϕ -features of the internal argument will be valued through an Agree-relation with the external argument. (19b) gives the structure after the application of the Agree operation. Observe that the relevant syntactic dependency between the internal and the external argument is established within a phase. The first phasal head to enter the structure, on our assumptions, will be AspP, dominating the vP. Locality restrictions on the Agree-operation are therefore respected.

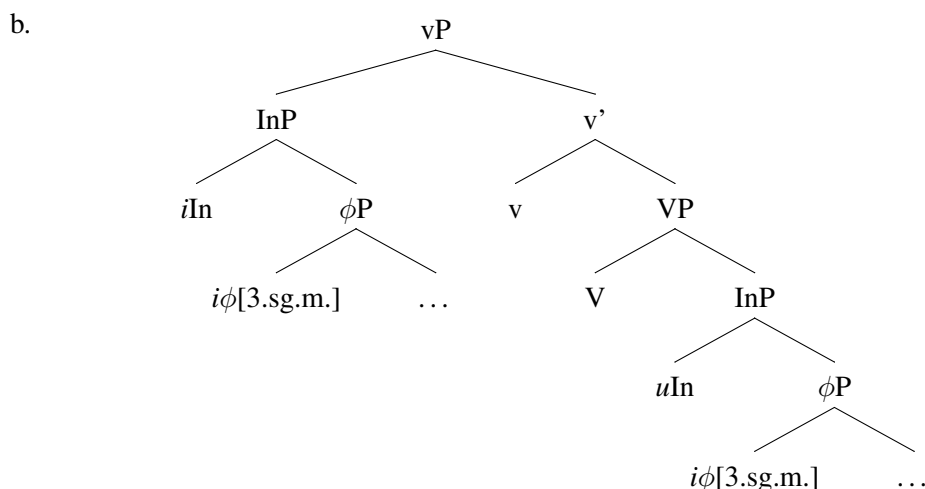
(19)

a.



¹³See the following section for a brief discussion regarding the nature of the feature system assumed here.

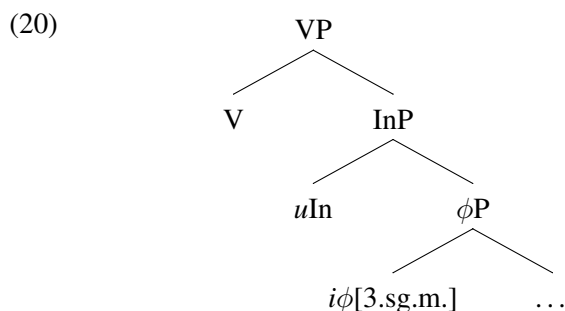
¹⁴For the sake of clarity, other features present in the nominal functional hierarchy that are not relevant for the discussion at hand are ignored.



At Spell-Out, the [*uIn*] will be lexicalized by the morpheme *self*, which contains this feature in its feature specification. The pronominal part of the complex anaphor will spell out the ϕ -features.

At the same time, the structure in (19b) will be sent off to the interfaces for interpretation. I have already suggested that the syntactic [*In*]-feature is interpreted at the semantic interface as an instruction to assign an index. Let us further assume that every noun phrase must bear an index in the semantic representation in order to be interpreted. The presence of the [*In*]-feature in the syntactic structure typically ensures that this requirement is met. However, in (19b), the internal argument bears [*uIn*], which is not visible at the interface. Therefore an index cannot be assigned by simply mapping the [*In*]-feature onto an integer value. I would like to suggest that in such cases, the noun phrase lacking a visible [*In*]-feature receives the same index as the noun phrase it enters into an Agree-relation with; i.e. the translation procedure mapping the syntactic structure onto a semantic representation copies the index of the NP which is in a ϕ -feature dependency with the phrase that has no [*iIn*]. The current proposal thus follows the intuition of some earlier approaches that an Agree-relation in ϕ -features plays an instrumental role in regulating the distribution of anaphoric expressions (e.g. Reuland (2001), Rooryck and Vanden Wyngaerd (2010)).

Note that the establishment of an Agree-dependency in configurations involving anaphors is crucial. A derivation that starts with the combination of [*uIn*] and valued ϕ , as illustrated below, would not yield an interpretable interface representation since the phrase bearing [*uIn*] would not be able to receive an index. On the one hand, direct mapping of [*uIn*] to an integer value is not possible since [*uIn*] is not visible at the interface; on the other hand, an index cannot be assigned via the proposed translation procedure since no dependency would be established with the antecedent DP.



Hence, on the approach pursued here, locality restrictions on the distribution of anaphors reduce to the requirement that a syntactic feature dependency be established between the anaphor and another c-commanding DP. In the previous chapter, I have argued that the phase is the local domain for binding, i.e. that an anaphor cannot be separated from its antecedent by a phase boundary. The account just sketched explains why this should be the case. As we have seen, an anaphor must enter into an Agree-relation with another DP capable of valuing its ϕ -features. No valuation of features can take place across a phasal boundary, given that the Agree-operation is sensitive to the PIC. Consequently, a noun phrase can function as a potential antecedent for an anaphor only if it occupies the same phasal domain.

It remains to be explained why anaphors are necessarily bound by the same DP that values their ϕ -features. In other words, why isn't it possible for an anaphor to get its ϕ -features valued in syntax via an Agree-relation with a local DP, and then be bound at the interface by a different, local or nonlocal DP? Recall that the translation procedure we have adopted mapping syntactic to semantic representations ensures that the anaphor always bears the same index as the DP it agrees with in ϕ -features. In other words, an anaphor always ends up coindexed with the DP it is syntactically dependent on. In the semantic literature on binding, it is commonly assumed that coindexation obligatorily results in semantic binding, or to put it differently, that the only interpretation of coindexation is the bound variable one (see for instance Reinhart (1983b), Reinhart and Grodzinsky (1993), Heim and Kratzer (1998)). We might implement this by following Heim and Kratzer (1998) in assuming that in cases of coindexation QR must apply. The rule of Predicate Abstraction, given in (21), then applies, yielding the desired bound variable interpretation.¹⁵

(21) *Predicate Abstraction Rule* (Heim and Kratzer, 1998, p. 186)

Let α be a branching node with daughters β and γ , where β dominates only a numerical index i . Then, for any variable assignment a , $\llbracket \alpha \rrbracket^a = \lambda x \in D. \llbracket \gamma \rrbracket^{a^{x/i}}$.

¹⁵See Heim and Kratzer (1998) for further details and Reinhart (1983b) and Büring (2005a) for alternative ways of deriving bound-variable interpretations from configurations involving coindexation. Since choosing between these alternatives is not crucial for the analysis sketched, I won't discuss the issue further.

It is well-known that anaphors obligatorily function as bound variables. Recall that one piece of evidence supporting this conclusion is the fact that anaphors contained within elided VPs do not display the same strict/sloppy ambiguity as pronominals under VP-ellipsis. Only the sloppy interpretation, which requires variable binding, is possible, while the strict reading which involves coreference is unavailable. A relevant example is repeated below.

- (22) John hit himself, and Bill did too.
 (=Bill hit Bill, *but* ≠ Bill hit John)

Under our current assumptions, the only way an anaphor can receive an index is by ‘inheriting’ it from an antecedent, thereby ending up coindexed with a c-commanding DP. Given that coindexation necessarily results in semantic binding, the analysis pursued here correctly predicts that configurations featuring anaphors should allow only the bound variable interpretation.

Pronominals on the other hand can, but need not, function as bound variables, as evidenced by the fact that both the sloppy and the strict reading under VP ellipsis are possible.

- (23) Felix thinks that he is smart, and Max does too.
 (=Max thinks that Max is smart, *or*
 =Max thinks that Felix is smart)

I assume that pronominals (as well as other DPs) differ from anaphors in being lexically specified with an interpretable [In]-feature. The [*i*In] is then directly mapped onto an integer value at the semantic interface. If a pronominal is assigned the same integer value as another c-commanding DP, a bound variable interpretation will be derived, in accordance with the assumption that coindexation leads to semantic binding. A pronominal can however be assigned an integer value distinct from those of other DPs present in the structure. The assignment function will then map the integer to an individual in D, i.e. the set of actual individuals in the real world (see Heim and Kratzer (1998), Büring (2005a) for details). Note that this does not ensure that a pronominal will end up disjoint in reference from other DPs present. A distinct integer value assigned to another DP may be mapped to the same individual in the real world, yielding a coreferential interpretation.¹⁶

Now that we have clarified the role of indexation in encoding binding relations, pursuing the practice we have adopted from the canonical binding theory of marking shared reference by coindexation becomes rather inconvenient. Note that in the canonical binding theory, employing shared subscripts as in (24a) is used to indicate referential dependency, with no distinction being made between the bound variable and the coreferential interpretation. Contra-indexing, as in (24b), typically signals that the indexed NPs are disjoint in reference. On the view pursued here however, coindexation necessarily implies semantic binding, so employing the notation in

¹⁶The distribution of pronominals will be discussed in more detail in section 8.3.

(24a) would be interpreted as signalling only the bound variable interpretation, which might be a problem if we want the discussion to encompass the coreferential reading as well. In addition, recall that in the system we have just outlined, conindexation can nevertheless yield shared reference.

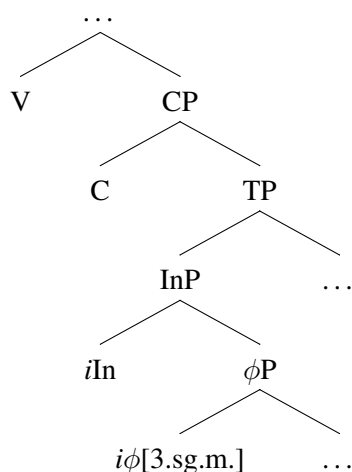
- (24) a. $Felix_i$ thinks that he_i is smart.
 b. $Felix_i$ thinks that he_j is smart.

To sidestep this problem, I will henceforth use italics to signal referential dependency in cases where I wish not to be explicit about whether we are dealing with a bound variable or a coreferential interpretation. Thus, italicizing *he* with *Felix* in (25) should be interpreted as signalling that *he* is either semantically bound or coreferential with *Felix*.

- (25) *Felix* thinks that *he* is smart.

The analysis just sketched assumes that pronominals bear an [*iIn*] in their featural specification, in addition to a valued set of ϕ -features. Since the [*In*]-feature is interpretable, there is no requirement that a syntactic dependency be established between a pronominal and another DP it is either semantically bound or coreferential with. The [*iIn*] can be directly mapped onto an integer value at the interface. In cases such as (25), the subject of the embedded clause would bear [*iIn*] and a valued set of ϕ -features.¹⁷ No feature dependency with the subject of the matrix clause would be established. At Spell-Out, the pronominal lexicalizes both the ϕ -features and the [*iIn*], as made possible by its feature specification.

- (26)



As we have seen, the absence of a syntactic dependency does not preclude the establishment of a bound variable dependency at the interface. If the pronominal's [*iIn*] is mapped onto an integer value shared by another c-commanding DP, the pronominal

¹⁷Other features not relevant for the discussion at hand are ignored.

will end up being semantically bound. Thus, although both anaphors and pronominals may be semantically bound (anaphors obligatorily so), only anaphors enter into a syntactic dependency with their antecedents. This explains why only configurations involving anaphors display syntactic locality effects. Pronominals on the other hand, are insensitive to locality restrictions even when they are interpreted as bound variables. They can appear inside embedded clauses, and even inside syntactic islands, as illustrated in (27) ((27a) is from Hornstein et al. (2007), the subsequent examples are from Johnson (2009)). In (27a), the bound pronominal is embedded in an adjunct island, in (27b) it is inside a complex NP and in (27c) inside a sentential subject. A bound-variable dependency can nevertheless be established. Recall that coreferential interpretation is impossible with quantified-NP antecedents. Thus, using quantified-NPs in (27) ensures that we are dealing here with the bound variable interpretation.

- (27) a. [Every boy]_i was pleased because Mary met him_i.
 b. [Every linguist]_i reviewed the claim that this interested her_i.
 c. [No linguist]_i realizes that to understand her_i isn't easy.

Let me now summarize the proposed analysis of anaphoric relations. The account developed here assumes that every NP must bear an index at the interface in order to be interpreted. The syntactic correlate of an index is the proposed [In]-feature, which is mapped onto an integer value at the interface. If a noun phrase receives the same integer value as another c-commanding NP, a bound variable dependency will necessarily be established, assuming that coindexation always results in semantic binding. If the integer value a particular NP receives is unique, the assignment function will map this value onto an individual in the real world. In such cases, coreference can nevertheless obtain, given that NPs with distinct integer values can be mapped onto the same individual. The presence of the [In]-feature and the mechanism of indexation behind it play a crucial role in encoding anaphoric relations. As we have seen, there are several properties that distinguish anaphors from pronominals (and other NPs). On the semantic side, anaphors are always interpreted as bound variables. On the syntactic side, anaphors require their antecedents to be sufficiently local; on our view, they must occupy the same phasal domain. On the analysis sketched, these properties are a consequence of the particular featural specification of anaphors and the way syntactic structure maps onto a semantic representation. A derivation yielding an anaphor starts with [*u*In] and unvalued ϕ . It is the combination of these features that ultimately obligatorily yields a bound variable dependency at the interface. In order to value anaphor's ϕ -features, an Agree relation must be established with another c-commanding DP. At Spell-Out, the [*u*In]-feature is lexicalized by the morpheme *self* in English, while the pronominal part of the complex anaphor spells out the now valued set of ϕ -features. Given that the morpheme *self* contains [*u*In] in its featural specification, this morpheme will necessarily appear whenever [*u*In] is present in the syntactic structure. [*u*In] is invisible at the interface and therefore an anaphor cannot be assigned an index by simply mapping the [In]-feature onto an integer value. In cases such as these, the translation procedure mapping syntactic onto semantic repre-

sensation ensures that the anaphor receives the same integer value as the noun phrase it agrees with in ϕ -features. The establishment of a syntactic feature dependency is therefore crucial for obtaining a legitimate interface representation. A consequence of this view is that an anaphor always ends up coindexed with its antecedent, and we thus correctly predict that it should always be interpreted as a bound variable. The presence of a syntactic dependency is also responsible for the observed locality restrictions on anaphoric binding. Given that Agree cannot operate across phasal boundaries, the relevant featural dependency must be established within the phase. Consequently, potential antecedents for an anaphor are only those noun phrases that occupy the same phasal domain.

Pronominals differ from anaphors in bearing an [*i*In] in their featural specification. The [*i*In] can be directly mapped onto an integer value and no syntactic dependency is required. As already noted, this mapping procedure may result both in a bound variable and a coreferential interpretation. No syntactic locality effects are expected even on the bound variable readings of pronominals since no syntactic dependency is established between a pronominal and its antecedent.

8.2.3 Some comments on the feature system assumed

Chomsky (2000, 2001) (henceforth *MI/DbP*) makes the following assumptions regarding the mechanism of Agree and the nature of the feature system involved, as summarized in Pesetsky and Torrego (2007).

- (28) *Agree*
 (i) An unvalued feature F (a probe) on a head H scans its c-command domain for another instance of F (a goal) with which to agree.
 (ii) if the goal has a value, its value is assigned as the value of the probe.
- (29) *Valuation/Interpretability Biconditional*
 A feature F is uninterpretable iff F is unvalued (at the point at which it enters the derivation).¹⁸
- (30) *Deletion of uninterpretable features*
 Once an uninterpretable feature is valued, it can and must delete.

The Agree operation provides values to unvalued features under appropriate conditions. While the valuation of an unvalued feature has consequences for PF-representations, any feature valued by Agree in the course of the syntactic derivation remains uninterpretable at LF. It is assumed that uninterpretable features are not tolerated at LF and therefore must be deleted by the Spell-Out operation. The system conjectures that all and only features which enter the derivation unvalued are uninterpretable. The hypothesized link between being unvalued and being uninterpretable is neces-

¹⁸The bracketed part is not stated in Pesetsky and Torrego (2007), but is necessary, I believe, to make the assumption behind Chomsky's system clear.

sary because the Spell-Out operation has no access to the interfaces and therefore does not ‘know’ which features to transfer to LF and which to delete. Once the link between feature values and feature interpretability is established, the decision about which features should be deleted can be made by simply examining whether the features have been valued or not. A problem that arises on this view, as Chomsky (2001) himself notes, is that once an LF-uninterpretable feature is valued in the course of the syntactic derivation it becomes indistinguishable from a feature which entered the derivation valued. In other words, although the relevant distinction exists prior to the application of Agree, it is lost once Agree has applied. Chomsky thus concludes that Spell-Out must apply shortly after the uninterpretable features have been assigned values, so that it ‘remembers’ that the relevant features were previously unvalued and delete them accordingly.

The analysis of binding relations in terms of Agree outlined in the previous section makes several departures from Chomsky’s assumptions regarding Agree and the nature of the feature system. These will be briefly mentioned here, but the reader is referred to the references that will be cited for more extensive discussion of the pertinent assumptions. The goal of this section is simply to point out that such departures from Chomsky’s MI/DbP’s model are neither entirely novel, nor completely unmotivated.

First of all, the proposed analysis rejects the Valuation/Interpretability Biconditional.¹⁹ Note that I have assumed that a derivation yielding an anaphor starts with an unvalued set of ϕ -features, which is nevertheless (once it has been valued) interpretable at the interface. Doubts about the plausibility of the Valuation/Interpretability Biconditional have also been raised by Pesetsky and Torrego (2007). The authors wonder why a link of that kind should exist between two such distinct properties of lexical items as interpretability (“does the item have a message to send to the semantics?”) and valuation (“are any syntactically relevant properties of the lexical item left unspecified?”). They suggest instead that interpretability and valuation are independent of each other. This leads to an enriched repertoire of features, but one which Pesetsky and Torrego argue is empirically supported. Allowing uninterpretable features into the system also makes it possible to simplify some aspects of Chomsky’s system. Recall that on Chomsky’s view there is a problem regarding how Spell-Out can distinguish features which enter the derivation valued (which are by assumption interpretable) from those that receive their value in the course of the syntactic derivation (by assumption uninterpretable). To circumvent this problem, Chomsky suggests that Spell-Out must apply shortly after valuation, and argues that this provides further motivation for the cyclic nature of Spell-Out. However, Epstein and Seely (2002) point out that the logical conclusion of this approach is that at any stage after feature valuation has taken place (whether ‘shortly’ after or not), the two types of valued features are indistinguishable, so the problem remains. Allowing uninterpretable features then provides a potential solution. Regardless of whether an uninterpretable

¹⁹Hicks (2009) also abandons this conjecture of Chomsky’s and assumes instead that *all* features are interpreted, either at LF or at PF.

feature is valued or not, its lexical status as being uninterpretable remains unaltered.

The current proposal also rejects the assumption stated in (30). On the view pursued here, uninterpretable features need not be deleted before reaching the interface. The same position has been taken and argued for recently by Epstein et al. (2010) (see also Carstens (2010)). The necessity to assume (30) in Chomsky's system stems from another stipulation, namely that the LF interface can recognize every feature appearing at LF. Hence the removal of uninterpretable features is required and delegated to the Spell-Out operation. However, as Epstein et al. (2010) point out, it is not necessary to stipulate that the interface is designed to recognize features that it cannot use. Instead, we may assume that the interface recognizes only the interpretable features. Uninterpretable features are simply invisible. The interface then by definition cannot recognize uninterpretable features, hence these need not be removed before reaching the interface. This considerably simplifies the internal mechanisms of the Spell-Out operation, given that inspection of syntactic structure and feature deletion are no longer necessary.

Finally, consider the conditions on the Agree operation stated in (28), in particular the configurational requirement on probe-goal agreement. Chomsky assumes that upon entering the derivation, unvalued features search within their c-command domain for an appropriate goal with a matching set of valued features. Crucially, probing is always downwards, that is the element bearing the unvalued feature(s) always c-commands the goal. However, precisely the opposite configurational relation between the probe and the goal seems to be required on the analysis of binding relations pursued here; the element bearing unvalued features (ϕ -features in our case) must be c-commanded by the goal, rather than the other way round as Chomsky assumes. Note however that this particular condition on Agree operation has not gone unchallenged in the literature (see Baker (2008), Rezac (2003), Hicks (2009), Zeijlstra (2010)). Baker (2008) for instance argues that agreement must be allowed to probe both upwards and downwards, suggesting that the c-command condition on Agree be revised as follows:

- (31) *The c-command condition* (Baker, 2008, p. 45)
 F agrees with XP, XP a maximal projection, only if F c-commands XP or XP c-commands F.

On the other hand, Zeijlstra (2010) has argued that Agree is in fact always upwards. He points to a number of empirical and conceptual problems that Chomsky's conception of Agree faces, claiming that these can be avoided if it is assumed that Agree applies in a configuration where the probe is c-commanded by the goal, rather than the other way round.

As Baker (2008) notes, Chomsky (2000) argues for restricting the search space of the probe to its c-command domain on the grounds that it reduces computational load, and thereby reflects a 'perfection' of human faculty. However, Baker points out that the independent restriction that the search space extends no further than the current phase is adequate enough to narrow the probe's search space and avoid computational

explosion. Thus, there is no conceptual motivation for assuming that Agree can apply only in configurations where the probe *c*-commands the goal. The empirical facts that these authors review furthermore suggest that such an assumption is also theoretically undesirable.

8.2.4 Minimality violations

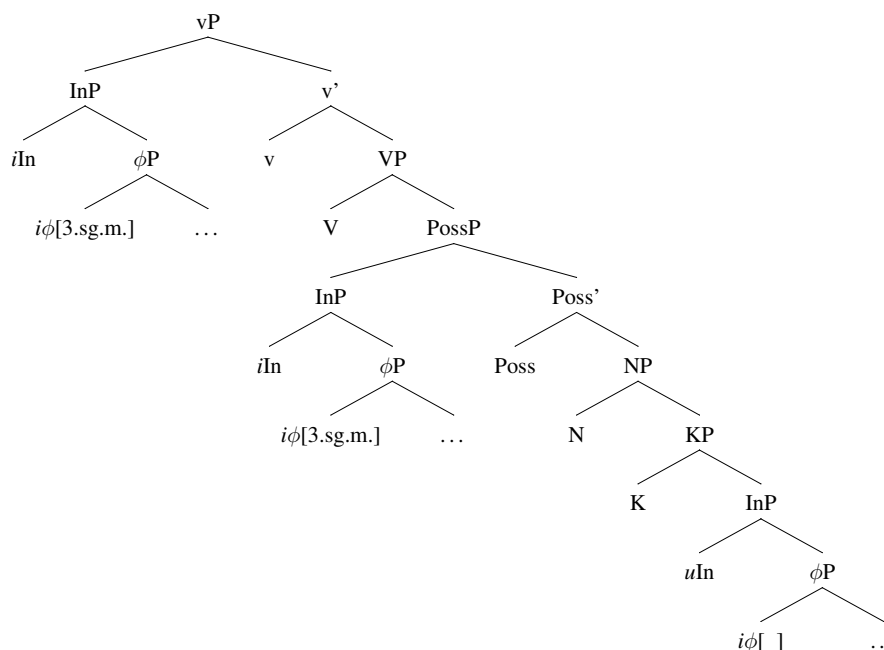
In the previous section I have argued that configurations where anaphors surface involve the establishment of an Agree relation in ϕ -features between the anaphor and its antecedent DP. It is generally assumed that Agree is sensitive to Minimality, i.e. that a probe cannot see past the closest potential goal bearing the relevant feature. If an Agree dependency lies at the heart of anaphoric relations, then we should see the effects of Minimality in binding configurations. In other words, if there are two potential antecedents available, i.e. two phrases bearing a valued set of ϕ -features, within the same phase, Minimality should block the establishment of an Agree-relation with the more distant phrase. Consider in this light examples involving picture-NPs. Recall that, as commonly noted in the literature, the presence of a possessor is crucial in determining the binding domain for an anaphor embedded inside a noun phrase. In examples like (32a), although *John* and the anaphor occupy the same phase at one point in the derivation, *John* cannot bind the anaphor. The only available antecedent is the possessor, as shown in (32b). When the possessor is absent, the clausal subject can bind the anaphor, as illustrated in (32c).²⁰

- (32) a. *John_i likes Bill's pictures of himself_i.
 b. John likes Bill_i's pictures of himself_i.
 c. John_i likes pictures of himself_i.

On the approach pursued here, the fact that the clausal subject cannot bind the anaphor in contexts like (32a) might be attributed to a Minimality violation. Consider a syntactic configuration with the possessor present.

²⁰Recall that within the canonical binding theory, these facts are handled by simply stating that a domain counts as a local domain for anaphors only if it contains a subject.

(33)



In (33), the closest phrase bearing valued ϕ -features is the possessor. The unvalued ϕ -features of the anaphor can thus receive a value through an Agree relation with the possessor, yielding the binding pattern observed in (32b). Given that the possessor is the closest goal bearing the relevant feature, Minimality prohibits valuation of the anaphor's ϕ -features via an Agree relation with the external argument. This accounts for the ungrammaticality of examples such as (32a).

In configurations where a possessor is present, a bound variable dependency can be established between the complement of the picture-NP and the external argument if the pronominal is used instead of the anaphor, as shown in (34a), or perhaps more clearly in (34b).

- (34) a. John_i likes Bill's pictures of him_i.
 b. [Every student]_i likes Bill's picture of him_i.

In such cases though, no syntactic dependency is established, and Minimality is therefore irrelevant. Recall that on our assumptions, derivations that yield a pronominal start with an [*iI*] and a valued set of ϕ -features. At the interface, the bound variable relation between the pronominal and the external argument can be established. As already pointed out, the interpretative relation established at the interface is not sensitive to syntactic locality restrictions.

When the possessor is absent, as in (32c), no intervention effects arise and the relevant Agree relation with the external argument can be established. Consequently, it is correctly predicted that in cases such as these the derivation yielding an anaphor

converges.

A potential problem now arises with binding relations established in double object constructions. Consider the following pair of examples.

- (35) a. John showed Bill_i himself_i in the mirror.
 b. John_j showed Bill himself_j in the mirror.

In (35), the anaphor surfaces regardless of whether *John* or *Bill* is chosen as the antecedent. As things stand now, the grammaticality of (35a) is predicted. The unvalued set of ϕ -features of the direct object can be valued via an Agree relation with the indirect object, which is the closest potential antecedent. That the anaphor appears in cases like (35a) is therefore unsurprising. That it should appear in (35b) is however unexpected. In this case, the external argument should not be able to value the ϕ -features borne by the direct object across the intervening indirect object. We rather expect that the bound variable interpretation should be derived from a configuration that would see the pronominal surfacing, i.e. the one starting with valued ϕ -features, as in the superficially similar (34a). Yet, the anaphor nevertheless surfaces.

Suppose however that there is an additional feature involved in configurations such as (35) (but crucially not in those like (34)), which only the *self*-morpheme can lexicalize. The appearance of the *self*-morpheme in (35b) would then be unsurprising, even if no Agree relation in ϕ -features is established. In the following section, I will argue that this is precisely the case, i.e. that in configurations where there is a referential dependency between two arguments of the same predicate, as is the case in (35), an additional feature is involved. This feature can be spelled out by the SELF-morpheme, but crucially not by the pronominal. Consequently, the *self*-morpheme will be obligatorily present in cases like (35b).

In order to find some evidence for the presence of this additional feature we will need to look beyond empirical facts from English and focus on a language which, unlike English, distinguishes two kinds of anaphors. The following section therefore discusses the distribution of anaphoric expressions in a language of this type, namely Dutch.

8.2.5 Encoding a three-way contrast in the anaphoric system: A case study of Dutch

8.2.5.1 Featural composition of Dutch anaphors

It is well-known that there are languages which, unlike English, have two types of morphologically distinct anaphors. For instance, Dutch distinguishes *zich* from *zichzelf*, and Norwegian has both *seg* and *seg selv*. Following Reinhart and Reuland (1993), I will use the terms SE and SELF anaphors to refer to these two types

of anaphoric expressions.²¹ In languages such as these, there are therefore three expressions which can serve as bound variables at LF: SELF anaphors, SE anaphors and pronominals. How could this three-way contrast in the anaphoric system be captured?

Reinhart and Reuland (1993) argue that the contrast between SE and SELF anaphors has to do with reflexivity. A predicate is defined as reflexive if (at least) two of its arguments are coindexed.²² Crosslinguistically, reflexive constructions have a special status, in that they must be linguistically marked as reflexive.²³ As Reuland (2001) notes, from the perspective of logical syntax, it is difficult to see why the reflexive counterpart of a structure we may represent as $DP_1 V DP_2$ or $DP_1 (\lambda x (x V DP_2))$ could not be obtained by simply replacing DP_2 by a pronominal since we know that pronominals can function as bound variables. What is more, in many languages it does not suffice to exchange the pronominal for just any anaphoric expression. In Dutch, for instance, the SE anaphor *zich* differs from the pronominal *hem* in that it must be bound. Thus, it appears to be a true anaphor. Nevertheless, (36a) is still ill-formed and cannot mean that everyone hated himself. To express this meaning the SELF anaphor *zichzelf* must be used instead of *zich*. In other words, a reflexive configuration must be specially marked by means of a SELF anaphor in this case. This fact follows neither from the canonical Condition B, nor from the canonical Condition A.

- (36) a. *Iedereen_i haatte zich_i.
 everyone hated SE
 b. Iedereen_i haatte zichzelf_i.
 everyone hated SELF

Reuland (2001) further notes that the special status of reflexive predicates also shows up in a pervasive contrast between transitive and ECM constructions.²⁴ Whereas the canonical binding theory treats binding of an object of a transitive verb and binding of an ECM subject in the same way, in many respects these two cases are not equivalent.

²¹It has become common practice to refer to SE and SELF anaphors as *long-distance* and *local* anaphors, respectively. I find this terminology misleading because it presupposes that SE anaphors are not bound within the local domain, a contention I believe is highly debatable. In fact, in this section I will argue that both SE and SELF anaphors are locally bound.

²²The analysis proposed in the previous section restricts coindexation to cases involving bound variable interpretation, which may give rise to some terminological confusion, as already pointed out. Note however that this understanding of coindexation does not conflict with Reinhart and Reuland's since they also assume that the binding theory only regulates the bound variable interpretation. Thus, using coindexation to mark anaphoric dependencies in this more limited sense is consistent with their view.

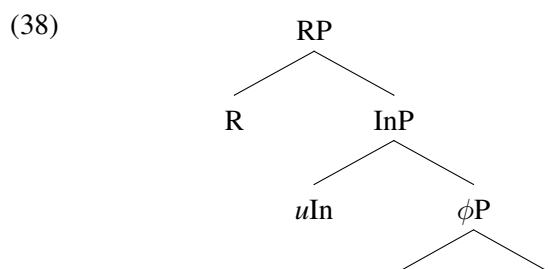
²³Languages employ various means of signalling reflexive configurations, the use of SELF-type anaphors being just one of these. Other strategies include reflexive clitics, verbal affixes, duplication of the bound element etc. (see Schladt (2000) for an overview). Each of these strategies merits extensive study by itself. I will limit my attention here just to SELF anaphors.

²⁴For another discrepancy see Reuland (2001) and Reinhart and Reuland (1993).

For instance, in Dutch (and Frisian), a SELF anaphor is required in canonical transitive constructions, as shown in (36), but not in cases of ECM, an example of which is provided in (37). Note that the configuration in (37) does not qualify as a reflexive configuration since the coindexed DPs are not arguments of the same predicate.

- (37) Henk_i hoorde zich_i/zichzelf_i zingen.
 Henk heard SE/SELF sing
 ‘Henk heard himself sing.’

In contexts of reflexivization, i.e. configurations where a binding dependency is established between coarguments of a predicate, Dutch requires that a SELF anaphor be used. How can this requirement be captured within the model of binding outlined so far? In answering this question, I will follow the strategy adopted in the previous section of linking the distribution of anaphoric expressions to their featural content. I will therefore assume that in constructions involving reflexivization, there is an additional feature present. Let us label the pertinent feature [R], which we may view for now as a shorthand for ‘Reflexive’. Based on the data from Dutch, I conclude that [R] is part of the nominal functional sequence for reasons that will become clear shortly. A possible arrangement of the relevant features in the nominal domain is given below.



Postulating the presence of an additional feature in reflexive configurations is in line with the observation that of the two anaphoric expressions available, the morphologically more complex one is used in these contexts. Observe that the SELF-anaphor, *zichzelf*, can be decomposed into a SE anaphor, *zich*, and a SELF morpheme, *zelf*. Since the SELF morpheme must surface in reflexive configurations, I will assume that *zelf* has the [R]-feature in its feature specification. The morpheme *zich* on the other hand, contains at least ϕ -features and the feature [*uIn*], but crucially no [R]-feature. The presence of [*uIn*] ensures that *zich* behaves as a true anaphor in Dutch. The distributional differences between SELF and SE anaphors can then be attributed to their featural composition.²⁵

²⁵Note that on this view there is no lexical item *zichzelf* in Dutch. In other words, Dutch has lexical items *zich* and *zelf*, rather than *zich* and *zichzelf*. For ease of exposition however, I will continue at times to use the term SELF anaphor as a descriptive label for the combination of the morphemes *zich* and *zelf*.

Before spelling out in more detail how the distributional facts can be captured, I will first briefly discuss the possible semantics of [R] feature. I will then examine a range of syntactic environments, showing how the observed distributional patterns can be accounted for. In 8.2.5.4 I reconsider the featural specification of English reflexives in light of the conclusions reached in this section. The section closes with a discussion of Reinhart and Reuland's (1993) account, highlighting some of the differences between that proposal and the one pursued here.

8.2.5.2 Semantics of the [R]-feature

Reinhart and Reuland (1993) do not discuss the semantic interpretation of SELF anaphors, noting only that the reflexivizing function is carried by SELF, whose semantics they take to be a restriction imposing identity on two arguments of a predicate. Note that on their view, SELF is a noun and it combines with a pronoun determiner or a SE determiner, as in (39).

(39) [NP Pron/SE [N' self]]

In an earlier paper however (Reinhart and Reuland (1991)), they are somewhat more explicit regarding the semantic analysis of SELF anaphors. They propose that SELF be analyzed as a two-place predicate denoting the relation of identity, as in (40a). When a SELF-NP occurs on a grid position, it is interpreted as a restriction on that grid, i.e. it restricts the range of the predicate interpretation. Thus, a VP like (40b) is interpreted as in (40c), where SELF specifies that the two arguments of the V grid must be identical variables. The SELF-N adjoins to V by head-to-head movement. This, they claim, yields an LF of the form (40c), without elaborating any further on the mechanisms of meaning composition involved.

(40) a. SELF $\langle x,y \rangle$
 b. ...[V $\langle y,x \rangle$... [... SELF $\langle x,y \rangle$]]
 c. SELF $\langle x,y \rangle$ & V $\langle y,x \rangle$...

I will follow a different line of thinking and assume that a SELF anaphor has the semantics of a generalized quantifier.²⁶ In particular, I assume that the postulated [R]-feature, which is part of the feature specification of the morpheme *zelf*, denotes a function of the type $\langle e, \langle \langle e, t \rangle, t \rangle \rangle$, i.e. it is a function from entities to generalized quantifiers. The semantic composition might then proceed as follows. The [R]-feature borne by the *zelf*-morpheme combines with the pronominal part of the complex anaphor to yield an NP of the type $\langle \langle e, t \rangle, t \rangle$. Such an NP in the object position of a transitive verb cannot combine with the verb directly due to a type mismatch (as the verb is of the type $\langle e, \langle e, t \rangle \rangle$). It must therefore undergo QR. The verb can then compose by function application with the trace and the subject, yielding a type t meaning. Predicate Abstraction applies to this node

²⁶Origins of this idea can be traced back to Cresswell (1973). See also ter Meulen (2000).

producing meaning of type $\langle e, t \rangle$, which is then a suitable argument for combining with the QR-ed object NP.

What kind of function does this feature denote? As we have seen, examples such as (41) are ungrammatical despite the fact that *zich* is an anaphoric expression which requires a syntactic antecedent.

- (41) *Oscar_i haat zich_i.
Oscar hates SELF

In other words, *zich* cannot be coindexed with another argument of the same predicate. Adding the morpheme *zelf* yields a well-formed output. The question is why. Reuland (2001) offers the following reasoning. *Zich*, being an anaphor, is translated as a bound variable at the interface. The semantic representation of (41) is thus as given below.

- (42) Oscar $\lambda x (x \text{ haat } x)$

The representation in (42) contains two tokens of the variable x . Ungrammaticality of (41) might thus be interpreted as showing that representations of the form $P(x,x)$ are illicit. According to Reuland, the reason for this is that two tokens of the same variable cannot be distinguished at the interface. Therefore, configurations such as (42) effectively contain only one semantic argument. This leads to a valency violation since *haaten* is a two-place predicate, and thus needs two semantic arguments.²⁷

The situation can be remedied by using the morpheme *zelf*. Reuland argues that adding *zelf* makes it possible to keep the two arguments formally distinct. The interface structure that arises from SELF marking is given in (43).

- (43) NP $\lambda x (x \text{ hates } f(x))$

The verb takes two arguments, one identical to x and the other to the value of some function of x . The interpretation of the second argument, $f(x)$, must therefore approximate that of the first argument, but be formally distinguishable from it. In other words, $\|f(x)\|$ must be able to stand proxy for $\|x\|$ wherever necessary.

I will follow this view and assume that *zelf*, or more precisely the [R]-feature which *zelf* lexicalizes, denotes a function that maps x onto an object that can stand proxy for x . Support for this kind of analysis can be found in contexts that allow both *zich* and *zichzelf*, such as ECM constructions illustrated already in (37). In such cases, a contrast in interpretation may arise, showing that the value of *zichzelf* can indeed be distinguished from the value of its antecedent, whereas the value of *zich* cannot. Reuland provides the following pair of examples, inspired by Jackendoff's (1992) Mme Tussaud's contexts, in which a famous person visits the wax museum.

- (44) a. Ze_i zag zich_i in een griezelige hoek staan.
she saw SE in a creepy corner stand

²⁷Consult Reuland (2001) for a more detailed discussion.

- b. Ze_i zag $zichzelf_i$ in een griezelige hoek staan.
 she saw SELF in a creepy corner stand

In (44a), the preferred reading is the one where the subject sees her own reflection in the mirror, i.e. the person seeing and the one being seen are indistinguishable. The interpretation where the subject sees a representation of herself, e.g. a statue, can only be obtained by using *zichzelf*. In other words, the person seeing and the one being seen are presented as distinguishable entities, which is achieved by interpreting *zichzelf* as a *representation* of the subject. This interpretation is often referred to as a *dissociation* or *proxy* reading.

A contrast along the same lines can be observed in the following examples (attributed to Voskuil and Wehrmann (1990)).

- (45) a. Münchhausen_{*i*} trok $zich_i$ uit het moeras.
 Münchhausen pulled SE out the swamp
 b. Münchhausen_{*i*} trok $zichzelf_i$ uit het moeras.
 Münchhausen pulled SELF out the swamp
 ‘Münchhausen pulled himself out of the swamp.’

Note first that assuming a small clause analysis of (45), (45) is predicted to pattern with cases like (37), rather than with examples such as (41). *Münchhausen* and *zich* are not arguments of the same predicate, and both *zich* and *zichzelf* should in principle be possible. We will return to the distributional facts in more detail below. Given the availability of both *zich* and *zichzelf*, interpretational differences arise. The German baron Münchhausen, famous for telling far-fetched stories about his adventures, claimed to have escaped from the swamp by pulling himself up by his hair. This improbable interpretation, but the one relevant for the story, can only be obtained by using *zichzelf*. With *zich*, the sentence receives the more natural interpretation where Münchhausen pulled himself out by holding on to something, such as a branch or a rope. In Reuland’s terms, the subject and object are indistinguishable in (45a), but presented as distinguishable in (45b). Thus, in both cases the sentence with *zichzelf* expresses a relation between *x* and *f(x)* that bears a systematic resemblance to *x*, but can be distinguished from it. More examples of proxy readings will be provided in the course of this section.²⁸

To summarize, I have argued that the proposed [R]-feature denotes a function of the type $\langle e, \langle \langle e, t \rangle, t \rangle \rangle$, which maps *x* onto an object that can be viewed as a representation of *x*. The [*uIn*]-feature, which is part of the featural specification of the *zich* morpheme, ensures that there is an anaphoric dependency between the anaphor and its antecedent, as discussed in the previous section. When the antecedent is an argument of the same predicate as the anaphor (i.e. in reflexive configurations), the output is ungrammatical because it yields an illicit interface representation in which two arguments of a predicate become indistinguishable. The use of the *zelf*-

²⁸See also Rooryck and Vanden Wyngaerd (2010) for a discussion of dissociation contexts, and a similar view regarding the semantics of the *zelf*-form.

morpheme, or more precisely the [R]-feature, provides a way of circumventing this problem and obtaining a grammatical output. Hence, the [R]-feature is expected to occur in configurations where two arguments of the same predicate are involved in a binding dependency.²⁹

8.2.5.3 Explaining the distribution of SE and SELF anaphors

Let me now discuss how the presence of the [R]-feature accounts for the distribution of SELF and SE anaphors. Recall that I have adopted the view that lexical insertion takes place post-syntactically and is regulated by the *Superset Principle*, repeated below for the sake of convenience.³⁰

(46) *The Superset Principle*

A phonological exponent of a Vocabulary item can spell out a fragment of a syntactic tree if its lexical entry contains all or a *superset* of features present in the structure that is to be spelled out.

I have also already made clear my assumptions regarding the featural specification of the morphemes *zich* and *zelf* in Dutch. Recall that the lexical item *zelf* has the [R]-feature in its feature specification, while the lexical item *zich* contains at least ϕ -features and [uIn].³¹ I have also argued that two local binding environments must be distinguished, namely reflexive and non-reflexive configurations. I will discuss each of these in turn.

Consider first reflexive configurations, i.e. contexts where two arguments of the same predicate are coindexed. In these cases, the [R]-feature must be present for reasons that were discussed in 8.2.5.2. We therefore predict that *zelf* should obligatorily surface in reflexive contexts since it is needed to lexicalize the [R]-feature. Lexicalization of the [R]-feature can't be accomplished by using *zich*, because the [R]-feature is not a part of the featural specification of this lexical item.

Simple transitive sentences are the canonical cases of reflexive configurations. As illustrated in (47), if the object of a transitive verb is bound by the subject, it must take the form of a SELF anaphor. This is precisely what we expect. Given that the subject binds another argument of the same predicate, an item is required which can

²⁹Given that [R]-feature can also appear in non-reflexive configurations, as I will argue, interpreting the label as standing in for 'reflexive' could cause some confusion. We might therefore think of [R] as a shorthand for 'representation', in light of the proposed semantics of this feature.

³⁰As far as I can tell, the analysis proposed here is rephrasable in terms of the Subset Principle of Distributed Morphology.

³¹I wish to point out here that lexical items *zich* and *zelf* might contain more features than noted here. I will restrict my attention only to features that are relevant for the discussion at hand, namely the [R]- and the [In]-feature and ϕ -features, and leave it as an open question what other features might be present in the nominal sequence and how they should be distributed between these two forms.

lexicalize the [R]-feature. Consequently, the morpheme *zelf* obligatorily surfaces in such configurations.

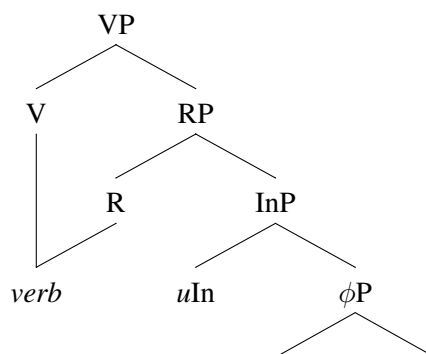
- (47) Max_i haat zichzelf_i / $*\text{zich}_i$. (Reinhart and Reuland (1993))
 Max hates SELF / SE
 ‘Max hates himself.’

However, the distributional patterns observed in (48) are unexpected given our assumptions so far. Structurally, these examples appear identical to those in (47), with two arguments of the same predicate in a binding dependency. Yet, using *zich* is not only allowed in these cases, but in fact required. (from Reinhart and Reuland (1993), and Reuland (2001), respectively):

- (48) a. Max_i schaamt $*\text{zichzelf}_i$ / zich_i .
 Max shames SELF / SE
 ‘Max is ashamed.’
 b. Oscar_i gedraagt $*\text{zichzelf}_i$ / zich_i .
 Oscar behaves SELF / SE
 ‘Oscar behaves (himself).’

As often noted, the distribution of anaphoric expressions observed in (48) hinges on the choice of the verb, and is only possible with *inherently reflexive* verbs. These verbs can never take an object distinct in reference from the subject. How can we account for the patterns in (48)? I will assume that inherently reflexive verbs have the [R]-feature in their feature specification. Consequently, in these cases, the verb can lexicalize the [R]-feature. This is represented schematically in (49).

(49)



Thus, there is no need for a *zelf* morpheme. The remaining features match the lexical entry of the morpheme *zich*, which therefore appears in these contexts.³² What

³²On this view, the verb ends up lexicalizing a part of the nominal structure. For other proposals where a lexical item of a particular category can spell out a part of the extended projection of another category see for instance Son and Svenonius (2008) (a verb lexicalizes

the examples in (48) and (47) have in common is that in both cases the [R]-feature is present. What makes them different is that only in the former case can the relevant feature be spelled out by the verb.³³

The proposal just sketched makes strong predictions regarding cases such as (50). The anaphors in these contexts appear in the complement position of a PP selected by the verb. The prepositions heading the PPs in question belong to the class of functional prepositions and it is plausible to assume that they are not predicative. If so, then the anaphor is in fact the argument of the verb and we again have a reflexive configuration. In light of this, the distributional patterns given below are expected: *zelf* must appear as it is required to spell out the [R]-feature.

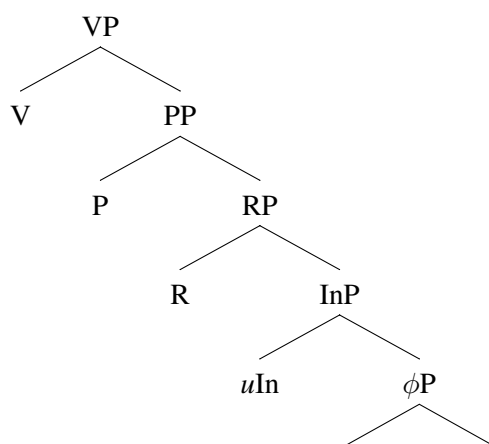
- (50) a. John_i schoot op zichzelf_i / *zich_i. (Koster (1997))
 John shot at SELF / SE
 b. Jan_i mopperde op zichzelf_i / *zich_i. (Everaert (1991a))
 Jan grumbled at SELF / SE
 c. Max_i praat met zichzelf_i / *zich_i. (Reinhart and Reuland (1993))
 Max speaks with SELF / SE

Recall now that the (sole) internal argument of the verb can surface as a SE anaphor provided the verb belongs to the class of inherently reflexive verbs. The analysis of inherently reflexive verbs that I have sketched above predicts that *zich* should never appear in cases involving functional prepositions. That is, even if the verb is of the right type and contains the [R]-feature in its feature specification, we should still see the *zelf* morpheme surfacing. To see why this is so, consider the structure in (51).

a piece of prepositional structure), Dékány (2009) (a preposition lexicalizes a projection in the nominal functional structure), Caha (2009) (a verb lexicalizes nominal functional projections).

³³Reinhart and Reuland (1993) assume that the analysis of configurations headed by inherently reflexive predicates involves an operation in the lexicon which reduces the valency of the predicate. This suggests another way of looking at examples like (48). We could argue that inherently reflexive verbs are listed as one-place predicates, without necessarily assuming that there is an intransitivization operation in the lexicon. If so, then the anaphor would not be the argument of the predicate. Consequently, we would not have a reflexive configuration and the appearance of the SE anaphor would be expected. On this view then constructions with inherently reflexive verbs would be intransitive, while the analysis sketched in the main text predicts that they should be syntactically transitive. Though both views have been argued for in the literature, I find Lekakou's (2005) arguments in favour of a transitive analysis reasonably convincing (for the opposite view see for instance Everaert (1986)). Secondly, on the intransitive analysis it remains to be explained what the function of *zich* is and why it appears at all, given that it is not an argument of the verb. For now I see no compelling reasons to adopt the intransitive analysis.

(51)



Recall that if a lexical item is to spell out a sequence of functional heads, the relevant functional heads have to be contiguous. Therefore, even if a verb's lexical entry contained the [R]-feature, such a verb could not spell out both V and R in the structure above since there is a functional head intervening, namely P.³⁴ Thus, in structures featuring functional prepositions, *zelf* would be called upon to spell out the [R]-feature, regardless of the verb's featural specification.

I know by now of no clear counterexamples to this prediction. At first blush, cases like (52) might seem problematic. As (52) shows, the complement of the preposition in this configuration cannot be distinct in reference from the subject, as in contexts with inherently reflexive verbs. Still, *zich* must be used rather than *zichzelf*, although there is a preposition present (example from Eric Reuland p.c. via Peter Svenonius).

- (52) Jan_i heeft het boek voor zich_i/*Piet.
 Jan has the book in.front.of SE/Piet

Nevertheless, I would argue that cases like (52) do not feature functional prepositions and are thus not expected to pattern with examples in (50) regarding the distribution of *zich* and *zichzelf*. Rather we are dealing here with a spatial PP. As we will see shortly, in contexts featuring spatial PPs, *zich* is the default choice. Some dialects also allow bound pronominals to appear in spatial PPs, and this is true of the configuration in (52) as well. Thus, (53) is also licit (Eric Reuland p.c. via Peter Svenonius).

- (53) Jan_i heeft het boek voor hem_i.
 Jan has the book in.front.of him

Note further that the relation between the verb and the preposition in (52) is not as tight as in cases involving functional prepositions. The preposition in (52) can be replaced by a number of other locative Ps, such as *naast* 'near', *achter* 'behind', etc.

³⁴The problem remains even if functional prepositions can be reanalyzed as case markers.

If (52) features a spatial PP, one might then wonder why coreference between the subject and the object of the preposition is required, since this is not generally true of cases involving spatial PPs. I won't try at this point to develop an analysis of (52) that would capture this restriction, but I think it is instructive to compare cases like (52) to similar examples in English discussed by Déchaine et al. (1995) (see also Rooryck and Vanden Wyngaerd (2007)). Déchaine et al. (1995) consider cases like (54), involving a possessive *have* and a spatial PP.

- (54) a. Mary has \$5 on her.
b. John has dirt on him.

An interesting property of such constructions is that the PP must contain a pronominal bound by the subject of *have*. Déchaine et al. (1995) account for this fact in the following way. They assume that *have* is morphologically complex, consisting of an (incorporated) abstract preposition and a form of BE, i.e. HAVE=P+BE. The BE-component is assumed to be a partitive operator, which is itself bound by the subject, but must at the same time bind a variable in its scope. The latter can happen under the following circumstances: (i) either the complement of *have* is a small clause containing a bound pronoun, as in (54), or (ii) an inalienably possessed noun is present. Such nouns, according to Déchaine et al. (1995), contain an empty argument position that must be bound by the possessor. Note that the latter option is also instantiated in Dutch. In cases like (52), *zich* can be replaced by an inalienably possessed noun, as illustrated in (55) (Eric Reuland p.c. via Peter Svenonius).

- (55) Jan heeft het boek bij de hand/onder z'n arm
Jan has the book by the hand/under his arm

I conclude therefore that the obligatory coreference observed in (52) is a property of the construction in question, and can, as we have just seen, be observed in similar constructions in other languages.³⁵ Whatever the correct analysis of the obligatory coreference in English examples like (54) turns out to be, I believe that the same analysis should be extendable to Dutch facts. Given these considerations, it seems highly plausible to me that examples such as (52) involve spatial PPs and therefore do not present a counterexample to our prediction concerning functional prepositions. I postpone a more extensive analysis of constructions such as (52) to a future occasion.

Let us now turn to non-reflexive, local binding configurations, i.e. to examples where the antecedent and the anaphor are not arguments of the same predicate but

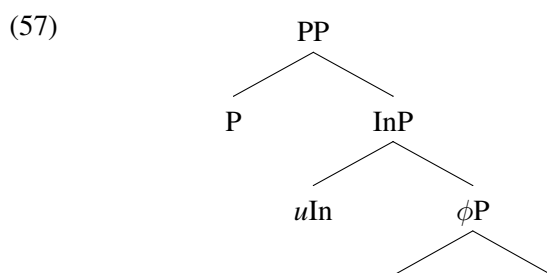
³⁵There are some differences concerning animacy effects between the relevant constructions in Dutch and English, which remain to be explained. Namely, Dutch does not have the analogue of possessive *have* with inanimate subjects, as pointed out by Rooryck and Vanden Wyngaerd (2007).

- (i) a. The table has a lamp on it.
b. *De tafel heeft een lamp erop/op zich.
the table has a lamp thereon/on SE

still appear within the same phase. Sentences involving spatial PPs are one relevant set of cases under the broadly-shared assumption that spatial Ps form their own predicate (see for instance, Marantz (1984), Baker (1988), Hestvik (1991), Reinhart and Reuland (1993), etc). Examples in (56) present the distributional pattern of anaphoric expressions in Dutch spatial PPs. The SE anaphor is the default choice. Some dialects of Dutch resemble English in that they also allow pronominals in these contexts. In that case, a destressed pronominal is used (examples from Rooryck and Vanden Wyngaerd (2007) and Koster (1987), respectively).

- (56) a. Jan_i zag een slang naast zich_i/%'m_i.
 John saw a snake near SE/'im
 'John saw a snake near him.'
 b. Jan_i keek achter zich_i/%'m_i.
 Jan looked behind SE/'im.

The distributional pattern observed in PPs is consistent with the approach developed here given that we are not dealing with reflexive configurations. The anaphor is the argument of the P, rather than of the verb. Hence, the sentences in (56) do not represent configurations where a binding dependency is established between two arguments of the same predicate. The presence of the [R]-feature is therefore not strictly required. In other words, there is nothing wrong with P embedding an InP (as shown below), rather than an RP.



The featural composition of the lexical entry for *zich* matches the features present in the syntactic structure, thus we predict that *zich* should surface in non-reflexive local binding configurations.

The issue of why a pronominal is in some dialects also licit in spatial PPs must be postponed until the following section. It will be suggested there that the distribution of pronominals in contexts such as these is regulated by an economy condition. As a result, pronominals will under certain conditions be allowed in configurations of this type.

As we have seen, contexts involving spatial PPs do not qualify as reflexive configurations, and therefore the presence of [R]-feature is not required in order to obtain a well-formed output. However, nothing we have said so far rules out the presence of [R]-feature either. Recall that the failure to employ the [R]-feature in configurations

where two arguments of the same predicate are coindexed leads to an illicit interface representation because the arguments become formally indistinguishable. The [R]-feature which we argued semantically denotes a function that maps x onto an object that can stand proxy for x , enables one to circumvent this violation. When the coindexed phrases are not arguments of the same predicate, an interface violation of this type does not arise. Hence, the [R]-feature is not required to obtain a licit output, but its presence is not ruled out. Thus, we expect that a SELF anaphor could in principle also surface in configurations involving spatial PPs. Rooryck and Vanden Wyngaerd (2007) note that *zichzelf* can indeed occur in these contexts. In such cases, the use of *zichzelf* gives rise to dissociation or proxy readings, which is consistent with our semantic analysis of the [R]-feature. Thus, (58a) is best interpreted as meaning that John saw the snake behind a representation of himself, such as on a picture or a video. Some additional dissociation contexts are provided in (58b) and (58c) (all examples from Rooryck and Vanden Wyngaerd (2007)).

- (58) a. Jan_{*i*} zag een slang naast zichzelf_{*i*}.
John saw a snake near SELF
- b. (In mijn droom) zag ik Robert_{*i*} naast zichzelf_{*i*}/*zich_{*i*}.
in my dream saw I Robert near SELF/SE
'In my dream, I saw Robert next to himself.'
- c. Piet_{*i*} viel over zichzelf_{*i*}.
Piet fell over SELF
'Piet fell over himself.'
(fine with *zichzelf* a statue)

Let us now turn to ECM and raising constructions which constitute another set of non-reflexive, local binding contexts. I will just briefly discuss the distribution of anaphoric expressions in Dutch ECM constructions, deferring a more detailed investigation of these configurations for some future occasion. According to the literature, both *zich* and *zichzelf* are licit as ECM subjects.

- (59) a. Jan_{*i*} hoorde zichzelf_{*i*} / zich_{*i*} zingen. (Reuland (2001))
Jan heard SELF / SE sing
- b. John_{*i*} zag zichzelf_{*i*} / zich_{*i*} vallen. (Koster (1987))
John saw SELF / SE fall

The ECM subject is not an argument of the matrix predicate. Since only in cases where two arguments of the same predicate are in a binding relation must the [R]-feature be employed, we again predict that the [R]-feature is not obligatorily present, but not excluded either. Consequently, both *zelf* and *zichzelf* might appear. In cases where *zelf* surfaces, the syntactic structure contains the [R]-feature. When only *zich* appears, the [R]-feature is absent. Since both forms are available, the presence or absence of this feature may yield interpretational differences. Thus, the contexts in which *zelf* occurs again favour proxy readings, as already noted in 8.2.5.2. I repeat

the relevant pair of examples from Reuland (2001). Recall that in (60a), the preferred reading is the one where the subject sees her own reflection in the mirror, i.e. the person seeing and the one being seen are indistinguishable, whereas in (60b) the subject sees a representation of herself, e.g. a statue.

- (60) a. Ze_i zag $zich_i$ in een griezelige hoek staan.
 she saw SE in a creepy corner stand
 b. Ze_i zag $zichzelf_i$ in een griezelige hoek staan.
 she saw SELF in a creepy corner stand

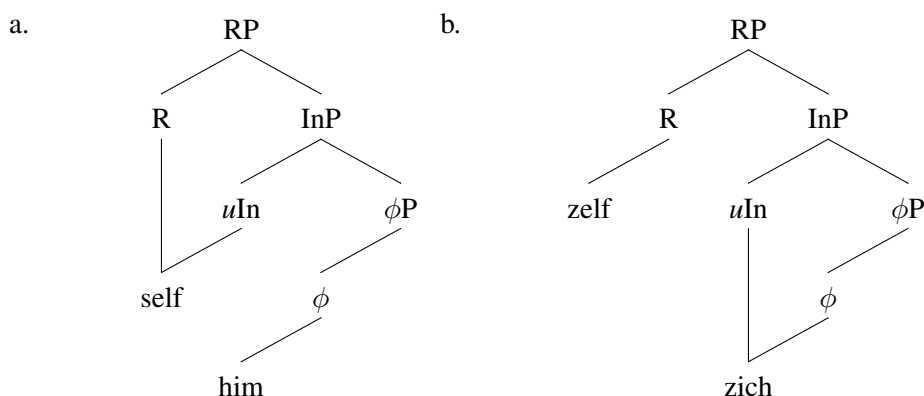
Interpretational effects of this type have also been studied by Voskuil and Wehrmann (1990) and Rooryck and Vanden Wyngaerd in a number of papers. (61) provides another example illustrating the same point, taken from Rooryck and Vanden Wyngaerd (1999). When *zich* is used the preferred reading is the one where Noam simultaneously speaks and hears himself through headphones. By contrast, the use of *zichzelf* is more appropriate if Noam either listens to a temporally dissociated broadcast of himself, or he is unaware of his own speech, as when drunk or dreaming or otherwise dissociated from his usual self.

- (61) $Noam_i$ hoorde $zich_i/zichzelf_i$ over honkbal praten.
 Noam heard SE/SELF about baseball talk
 'Noam heard himself talk about baseball.'

8.2.5.4 Featural composition of English reflexives

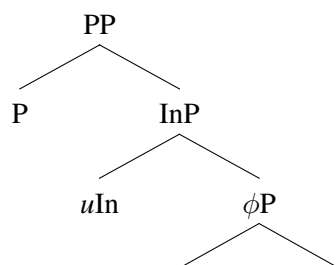
Turning back to English, we must now reconsider the featural specification of English reflexives in light of the revisions made in this section. Given that English anaphors occur in reflexivization contexts, they must be able to spell out the [R]-feature. I conclude therefore that like its Dutch counterpart, the English *self* morpheme has the [R]-feature in its feature specification. However, unlike Dutch, English does not have a distinct form lexicalizing the [u In]-feature, namely the SE anaphor. Rather, the lexical entry for the morpheme *self* in English contains both [R] and [u In]. The ϕ -features in the nominal sequence are spelled out by the pronominal. This is schematized below. The Dutch lexicalization pattern is given in (62b).

(62)



Consequently, the English morpheme *self* is predicted to surface also in those local binding configurations where only the [*uIn*], but no [*R*]-feature is present, as in ECM constructions or spatial PPs. Thus, in contrast to Dutch, in configurations such as (57), repeated below, a SELF morpheme will surface.

(63)



In such contexts, features spelled out by *self* will constitute a subset of features present in its lexical entry, which is in accordance with the Spell-Out principle we have adopted. The difference in the inventory of anaphoric expressions between Dutch and English has therefore important consequences for the distributional properties of these anaphoric forms.

8.2.5.5 Reinhart and Reuland (1993)

The analysis developed in this section takes as its starting point the observation that cross-linguistically reflexive constructions seem to require special licensing, which in Dutch is instantiated in the choice between a SELF and a SE anaphor. The requirement for special licensing of reflexive configurations lies also at the heart of the extremely influential proposal by Reinhart and Reuland (1993). I will finish off this section by contrasting their account with the one proposed here.

On the analysis pursued here, the observation that a SELF anaphor, or more precisely the *self* morpheme must be used in reflexive configurations reduces to the fact that the proposed [R]-feature is obligatorily present in these contexts. Since neither the pronominal nor the SE anaphor can lexicalize this feature, a SELF morpheme is required (except in cases featuring inherently reflexive verbs). This however does not mean that SELF morphemes will only surface in reflexive configurations. We have seen that the analysis allows us to capture the fact that SELF anaphors can also appear in non-reflexive local binding contexts, such as ECM constructions and cases involving spatial PPs. The situation is different on R&R's view.

Reinhart and Reuland (1993) assume that a reflexive predicate must be reflexive marked and argue that SELF anaphors function as reflexive markers. Recall that a predicate is reflexive if two of its arguments are coindexed. On R&R's view therefore, whenever we see a SELF anaphor, we are either dealing with a reflexive configuration, or the predicate is not reflexive and the anaphor is logophoric. Given these options, the strategy they adopt for dealing with SELF anaphors in ECM constructions is to argue that ECM constructions qualify as reflexive configurations, while SELF anaphors embedded in spatial PPs are treated as logophoric. Both solutions are problematic, as I will argue in the rest of this section. The first requires considerable complications of the theory, and particularly of the definition of argumenthood. The second, as I have already argued, wrongly pushes PP-anaphora out of the domain of grammar proper and into the realm of the discourse component.

Consider first cases involving ECM verbs. As R&R note, when they function as ECM subjects, anaphors do not show properties associated with logophoric uses of anaphora and therefore cannot be treated as logophors. On their account, the only remaining option is to analyze them as markers of reflexivity in these contexts. However, assuming a standard definition of argumenthood in terms of theta-marking, the matrix predicate in (64) would not be considered reflexive as the ECM subject is not theta-marked by the matrix verb. Thus sentences such as (64) seem to violate their Condition A, since the anaphor would reflexive-mark the predicate, but the predicate would not be reflexive.

(64) Max_i expects himself_i to pass the exam.

(65) *Condition A*
A reflexive-marked syntactic predicate is reflexive.

To circumvent this problem, R&R propose an alternative definition of argumenthood, assumed to be relevant for the Binding theory. First, they argue that Condition A applies to *syntactic* predicates and then suggest the following disjunctive definition of *syntactic arguments*:

(66) The *syntactic arguments* of P are the projections assigned θ -role or Case by P.

With these modifications, ECM cases such as (64) no longer pose a problem. The

ECM subject does not receive its theta-role from the matrix predicate, but it is assigned case by it, which according to R&R's definition, suffices to make ECM subjects count as arguments of the matrix predicate. If so, then Condition A is met in (64) given that the syntactic predicate of *expect* has two of its syntactic arguments coindexed.

In contrast to R&R's account, on the analysis pursued here ECM constructions are not treated as reflexive configurations. This creates no problems since SELF anaphors are not limited to appearing in reflexive configurations. The featural specification of the English morpheme *self* allows it to surface also in configurations where no [R]-feature is present, in accordance with the adopted Spell-Out principle. In ECM constructions, the *self* morpheme will obligatorily surface in English not because it is spelling the [R]-feature, but because it is lexicalizing [μ In]. In addition, the [R]-feature itself, which is obligatory in reflexive configurations, can also appear in non-reflexive constructions and thereby cause a SELF-morpheme to surface.

For languages which, unlike English, make a two-way distinction between SE and SELF anaphors, the analysis I have been pursuing also predicts that a SE anaphor should be licit in ECM contexts given that these are not reflexive configurations. As we have seen, this prediction is borne out in Dutch. For R&R, the fact that a SE anaphor can appear in ECM constructions constitutes one of the major reasons for assuming that Conditions A and B do not apply to the same types of predicates, as stated below (*italics mine*).

(67) *Condition A*
A reflexive-marked *syntactic* predicate is reflexive.

(68) *Condition B*
A reflexive *semantic* predicate is reflexive-marked.

If Condition B, like Condition A, also applied to syntactic predicates, then it would rule out the SE anaphor in ECM constructions like (69).

(69) Henk_i hoorde zich_i zingen.
Henk heard SE sing

Zich here counts as a syntactic argument of the matrix verb since it is assigned case by it. Therefore, this is a reflexive construction in R&R's view and should be reflexive marked, but *zich* cannot function as a reflexive marker. This problem is avoided by making Condition B apply to semantic, rather than syntactic predicates. In (69), the whole embedded IP, rather than its subject, is the semantic argument of the matrix verb. Thus, there is no reflexive semantic predicate and reflexive marking is not necessary, making *zich* licit in these configurations.

Turning to configurations which involve spatial PPs, R&R treat SELF anaphors in these environments as logophoric in nature. As I have already extensively argued in the previous chapter, this seems to be a problematic conclusion for English, and one that leads to complications of the theory of variable binding. Moreover, it is

called into question by the fact that in certain languages anaphors which can function as reflexive markers also occur in PP-environments even though they generally do not have logophoric uses. In contrast to R&R's proposal, the analysis pursued here allows us to subsume PP-anaphora under cases of local binding.

8.2.6 Summary

In this section, I have argued that the behaviour of anaphors can be understood on the basis of the features they possess and the way these interact with properties of the grammatical system. The proposal takes as its starting point the semantic distinction between variable binding and coreference in interpreting anaphoric relations. As pointed out by Reinhart (1983b) and much subsequent work, the failure to distinguish bound variable from coreferential uses of anaphora has led to many stubborn problems in the canonical binding theory. Binding theory, Reinhart (1983b) argues, should only regulate bound variable uses of anaphora. Coreference, on the other hand, is subject to the principles of the discourse component and beyond the purview of grammar proper. This is the view I have adopted in this and the preceding chapter. I have then assumed that a combination of [*u*In] and unvalued ϕ yields a bound variable dependency at the interface. The ϕ -features must be valued in the course of the syntactic derivation in order to be interpretable at the interface. This establishes a syntactic dependency between an anaphor and the antecedent. The [In]-feature is interpreted at the interface as the instruction to assign an integer to the noun phrase bearing it. When an [*u*In] is inserted into the derivation, as in the case of anaphors, it will not be visible at the interface. In such cases, the translation procedure mapping syntactic to semantic representation ensures that the anaphor receives the index of the noun phrase it enters into ϕ -agreement with. The anaphor thus always ends up coindexed, and semantically bound, by its antecedent. The requirement that anaphors find antecedents follows from the need to establish the relevant Agree-relation between the features lexicalized by the antecedent and those lexicalized by the anaphor. Since Agree is constrained by the PIC, it also follows that the phase should act as the local binding domain for anaphors, as argued extensively in the previous chapter. If no local antecedent is available, i.e. no antecedent within the same phasal domain, only the derivation which starts with an [*i*In] will yield a licit interface representation. [*i*In] can then be lexicalized by the pronominal at the point of Spell-Out. No syntactic locality effects are expected in this case, given that no Agree relation is established between the binder and the bindee in the course of the derivation. As we have seen, this is the correct result since pronominals can be associated with their antecedents across phasal boundaries, and even syntactic islands.

The distribution of anaphors and pronominals is thus tightly linked to their featural make-up and general principles of syntactic computation. Relying on assumptions about the featural composition of particular lexical items paves a way for capturing at least some of the differences in distributional patterns between various anaphoric expressions across languages. Thus, while in English the pertinent [*u*In]-feature is

argued to be part of the featural specification of the *self*-morpheme, forcing its presence in all local binding configurations, in Dutch this feature is arguably lexicalized by the anaphor *zich*, rather than by *zelf*, the cognate of the English SELF-morpheme. Consequently, in a number of local binding environments in Dutch, a SELF anaphor is not required and it suffices to use *zich* instead.

Investigation of anaphoric dependencies in Dutch led us also to further refinements regarding how binding relations are encoded. It has been shown that two types of local binding configurations must be distinguished, those where the anaphor and its antecedent are coarguments of the same predicate, and those where they are not, but are still within the same local domain, namely the phase. Following a relatively well-established tradition, I have used the term *reflexive* configuration to refer to the former type. In languages that make a distinction between SE and SELF anaphors, the SELF anaphor must appear in reflexive configurations. In other local binding environments, either anaphoric expression can be used.

On the analysis developed here, the obligatory presence of the SELF anaphor, or more concretely the SELF morpheme, is tied to the obligatory presence of a particular feature, labelled as [R], in reflexive configurations. Adopting the view that lexical insertion happens post-syntactically and assuming that the lexical entry for the SELF-morpheme contains the proposed [R]-feature, the SELF morpheme is predicted to occur whenever the [R]-feature is present in syntactic structure. The relevant question to ask then is why [R]-feature must be present in reflexive configurations. In answering this question, I have followed an earlier proposal by Reuland (2001), who argues that coreference between two arguments of the same predicate results in an illicit interface representation where the two arguments can no longer be distinguished as distinct semantic objects. In other words, cases where $P(x,x)$, i.e. where the predicate takes two tokens of the same variable, are illicit. The violation can be circumvented, while retaining the bound reading, if the predicate takes two arguments, one identical to x , and the other to the value of some function of x , i.e. $P(x, f(x))$. This is achieved by SELF-marking. Adopting this view, I have thus argued that the proposed [R]-feature, lexicalized by a SELF morpheme, denotes a function which maps x onto an object that can stand proxy for x , i.e. it yields a ‘representation’ of x . The [*uIn*]-feature is still needed though to ensure anaphoric dependency.

The proposed account allows us to capture the fact that a SELF morpheme must surface in reflexive configurations, but at the same time it does not predict that these are the only environments in which a SELF morpheme can appear. In non-reflexive contexts, the [R]-feature is not required, but not ruled out either. Consequently, a SELF morpheme may surface also in non-reflexive configurations, whenever this feature is present in the syntactic structure. In Dutch, where both *zich* and *zichzelf* are available in non-reflexive contexts, the choice of one anaphoric expression over the other may yield interpretational differences. In such cases, the use of *zichzelf* yields a dissociation or a proxy reading, which is consistent with the semantic analysis of the [R]-feature which *zelf* lexicalizes. In English, the *self*-morpheme is assumed to be able to lexicalize both the [R]- and the [*uIn*]-feature, and is therefore predicted

to occur also in contexts where no [R]-feature is present. Such a treatment of English reflexives allows us to rather straightforwardly account for their presence in ECM/raising constructions and contexts involving spatial PPs, without encountering problems that Reinhart and Reuland's (1993) analysis faces.

8.3 The issue of (non)complementarity between anaphors and pronominals

In the canonical binding theory, the distribution of anaphors and pronominals is regulated by the following binding conditions:

- (70) *Binding Condition A*
An anaphor is bound in its governing category.
- (71) *Binding Condition B*
A pronominal is free in its governing category.

So far I have focused almost exclusively on anaphors, arguing that their distribution can be reduced to their featural make-up, thereby eliminating Condition A. Anaphors surface in cases where the derivation starts with the combination of [μ In] and unvalued ϕ . The ϕ -features are valued in the course of the syntactic derivation through an Agree relation with a local antecedent. Given that the application of Agree is constrained by the PIC, the locality conditions on anaphoric binding follow. In other words, the proposal predicts that the phase is the local domain for binding of anaphors, as argued extensively in chapter 7.

What about Condition B effects? In the previous sections, I have made explicit my assumptions regarding the featural specification of pronominals. One might ask whether the distribution of pronominals can also be made to follow from their featural composition. Given that we have shown Condition A to be eliminable, maintaining Condition B would be highly undesirable. In this section, I will investigate to what extent the distribution of pronominals can be derived from their featural specification. We will see that although relying solely on the featural specification and the theory of lexical insertion that we have adopted can take us a long way in accounting for the distribution of pronominals, these factors alone are not sufficient to capture all Condition B effects. I will then briefly discuss what additional factor(s) might be involved, leaving a more detailed account of Condition B effects for another occasion.

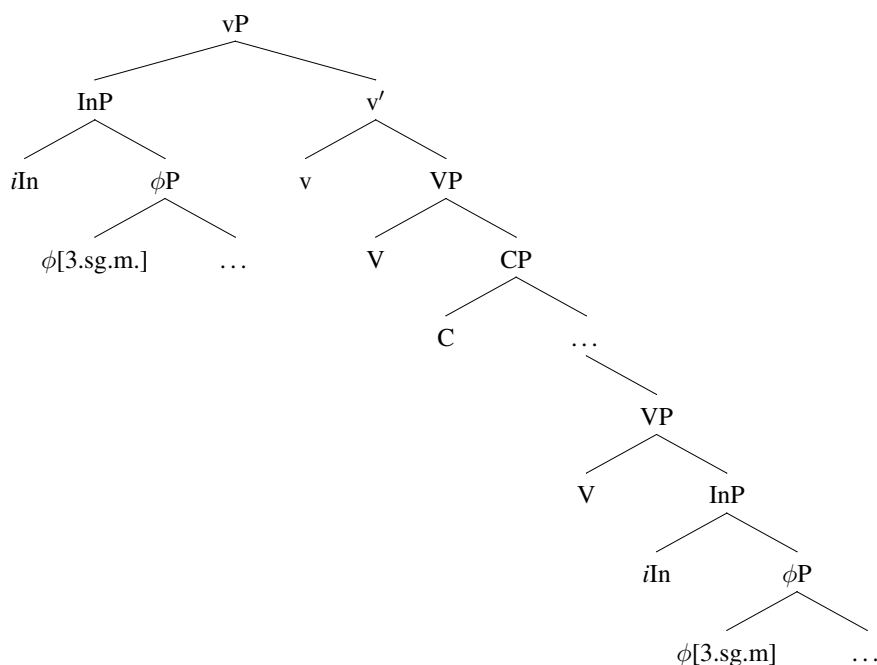
8.3.1 Non-local binding configurations

In configurations where a binding relation spans a phasal boundary, as in (72), the bindee takes the form of a pronominal.

(72) John_{*i*}/[Every classmate]_{*i*} knows that we invited him_{*i*}/*himself_{*i*}.

The model developed here predicts that this should be the case. Recall that an anaphor will surface in configurations where [*uIn*] is present, given that it bears [*uIn*] in its featural specification.³⁶ Since the [*In*]-feature is uninterpretable, an Agree-relation in ϕ -features must be established in the course of the syntactic derivation in order to obtain a licit interface representation, as discussed in detail in the previous section. The required syntactic dependency cannot however be established in cases like (72) since there is a phase boundary intervening between the binder and the bindee and Agree is constrained by the PIC. Therefore, in order to get the object of *invite* in (72) bound by the matrix clause subject, the object of *invite* must enter the derivation bearing [*iIn*], as sketched below.

(73)



[*iIn*] can be directly mapped onto an integer value at the interface, and no syntactic dependency with the antecedent needs to be established. Consequently, the fact that the binder and the bindee are separated by a phasal boundary causes no problems. At the point of Spell-Out, the pronominal will lexicalize both the [*iIn*] and the ϕ -features, in accordance with its featural specification. Recall that on the assumptions spelled out in the previous section, pronominals contain at least [*iIn*] and [ϕ]-features in their featural specification. Since in cases like (72), the features present in the syntactic structure match the features that the lexical entry for the pronominal contains,

³⁶More precisely, [*uIn*] is part of the featural specification of the morpheme *self* in English, and the morpheme *zich* in Dutch.

a pronominal will be inserted in this configuration. Note that the *self*-morpheme cannot be inserted in this case since the features it is lexically specified for, namely [*uIn*] and [R], are not present in the syntactic representation of (72). At the interface, the configuration in (73) may give rise to a bound variable interpretation, as discussed in the previous section. We have already seen that the interpretative relation at the interface is not sensitive to syntactic locality restrictions. The analysis pursued here thus correctly predicts that a pronominal should be used in nonlocal binding configurations, i.e. configurations where the binder and the bindee are not within the same phase.

The situation is different when it comes to local binding contexts. We have seen in the previous section that two types of local binding configurations must be distinguished: reflexive and non-reflexive ones. I will discuss each of these in turn.

8.3.2 Reflexive local binding configurations

In section 8.2.5, we have distinguished two types of local binding configurations: those where the phrases in a binding relation are arguments of the same predicate and those where they are not, but where they still occupy the same phasal domain. In a language like Dutch, the distinction is reflected in the choice of the anaphoric type. Namely, in reflexive configurations a SELF anaphor must be used, whereas in other local binding contexts both a SELF and a SE anaphor is in principle possible. We have captured this distributional difference by postulating the presence of an additional feature in reflexive configurations, namely the [R]-feature. Failure to employ this feature in cases where the binding relation involves two arguments of the same predicate has been argued to result in an illicit interface representation. Assuming then that the pertinent feature is lexicalized by the SELF morpheme, we correctly predict that a SELF anaphor would be obligatory in such environments. In other local environments, the presence of [R]-feature is not required and therefore a SELF-morpheme need not surface. The presence of [*uIn*] ensures the establishment of an anaphoric dependency. In Dutch, we have argued, [*uIn*] is spelled out by the SE anaphor, *zich*.

In English, the distinction between reflexive and non-reflexive local configurations is blurred by the fact that English only has one type of anaphors, namely SELF anaphors. The *self*-morpheme in English lexicalizes both the [R] and the [*uIn*]. Thus, it is predicted to appear in all local binding configurations. Assuming that pronominals do not have the [R]-feature in their feature specification, we also predict that the *self*-morpheme is obligatory in reflexive configurations. An example of the relevant context is repeated below.

(74) John_{*i*} likes himself_{*i*}/*him_{*i*}.

Given that the binder and the bindee are coarguments in (74), the presence of the [R]-feature is required to obtain a licit result. Since only the *self*-morpheme, but crucially not the pronominal, can spell out the [R]-feature, it is predicted that the

self-morpheme will surface in cases such as these.

The proposed analysis therefore captures a subset of Condition B effects by relying on the featural composition of pronominals and the *self*-morpheme. In local binding configurations involving two arguments of the same predicate, the need to lexicalize the [R]-feature ensures that we never see simply a pronominal appearing. The *self*-morpheme is always required in such contexts.

8.3.3 Non-reflexive local binding configurations

Let us turn now to non-reflexive local binding configurations. Recall that these are contexts where the binding relation is established between two phrases which are not arguments of the same predicate but nevertheless occupy the same local domain, namely the phase. Two environments of this type that we will focus on at this point are ECM/raising constructions and spatial PPs. Interestingly and quite surprisingly, we observe different distributional patterns in these two contexts. In ECM and raising constructions, only anaphors are licit, whereas inside spatial PPs both anaphors and bound pronominals can be used in English. Some representative examples are repeated below.

- (75) a. John_{*i*} considers himself_{*i*}/*him_{*i*} to be smart.
 b. John_{*i*} seems to himself_{*i*}/*him_{*i*} to be smart.
- (76) a. John_{*i*} saw a snake near himself_{*i*}/him_{*i*}.
 b. John_{*i*} looked around himself_{*i*}/him_{*i*}.
 c. John_{*i*} pulled the blanket over himself_{*i*}/him_{*i*}.

Within the canonical binding theory, the availability of both anaphors and pronominals in spatial PPs is *prima facie* unexpected. Conditions A and B taken together predict that anaphors and pronominals should be in complementary distribution, i.e. in environments where anaphors are licit, pronominals should be ruled out. This prediction is clearly not borne out in cases like (76). The fact that the expected complementarity breaks down in certain configurations has led to a number of revisions of the classical binding theory (see for instance Chomsky (1986b), Hestvik (1991)).

While cases like (76) are problematic for the canonical binding theory, it is in fact examples such as (75) that are troublesome on the approach pursued here. Let us consider why. Note that the coindexed phrases in (75) and (76) are not arguments of the same predicate: the antecedents are arguments of the matrix verb, while the anaphors/pronominals are arguments of the embedded verb in (75) and of the preposition in (76). Therefore, we are not dealing with reflexive configurations in these cases and the presence of the [R]-feature is not strictly required. This means that the anaphoric relations given above can be encoded solely in terms of the [*u*In]-feature (and ϕ -features). Observe also that the anaphors/pronominals and their antecedents are within the same local domain. In (75), the two do not initially occupy the same local domain, but come to do so as a result of displacement. In (76), the

anaphors/pronominals and their antecedents are within the same domain when both occupy their base-generated positions, given the assumption that PPs do not define phases. Consequently, one possible derivation of (75) and (76) would start with [uIn] and unvalued ϕ . ϕ -features can then be valued through an Agree operation with a local antecedent. At Spell-Out, [uIn] would be lexicalized by the *self*-morpheme. The availability of anaphors in both (75) and (76) is therefore expected.

Consider now an alternative derivation of (75) and (76) that would start with a valued set of ϕ -features, accompanied by an interpretable [In]-feature.³⁷ In that case, no syntactic dependency would be created between the binder and the bindee. At Spell-Out, the bindee could then take the form of a pronominal, given that, in accordance with its feature specification, a pronominal can lexicalize [iIn] and ϕ . At the interface, this configuration can result in a bound-variable relation. Nothing we have said so far rules out such a derivation. When the derivation starts with an interpretable [In], it is thus irrelevant whether the potential antecedent is in the same local domain or not since no syntactic dependency needs to be established. Consequently, we predict that in cases like (75) and (76), both anaphors and pronominals should in principle be available. This prediction is however not borne out in (75). Hence, on the approach pursued here, the question is how to rule out pronominals in configurations such as (75) and how to capture the difference between (75) and (76). It is clear that relying solely on the featural composition of anaphors and pronominals does not suffice to capture their distribution in non-reflexive local binding configurations.

Following some earlier suggestions, in the following subsection I will argue that this residue of Condition B might be captured by appealing to an economy condition. In 8.3.3.2, I will then revisit PP-binding environments and briefly discuss what factors might be responsible for yielding different distributional patterns in (75) and (76).

8.3.3.1 Economy considerations

A recurring idea in the literature among those trying to eliminate Condition B is to reduce it to an economy condition of a sort (Taraldsen (1996), Koster (1997), Hornstein (2001), Reuland (2001), Hicks (2009), etc.). On these accounts, pronominals are considered less economical for some reason or other, and therefore disfavoured in contexts where an anaphor is licit. Approaches such as these are consistent with the emphasis placed on economy in the Minimalist Program.

Koster (1997), for instance, formalises this idea as the *Principle of Maximal Specialization*, a general grammatical principle which subsumes Condition B.

(77) *Principle of Maximal Specialization*

In a grammatical dependency relation R, select the most specialized form.

A form is more specialized if it ‘can fulfil fewer functions’. The consequence of this is that reflexives, which are more specialized than pronominals, will be chosen in

³⁷Recall that combining [uIn] with a valued ϕ would yield an illicit interface representation. Refer back to section 8.2.2 for discussion.

contexts where they are licensed. One concern that this approach raises is how ‘being a more specialized form’ is to be properly defined so that the desired empirical facts are derived.

Hicks’s (2009) economy-based elimination of Condition B is a formalization of the idea that ungrammaticality in Condition B configurations arises from the fact that Agree *could* have applied were it not for the fact that pronominals bear valued [VAR] features and so are incapable of triggering Agree. He takes as his starting point Ćitko’s (2006) proposal regarding Across-the-Board (ATB) movement constructions in Slavic. Namely, Ćitko suggests that economy prefers derivations arising from numerations with fewer instances of the same lexical items over those with more. Such derivations result in maximal structure-sharing. Hicks suggests that this preference for structure-sharing should be broadened to a preference for feature-sharing. In particular, feature sharing is maximised not only when the number of tokens of identical lexical items is reduced, but also when as many features as possible enter the derivation unvalued. He proposes that this type of economy consideration can be formalized in the following way:

- (78) *Maximize Featural Economy*
 Establish dependencies via syntactic operations where possible.

Hicks then argues that this economy condition ensures that when two matching features with identical values are in a configuration in which Agree can apply between them, Agree must have applied, establishing feature-sharing by syntactic means. Recall that on Hicks’s view bound pronominals do not receive the value of the [VAR]-feature in the course of the syntactic derivation, but just happen to enter the derivation bearing the same value of the [VAR]-feature as their antecedents. Consider then (79). Both the pronominal and its local antecedent enter the derivation with an identical value of the [VAR]-feature. However, if the [VAR]-feature were unvalued, precisely the same output could be derived by Agree, so (78) applies to block (79). Employing a pronominal that enters the derivation with a [VAR] already specified with the same value as its local antecedent is ruled out as uneconomical.

- (79) *John_i loves him_i.
 [TP John_[Var:x] [VP < John_[Var:x] > likes him_[Var:x]]]

Though the intuition behind Hicks’s approach is clear and not novel, the downside of Hicks’s proposal is that the formulation in (78) is rather vague. This particularly concerns the clause “where possible”. It is also not obvious that the step from Ćitko’s economy condition to the one that Hicks proposes is as straightforward as Hicks presents it to be. While it seems intuitively plausible that having a numeration with a single token of a lexical item would be more economical than employing several tokens of the same item, it is more difficult to see why choosing unvalued features and then valuing them in syntax should be more economical than simply choosing the valued version of the same feature. Hicks offers no discussion of this

point. Note that the condition in (78) seems to imply that establishing dependencies by syntactic means is more economical than establishing dependencies via some other mechanisms. However, it is unclear what other mechanisms (78) is intended to rule out. What is more, if I understood Hicks's proposal correctly, there is in fact no formal dependency whatsoever between the pronominal and its antecedent and thus no mechanism for encoding dependencies is required: upon the selection for the numeration, a particular value for the [VAR]-feature borne by the pronominal is chosen and this value just happens to be the same as the value of another c-commanding DP.

An earlier proposal by Reuland (2001) is more explicit in this respect. According to Reuland, the relevant 'cost' is linked to the component of grammar involved. In particular, he argues that syntactic processes are more economical than processes at the interface, and the latter are more economical than processes involving the discourse storage. The proposal builds on Reinhart (1983b) and Reinhart and Grodzinsky (1993), who discuss the availability of bound variable and coreferential interpretations of pronominals. Observe that an example like (80) has in fact two representations at the interface, one where *him* is a variable bound by *Bill*, the other where *him* is referential.

- (80) **Bill* adores *him*.
 a. Bill λx (x adores x)
 b. Bill λx (x adores a)

As Reuland points out, since the value of *a* can be freely chosen, one interpretation of (80b) is (81).

- (81) Bill λx (x adores Bill)

According to Reinhart (1983b) and Reinhart and Grodzinsky (1993), Binding theory regulates only the bound variable interpretation. Coreference is a matter of the discourse component. Thus, Condition B can be called upon to rule out (80a), but not (80b). Although we might want to rule out (81) in general, the acceptability of examples such as (82) shows that it would be wrong to exclude this interpretation under all circumstances.

- (82) a. I know what Mary and Bill have in common. Mary adores him and *Bill* adores *him* too.
 b. I dreamt that I was Brigitte Bardot and *I* kissed *me*. (due to George Lakoff, discussed in Heim (1998))

Reinhart (1983b) and Reinhart and Grodzinsky (1993) thus formulate the following interpretative condition to express when the coreference option is blocked:

- (83) *Rule I: Intrasentential Coreference*
 NP A cannot corefer with NP B if replacing A with C, C a variable A-bound by B, yields an indistinguishable interpretation.

The rationale behind (83) is that coreference is not grammatically impossible in (80), but this option cannot be used because another available process is preferred for independent reasons.³⁸ Reuland (2001) points out that Rule I uses in fact an economy principle, reflecting a division of labour within the linguistic system: the process of encoding a dependency in the semantic structure by variable binding is more readily accessible than establishing coreference by using the discourse storage. Reuland then argues that the choice of anaphors over pronominals involves the same logic, but different components - namely, operations in narrow syntax versus operations at the interface. Given that both pronominals and anaphors can function as bound variables and thus yield the same interface representations, logical binding should be compared with some other way of establishing a dependency, which, in the end, should be preferred. The alternative mechanism in question is feature-valuation in narrow syntax. This intuition is then stated as Rule BV, where *R* stands for a syntactic feature-checking relation, an A-bound in this context refers to Reinhart's (1995) logical-syntax-based definition of binding.³⁹

- (84) *Rule BV: bound variable representation*
 NP A cannot be A-bound by NP B if replacing A with C, C an NP such that B R C, yields an indistinguishable interface representation.

Taken together, Rule I and Rule BV imply a ranking, with processes within narrow syntax being most and those in the discourse component least economical. Other considerations also speak in favour of such a ranking.⁴⁰ It is often assumed that processes in the narrow syntax are automatic, and hence plausibly cheap. As Reuland points out, computations within the interpretative component are automatized to a lesser extent, as evidenced by difficulties that speakers face in interpreting complex quantificational structures. The same has been argued for processes involving the discourse storage, as these require access to various linguistic and non-linguistic contextual information. Rule I and Rule BV are thus a consequence of more general economy considerations that prefer the establishment of dependencies in one linguistic component over the other.

I will follow Reuland (2001) in assuming that the choice between anaphors and pronominals reflects a more general preference for employing narrow syntactic, rather than interface operations in establishing dependencies. In (85a), the dependency be-

³⁸The process in question is the "early closure" of an open expression. See Reinhart (1983b) for details.

³⁹(i) gives Reinhart's (1995) definition of binding:

- (i) *A-binding*
 α A-binds β iff α is the sister of a λ -predicate whose operator binds β .

⁴⁰Reuland (2001) argues that the relevant ranking need not be stipulated, but can in fact be derived and proposes an economy metric that is intended to achieve that. I refer the reader to Reuland (2001) for details regarding this economy metric.

tween the two coindexed noun phrases is expressed within narrow syntax, by means of the Agree operation involving ϕ -features. In (85b), the dependency is expressed in the semantic structure, but not encoded in the syntactic component. Deriving (85c) from (85a) via a syntactic Agree-relation is thus preferred for reasons of economy.

- (85) a. John_i considers himself_i to be smart.
 b. *John_i considers him_i to be smart.
 c. John λx (x consider (x smart))

On this view, the derivation yielding (85c) from (85b) does not crash, as no rule of the narrow syntactic component is violated. As Reuland (2001) points out, this is a welcome result because only if it converges can such a derivation block (86). (86) represents the coreferential interpretation, which involves accessing the discourse storage. If (86) were allowed, examples such as (85b) would be predicted grammatical on a coreferential interpretation.

- (86) John λx (x consider (a smart)) & a=John

Note that economy conditions only compare convergent derivations (Chomsky (1995), pp. 220-221). A convergent derivation could never compete with a non-convergent one. Hence, a derivation yielding a pronominal will be licit from the point of view of economy, whenever the relation between the anaphor and its antecedent cannot be syntactically encoded. In these instances pronominals can, and must be used instead of anaphors.

Before concluding this section, let me add a few comments regarding the kind of economy we must appeal to in these cases. Observe that economy considerations the proposed account relies on are not local in character. I have argued that in certain cases both the derivation yielding a pronominal and the one yielding an anaphor converge and the decision about which derivation to choose is made on the basis of information available at the interface. In the current minimalist theorizing, the status of global economy conditions, such as this one would be, is a matter of debate. It has been argued both on conceptual and empirical grounds that only local economy conditions should be permitted, i.e. those that determine well-formedness of a derivation D by relying solely on information available in the set of syntactic objects that are part of D, with no appeal to alternative derivations, convergent or non-convergent ones, allowed. Such a view arose in particular as a reaction to the extensive role that economy considerations played in the earlier versions of the Minimalist program (see for instance Chomsky (1995)), which gave rise to growing concerns that incautious use may lead to computational blowup. Collins (2005), one of the fervent advocates of local economy, has shown that a number of empirical phenomena previously argued to involve global economy considerations could be successfully reanalyzed in terms of local economy. Nevertheless, global economy conditions are still appealed to in many recent accounts of various linguistic phenomena (see for instance Reinhart (2000), Fox (2000), Rezac (2007), etc.). Reinhart (2000) and Fox (2000) make

a particularly strong case, in my opinion, that at least in the domain of binding phenomena recourse to global economy (i.e. reference-set computation) is required. It is yet to be shown whether facts they discuss, which involve properties of the interface such as scope-taking and coreference, could successfully be integrated into purely local computation. For now, cases such as these might provide the necessary empirical justification that a computationally more complex comparison of derivations is needed. As already noted, supporting evidence for this view is provided by increased processing difficulties associated with derivations involving reference-set computation, as reported in acquisition studies. If the derivations of the pertinent examples indeed involve a more complex strategy, rather than a mechanical syntactic rule, then this might explain the observed processing difficulties that arise with such examples.

A common conceptual argument against global economy is that it greatly increases computational complexity since alternative derivations must be constructed and compared. Though this is a valid concern, how grave it is depends on how one construes interface access to the computation. On a phase-based view of computation as assumed here, we do not need access to the entire derivation to determine whether an anaphor or a pronominal will be used. It suffices to inspect a phasal domain. It is only within a phase that the pertinent syntactic dependency can be established (in cases where the bindee bears [*uIn*]), resulting in a legitimate output that can be sent off to the interfaces. The relevant part of the structure can then be evaluated for economy. If the valuation cannot be achieved within a phase containing the phrase bearing the unvalued ϕ -features, the derivation will not converge and the only alternative that remains is to resort to the derivation which would yield a pronominal. No appeal to economy is then required. Thus, only a part of the structure needs to be considered, which considerably reduces computational burden. I won't discuss the issue further at this point. For a more detailed justification of the need for global economy conditions within the domain that concerns us here, namely that of binding phenomena, I refer the reader to Reinhart (2000), Fox (2000), Reuland (2001).

8.3.3.2 PP-binding contexts

The economy-based solution sketched out in the previous section captures the previously problematic contrast between the (a) and (b) examples below. Examples where the *self*-morpheme surfaces involve the derivation where the dependency between the anaphor and its antecedent is syntactically coded via an Agree relation in ϕ -features, which in turn is required due to the presence of [*uIn*] in the structure. No such syntactic dependency exists in examples featuring pronominals. Economy considerations then favour derivations yielding anaphors.

- (87) a. John_{*i*} considers himself_{*i*} to be smart.
 b. *John_{*i*} considers him_{*i*} to be smart.
- (88) a. John_{*i*} seems to himself_{*i*} to be smart.
 b. *John_{*i*} seems to him_{*i*} to be smart.

Whereas the distributional patterns observed above are now captured, problems however arise with binding configurations involving spatial PPs. We expect now that economy should block the use of pronominals in PP-binding contexts, on a par with examples involving ECM and raising verbs. Nevertheless, as we have already seen, both anaphors and bound pronominals can appear as complements to spatial PPs in languages like English. Some representative examples are repeated below.

- (89) a. John_i saw a snake near himself_i/him_i.
 b. John_i looked around himself_i/him_i.
 c. John_i pulled the blanket over himself_i/him_i.

One common line of approach in dealing with examples like (89) is to assume that local domains for anaphors and bound pronominals are not the same. Hicks (2009) appeals to this type of solution within the phase-based approach to binding phenomena. Recall that Hicks makes a distinction between LF- and PF-phases. On his view, anaphors must find an antecedent within their minimal LF-phase, while pronominals are sensitive only to the presence of a PF-phase boundary. The ‘core’ phases vP and CP count as both LF- and PF-phases, which, Hicks argues, created the illusion of full complementarity between pronominals and anaphors. However, both anaphors and bound pronominals become available in environments involving spatial PPs. According to Hicks, this happens because PPs are only PF-phases, but crucially not LF-phases. Thus, in cases like (89), the pronominal is acceptable because it remains unbound within its minimal PF phase, namely the PP. On the other hand, since only LF-phases are relevant for the distribution of anaphors, an anaphor is free to search for an antecedent outside the PP. Hence, the use of anaphors in (89) is also licit.

The question that arises on this view is why anaphors and pronominals should care about LF- and PF-phases, respectively. The reasons why anaphors are sensitive to LF-phases have already been discussed in section 7.9. Recall that, on Hicks’s assumptions, anaphors enter the derivation with an unvalued [VAR]-feature which must be valued before the relevant chunk of syntactic derivation is transferred to the interfaces. At the LF-phasal boundary, it is semantic syntactic features that are transferred. Since the [VAR]-feature belongs to this class of features, the sensitivity of anaphors to LF-phases is explained. It is more puzzling however why PF-phases should be the relevant local domains for bound pronominals. Once a PF-phase is complete, morphosyntactic features are sent to the PF-interface. Semantic syntactic features are unaffected. The question then is why the transfer of morphosyntactic features to the PF should affect semantic interpretation, i.e. the binding possibilities of pronominals. To explain the relevance of PF-phases for Condition B, Hicks appeals to the Activity condition. Namely, Chomsky (2000) argues that only DPs bearing unvalued Case features are syntactically active and capable of entering into an Agree operation. Hicks proposes that “the mechanisms giving rise to Condition B effects are sensitive to whether the features of the pronoun are active at the point when another DP sharing the same value for [VAR] enters the derivation” (Hicks (2009), p.

201). But it is unclear why they should be. The answer can't be that a valued case feature makes other features inactive and precludes a successful Agree relation. This is because pronominals enter the derivation with a valued [VAR]-feature (the value of which happens to be the same as that of its local antecedent), and are thus independently incapable of triggering Agree. It is anaphors which enter into an Agree relation with their antecedents, so if anything we should expect anaphors to be sensitive to the value of case features. Nevertheless, Case plays no role in Hicks's account of Condition A effects, but is apparently crucial for Condition B effects.

Given these concerns, one wonders whether the distributional pattern observed in (89) can be accounted for without enriching the typology of phases. Let me briefly sketch two possibilities that are consistent with the assumptions we have adopted so far.

In this section I have argued that economy considerations dictate that a derivation which requires the morpheme *self* to surface is preferred in nonreflexive local binding configurations. Note that only derivations that yield nondistinct interface representations are assumed to enter into economy calculations. This is made explicit in Rule I and Rule BV, both of which require that the derivations considered result in indistinguishable interface representations. If the use of the pronominal in place of an anaphor in PP-binding contexts has distinct interpretational effects, then the two derivations, the one yielding the pronominal and the one yielding the anaphor, would no longer be in competition and both derivations would be allowed.

In fact, it has long been noted that choosing a pronominal or an anaphor in configurations involving binding into PPs can affect interpretation. These interpretational effects have recently been discussed by Rooryck and Vanden Wyngaerd (2007), who formulate several generalizations on the basis of such observations.

Already in the 70s, Cantrall (1974) has observed that the choice of an anaphor or a pronominal may reflect a shift in viewpoint. Consider the following pair of examples from the classic paper by Lees and Klima (1963):

- (90) a. *They placed their guns in front of them.*
 b. *They placed their guns in front of themselves.*

Lees and Klima (1963) mark (90b) as unacceptable. Cantrall (1974) however notes that for many speakers both examples are fine, though with a difference in meaning which can be perceived as a shift in viewpoint. Insertion of a parenthetical indicating viewpoint makes this difference clearer.

- (91) a. *They placed their guns, as they looked at it, in front of themselves / *them.*
 b. *They placed their guns, as I looked at it, in front of *themselves / them.*

According to Rooryck and Vanden Wyngaerd (2007), the generalization that emerges on the basis of these and similar examples is that the use of the pronominal correlates with the speaker perspective, while the use of the reflexive shifts the perspective to

that of the referent of the anaphor.

A different interpretational effect has been noted by Kuno (1987). Consider as an illustration the following pairs of examples:

- (92) a. *John hid the book behind himself.*
 b. *John hid the book behind him.*
- (93) a. *John pulled the blanket over himself.*
 b. *John pulled the blanket over him.*
- (94) a. *John put the blanket under himself.*
 b. *John put the blanket under him.*

According to Kuno, the difference in (92) has to do with whether or not there is physical contact between John and the book. In (92a), the book is in John's hands, touching his back. (92b) on the other hand does not require direct contact between the book and John: the book might be on a chair, with John standing in front of the chair so that the book cannot be seen. (93a) implies that the blanket is over John's entire body. If his head is sticking out, (93b) is preferred. Similarly, (94a) implies that John was trying to hide the blanket by covering it with his body, while (94b) lacks such an implication. Rooryck and Vanden Wyngaerd (2007) argue that considerations of this type support the following generalization: the *self*-form requires a locative relation of concrete, close, total bodily proximity, while the pronominal allows for a more abstract locative relation: vicinity, proximity, partial coverage.⁴¹

Relying on interpretative differences such as those noted above, we might then argue that derivations yielding anaphors and those yielding pronominals in PP-binding configurations do not result in indistinguishable semantic representations. Consequently, the two derivations would not enter into competition and economy would

⁴¹Kuno's generalization is somewhat different, as he tries to extend these observations beyond contexts involving locative PPs. For Kuno, the (non)complementarity between pronominals and anaphors is a consequence of the following semantic filter:

- (i) *Semantic Constraint on Reflexives*
 A [+reflexive] NP that ends with *-self/selves* can be used in English if and only if its referent is the direct recipient of the actions or mental states represented by the sentence.

According to Kuno, both reflexives and pronominals are in principle acceptable in what are standardly viewed as local binding contexts. The reason why reflexive is always used in canonical cases of local binding such as (ii) is because in such contexts the referent is always the direct target of the actions represented by the sentence.

- (ii) *John killed himself / *him.*

Rooryck and Vanden Wyngaerd (2007) restrict their generalization only to locative contexts, which I believe is the correct step to take.

not block the derivation in which the pronominal surfaces.

An alternative solution might be developed by building on the proposal by Rooryck and Vanden Wyngaerd (2007) (henceforth R&VW). R&VW's account of binding patterns in PP-environments is tightly linked to the locative nature of the binding contexts in question. Following Svenonius (2006), R&VW assume that prepositions project an AxPartP when used in their locative sense. Semantically, AxPart identifies a region along a particular axis of the object, such as the up-down, or the front-back axis. The consequence of this assumption is that there is a structural difference between locative contexts and other configurations including non-locative PPs, namely only locative PPs contain AxPartP. The choice between pronominals and anaphors in locative PPs then derives from the following assumptions regarding their inherent featural composition:

- (95) a. pronominals lack grammatical axial dimensions
 b. *self* contributes grammatical axial dimensions to the pronominal form it attaches to

This amounts to saying that the complex *self* anaphor has intrinsic AxParts, while the pronominal *him* lacks intrinsic AxParts. Thus, the *self*-form is semantically rich enough to provide the semantics of axial dimensions in locative contexts.

When a reflexive is used in a locative PP, an Agree-relationship is established between AxPart and the axial dimension of the *self*-form.

- (96) *John* put the book behind *himself*.
 [Evid **Speaker**_{1P.SG} John put the book [PP behind [AxPart \emptyset {**front-back**} [D himself {**front-back**}]]]]

According to Rooryck and Vanden Wyngaerd (2007), this gives rise to the intrinsic frame of reference. The axial dimensions provided by *self* also account for a strictly locative interpretation, involving typically direct, physical contact.

When pronominals are used axiality cannot be determined inside the PP projection, given that pronominals lack axial features by assumption. In this case, R&VW argue that the Speaker, grammatically represented in the left periphery of the clause, anchors AxParts via variable binding.

- (97) *John* saw a snake near *him*.
 [Evid **Speaker**_{1P.SG} John put the book [PP behind [AxPart \emptyset **Speaker** [D him]]]]

Binding of AxPart by the Speaker yields the extrinsic or speaker-oriented frame of reference. Since pronominals lack AxParts and therefore spatial dimensions, configurations involving pronominals allow a more abstract interpretation of locative relations.

According to R&VW, the proposed derivations also have important ramifications for the calculation of the local binding domain. They are however not very explicit regarding the version of the binding theory they endorse. With regards to (96) they

only note that “the domain for the *self*-form simply is the minimal domain with an accessible subject, i.e. the IP that has *John* as a subject.” (Rooryck and Vanden Wyngaerd (2007), p. 46). In (97), they argue that the Speaker-bound AxPart variable can be viewed as a second occurrence of Speaker within the same sentence. “The Speaker *a fortiori* creates an opaque domain for the application of Condition B of the binding theory.” (Rooryck and Vanden Wyngaerd (2007), p. 47).

Embedding this proposal within the model of binding pursued here, suppose then that the presence of the Speaker in AxPartP creates an intervention effect in a manner similar to the one discussed in the previous section, blocking the valuation of ϕ -features. The only alternative hence would be to select [*i*In] and valued ϕ for the numeration, and avoid the necessity of establishing the relevant Agree relation. Such a derivation would result in a configuration featuring a pronominal. In (96), where an Agree relation is established between AxPart and the axial features of the *self*-form, no intervention effects would arise. Recall that on our assumptions, an anaphor embedded in a PP occupies the same phasal domain as the clausal subject, when the latter is in its base-generated position, SpecvP. All the conditions are therefore met for an Agree relation to be established in the course of the syntactic derivation. Consequently, a derivation starting with [*u*In] and unvalued ϕ would yield a licit output. At Spell-Out, the [*u*In] would be spelled out by the morpheme *self*. The choice between an anaphor and a pronominal could thus potentially be reduced to a minimality effect, though precise details of implementation remain to be worked out.

Summing up, I have sketched two potential ways of accounting for availability of both bound pronominals and anaphors in spatial PPs in English. Both proposals rely on the presence of interpretative differences between the configurations featuring anaphors and those that feature bound pronominals. Since, as I have argued, the choice between the derivation yielding a pronominal and the one yielding an anaphor in local nonreflexive environments hinges on economy considerations, one possibility is to argue that due to the presence of interpretative differences, the two derivations do not enter into competition. Recall that by assumption economy only considers derivations that lead to indistinguishable interface representations. An alternative is to rely more closely on a recent analysis of PP-binding configurations by Rooryck and Vanden Wyngaerd (2007), arguing that the syntactic environments in which anaphors and pronominals surface are in fact different. In cases where an anaphor appears, there would be no intervenor present. On the other hand, the derivation yielding a pronominal would involve an intervenor, namely the Speaker inside the PP, blocking thereby the syntactic encoding of the relevant binding relation.

8.3.3.3 Revisiting the typology of PPs

In light of the suggested analysis of binding into PPs, let us return for a moment to the typology of PPs that we took as our point of departure in the previous chapter. Recall that we started by distinguishing functional prepositions such as those in (98) that do not allow bound pronominals as their complements, from those in (99) that

allow both an anaphor and a pronominal.

- (98) a. *John* sent a letter to *himself*/**him*.
 b. *John* always talks about *himself*/**him*.
- (99) a. *John* pulled the blanket over *himself*/*him*.
 b. *John* looked around *himself*/*him*.

I then speculated that the relevant difference between these two types of Ps might be that the functional prepositions are simply the Spell-Out of case-layer(s), rather than instantiations of category P. However, the investigation of mechanisms responsible for encoding binding dependencies undertaken in this chapter, and in particular the distribution of anaphoric expressions in Dutch led us to the conclusion that the relevant difference between functional and spatial Ps responsible for the observed binding patterns is the fact that only the latter type of Ps are predicative in nature. Functional Ps do not form a distinct predicate and hence the complement of such Ps is an argument of the verb. In other words, contexts featuring functional Ps are in fact reflexive configurations. As a result, the [R]-feature will be present in cases like (98), forcing a SELF anaphor to appear. Employing just a pronominal in such contexts is illicit since the pronominal cannot lexicalize the [R]-feature. The suggestion from the preceding chapter that functional Ps might not be instantiations of category P but simply spell out layers in the nominal functional sequence is compatible with the view that these Ps are non-predicative. If functional Ps are simply case markers, then it would be unsurprising that their complements should in fact be arguments of the verb. However, even if the analysis of functional Ps along these lines proves plausible once developed in more detail, a task which I will leave for some other occasion, there might be reasons not to extend such an analysis to all functional Ps. In the preceding chapter, I have suggested that treating *to* as the Spell-Out of case layer(s), rather than as a member of the category P, might also explain why binding out of *to*-phrases is possible despite the apparent lack of c-command. Note however that *to*- and *about*-phrases differ when facts pertaining to the c-command puzzle are reviewed, although they pattern alike in allowing anaphors and disallowing bound pronominals as complements. Already Postal (1971) observed that while antecedents embedded in a *to*-phrase do seem to c-command out of that phrase for the purposes of binding, potential antecedents embedded in *about*-phrases do not, as shown in (100). Note that when no binding is involved, rearranging the order of *to*- and *about*-phrases does not produce an ungrammatical output, as (101) illustrates.

- (100) a. I talked to *Bill* about *himself*.
 b. *I talked about *Bill* to *himself*.
- (101) a. I talked to Bill about communism.
 b. I talked about communism to Bill.

We might interpret this pattern as indicating that *to*-phrases and *about*-phrases differ with respect to whether they instantiate the category P. If *to* spells out case layers, then in effect there would be no PP layer dominating *Bill* in (100a) and thus no problem with c-command would arise. In this case, *to* would be simply spelling out a projection within the nominal functional sequence. On the other hand, we might argue that the preposition *about* does head a PP phrase, in which case *Bill* would not c-command and therefore could not bind the reflexive. Reflexives would still be required though, on the assumption that the preposition *about* is non-predicative in nature.⁴²

This view would then lead to a three-way typology of Ps: *to*-type prepositions, *about*-type prepositions, and locative prepositions. A representative example of each class is given below.

- (102) *John* always talks to *himself*/**him*.
- (103) *John* always talks about *himself*/**him*.
- (104) *John* pulled the blanket over *himself*/*him*.

On this view, members of the first class would be part of the nominal functional structure, rather than instantiations of category P. This would explain why they differ from the latter two types in apparently not counting for purposes of calculating the relevant c-command domain of the DP that they seem to take as their complement. On the other hand, functional Ps like *about* would instantiate the category P, but would be non-predicative in nature, explaining why they do not allow bound pronominals as their complements. Finally, spatial prepositions form their own predicate. Consequently, bound pronominals as complements of spatial Ps are not categorically ruled out and might surface under certain conditions, as discussed in this section. Many issues arise and the details of this type of approach need to be carefully attended to. I won't pursue such an analysis further at this point, and will leave a more detailed investigation of functional Ps for some future occasion.

8.4 Summary and some consequences

Let me now summarize the proposed analysis and the main conclusions we have reached in the preceding sections. I have pursued an Agree-based account of anaphor binding, arguing that anaphors enter into a feature dependency with their antecedents in the course of the syntactic derivation. I have assumed that the relevant features involved are ϕ -features and the feature [In], which is mapped onto an integer value at the interface. The *self*-morpheme surfaces in a derivation that starts with [*u*In] and unvalued ϕ . An Agree-relation is then established with another c-commanding

⁴²Alternative ways of analyzing the pattern in (100) have been suggested in the literature. See for instance Pesetsky (1995).

noun phrase in order to value the unvalued set of ϕ -features. This must be achieved before the phase is complete, i.e. before the relevant part of the syntactic structure is sent to the interfaces for interpretation, which explains why anaphors need local antecedents. The translation procedure mapping the syntactic to semantic representation ensures that the phrase bearing [*uIn*] receives the same index value as the phrase it enters into an Agree-relation with. Consequently, a bound variable dependency will be established at the interface, assuming that coindexation always results in semantic binding. When the derivation starts with [*iIn*], the *self*-morpheme does not surface. A pronominal then lexicalizes both the ϕ -features and the [*iIn*]. In such cases, no syntactic dependency needs to be established and no locality restrictions are therefore observed. The configuration in which a pronominal surfaces might give rise to a bound variable dependency at the interface, but it need not to, since the pronominal does not necessarily end up coindexed with its antecedent. Hence, the relation between a pronominal and its antecedent may also be one of coreference.

Investigation of anaphoric dependencies in Dutch, a language which makes a three-way distinction in the anaphoric system, suggests that not all local binding configurations are alike. I have argued that in reflexive configurations, i.e. contexts where a binding relation is established between two arguments of the same predicate, the presence of what I have labelled [R]-feature is required in order to circumvent a violation at the interface. In such configurations therefore, a lexical item must be used which can spell out this feature. In languages under discussion, this can be achieved by using a SELF anaphor. Pronominals on the other hand, are strictly ruled out in reflexive configurations, since they cannot lexicalize the [R]-feature.

In non-reflexive local binding configurations, the distribution of anaphors and bound pronominals is regulated by an economy principle which favours the establishment of dependencies in the course of the syntactic derivation. As noted, a syntactic dependency always exists between an anaphor and its antecedent, which then results in an interpretative dependency at the interface. When a pronominal is used, the same interpretative dependency can be obtained at the interface, without there being any syntactic dependency between the pronominal and its antecedent. Consequently, anaphors are preferred over bound pronominals for reasons of economy. This is however only true of local binding configurations. When the anaphor's unvalued features cannot be valued within the local domain, i.e. the phase, the only derivation that converges is the one which yields a pronominal. A consequence of this view is that bound pronominals in non-reflexive local binding configurations are not ungrammatical per se, but generally do not surface due to economy considerations. This is a welcome result given that (i) as Reuland (2001) points out, only if such a derivation converges can it block the coreferential interpretation in the relevant contexts, and (ii) pronominals do sometimes surface in local binding configurations, as is the case in certain PP-binding contexts. We have then briefly considered two potential ways of accounting for the availability of both anaphors and bound pronominals in environments involving spatial PPs. Both suggested solutions build on the existence of certain interpretative differences correlating with the use of an anaphor or a pronominal.

inal.

In the previous chapter, I have argued that binding can take place either in the course of the syntactic derivation, at the semantic interface, or at the level of discourse. Now that we have discussed how binding relations might be encoded, these claims can be further clarified. Recall that I have argued that, semantically, binding in syntax and binding at the semantic interface behave alike, i.e. they both involve variable binding. However, only the former type of binding relations must be established within the local domain, i.e. the phase. As we have seen in this chapter, this locality restriction is due to the requirement that a syntactic feature dependency be established. Thus, ‘binding in syntax’ involves a formal dependency which translates into a bound variable representation at the interface. What I have previously called ‘binding at the semantic interface’ would then involve the same interpretative dependency, without the corresponding syntactic dependency. Discourse ‘binding’ is crucially different since it employs a distinct strategy for establishing overlapping reference, namely *coreference*. Coreference is subject to the principles of the discourse component and not under the purview of grammar proper.

This chapter has aimed to provide an answer as to why syntactic locality restrictions should play a role in constraining binding relations, relations which are essentially interpretative in nature. I have argued that sensitivity of anaphor binding to syntactic factors is due to the fact that an anaphor enters into an Agree-relation with its antecedent in the course of the syntactic derivation. On this view, structural restrictions on binding of anaphors thus reduce to structural conditions on the Agree operation. We therefore predict that anaphors and their antecedents should occupy the same phase, given that Agree cannot operate across phasal boundaries. In other words, we explain the relevance of phases as locality domains on anaphoric binding. We also correctly predict that anaphors and their antecedents should be in a c-command relation, given that c-command is a structural requirement on the Agree-operation.

On the view pursued here, binding of pronominals, unlike binding of anaphors, is not encoded in syntax. Consequently, we capture the fact that the relation between a bound pronominal and its antecedent can span phasal boundaries. Note however that bound pronominals, which I have argued are related to their antecedents only via an interpretative dependency, must also be c-commanded by their antecedents. In this case, the c-command requirement cannot be reduced to a condition on syntactic relations since no syntactic dependency is established between a bound pronominal and its antecedent. We are therefore led to the conclusion that there is an independent c-command requirement on variable binding. However, it has been argued in the literature that the c-command requirement on variable binding need not be stipulated, but can be derived from general properties of the interpretative process. I refer the reader to Reinhart (1983b, 1995) and Reuland (1998) for relevant discussion. Thus, on the view pursued here, anaphors end up being subject to a double c-command requirement: the c-command requirement which is the consequence of the established Agree-relation, and the c-command requirement as a consequence of the bound vari-

able relation at the interface. That this might be a desirable result, rather than an inconvenient redundancy, has been argued by Reuland (1998) and Hicks (2009). As noted already by Reinhart (1976, 1983b), there are some well-known cases of variable binding which seem not to involve surface c-command. For instance, for many speakers, a bound pronominal can take as its antecedent a possessor embedded in a DP ((105a) is from Reinhart (1983b), (105b) is from Reuland (1998)).

- (105) a. Everyone_{*i*}'s mother kissed him_{*i*}.
 b. Every girl_{*i*}'s father admires her_{*i*}.

Since the antecedents in (105) are quantified DPs, we know that the relation between the pronominals and their antecedents in these cases must be one of variable binding, not coreference. Such examples thus seem to suggest that for reasons that are not yet very well understood, the c-command requirement on variable binding can be relaxed in some cases.⁴³

However, replacing the bound pronominals with anaphors in contexts like (105) yields a deviant result. In other words, a possessor embedded in a DP cannot apparently serve as an antecedent for an anaphor.

- (106) *Every girl_{*i*}'s father admires herself_{*i*}. (from Reuland (1998))

It appears therefore that the c-command requirement is less strict on pronominal binding than on anaphor binding. Since both anaphors and pronominals can function as bound variables, it is unlikely that the contrast is due to restrictions on variable binding. However, if anaphors, unlike bound pronominals, must also be locally c-commanded by their antecedents in narrow syntax, then this discrepancy in the behaviour of anaphors and bound pronominals becomes less mysterious.

⁴³See Reuland (1998) for a proposal on how this might be captured.

Chapter 9

Conclusion

The intuition that common syntactic factors constrain both movement operations and the relation between anaphoric expressions and their antecedents is an old one and has been pursued in many accounts of movement/binding phenomena. As I have noted, in Chomsky's highly influential work, *Lectures on government and binding*, for instance, ordering restrictions on movement operations followed from binding-theoretic properties of traces. On this view, the same principles, namely the binding conditions, were responsible for deriving both the improper movement effects and the distribution of anaphoric expressions. This dissertation offers a new angle on the link between these two empirical domains. The main focus is no longer on the Binding theory as the unifying factor, but rather on the nature of syntactic derivation and featural composition of the elements involved.

The first part of the dissertation investigated improper movement phenomena. The empirical foundation for subsequent discussion was laid out in chapter 2. Whereas the original formulation of the relevant generalization regarding possible feeding/-bleeding relations between movement operations encompasses only cases involving consecutive movements of the same phrase, it was shown that the phenomenon is much broader in scope. Namely, the same interactions between movement operations can also be observed in cases involving remnant movement and subextraction from moved phrases, as argued in several recent analyses of the phenomena. Different ways of capturing these ordering restrictions suggested in the literature were reviewed in chapter 3. Although the noted shortcomings of these proposals prompted us to search for a different solution, they nevertheless provided important insights that shaped the analysis developed in the following chapters.

Chapters 4 and 5 investigated how the restrictions on ordering of movement operations might be derived from the independently needed hierarchy of functional projections. A first attempt in this direction was outlined in chapter 4. Drawing on the work of Sternefeld (1993), Williams (2003), Abels (2007), I formulated a constraint on movement requiring the landing site of any movement operation to be higher than its launching site in terms of the functional hierarchy. I have referred to this constraint as the Generalized Ban on Improper Movement (GBOIM). It has then been shown

that an analysis relying solely on the GBOIM fails to capture a broad range of relevant cases, but that the restrictive power of such an analysis increases considerably once GBOIM is combined with a phase-based theory of locality. It is important to note that such a move does not amount to adding new technical machinery, given that phases play a significant role in accounting for other properties of language as well. Among these are distributional properties of anaphoric expressions, as I have argued in chapters 7 and 8. I have also shown that despite initial appearances, the GBOIM is compatible with the analysis of ECM-constructions in terms of raising-to-object, under the assumption that successive-cyclic A-movement proceeds via the Specs of v/VPs, rather than Specs of TPs. We have seen that most, if not all the evidence for intermediate stop-over in Spec of TP can be reanalyzed as involving the Specs of v/V instead. While the analysis of ECM-constructions in terms of raising-to-object can be reconciled with the GBOIM, some other aspects of the proposal remain problematic. In particular, without additional stipulations, no restrictions on ordering of movement operations are predicted in cases involving subextraction out of moved phrases. While in this respect the proposal seems to be too permissive, closer consideration reveals that a theory adopting both the GBOIM and phase-based locality is in fact overly restrictive since it rules out all cases of long movement out of the CP.

In light of these concerns, in chapter 5 I have outlined an alternative, though related analysis of improper movement. In particular, drawing on a proposal by Svenonius (2004) and assuming a close correspondence between the nominal and clausal functional hierarchy, I have argued that restrictions on ordering of movement operations can be derived by relying on the internal featural composition of the moving phrase, again in conjunction with a phase-based locality. To this end, I have suggested that the standard view regarding the nature of Spell-Out be modified so that upon the merger of H, H a phasal head, all features/projections lower than H on the functional hierarchy spell out. A consequence of this view is that the phrase, or part of the phrase, sitting in the specifier of the phasal head might spell out, not just the complement of the phasal head, as standardly assumed. It was then shown that, with these assumptions in place, a broad range of violations which might be viewed as falling under the rubric of improper movement can be captured. This alternative account not only avoids the problems that plagued the GBOIM-based approach, but it also does not require that an independent constraint on movement, such as the GBOIM, be stated. The ordering restrictions fall out from more general constraints on syntactic derivations and particular assumptions regarding the featural make-up of the elements involved.

The same assumptions about the nature of syntactic computation have then been carried over to the realm of anaphoric binding. In chapter 7, I have explored the hypothesis that the phase constitutes the relevant locality domain for binding relations. Though many details remain to be clarified, I hope to have shown that pursuing this line of enquiry can take us a long way in accounting for empirical facts. It has also been shown that the particular modifications to the assumptions regarding the nature of Spell-Out proposed in the previous chapter and motivated by the analysis of im-

proper movement facts lead to some novel predictions in the domain of anaphoric binding, which I have argued are empirically supported. I have also argued that assuming the phase to be the relevant binding domain makes the correct split between local and nonlocal binding relations. I went on to show that there are grounds for distinguishing two types of nonlocal binding relations, thereby further refining the typology of referential dependencies. Finally, considerable attention was dedicated to the task of capturing the ban on A'-binders within the current model of binding.

The goal of the final chapter was to explain the relevance of phases for anaphoric binding and to further clarify the nature of binding relations. I have pursued the view that a binding relation between an anaphor and its antecedent involves the establishment of an Agree-relation in the course of the syntactic derivation. Locality restrictions on anaphoric binding therefore reduce to the locality restrictions on the operation Agree, which is constrained by the PIC. On the analysis developed here, features that are involved in encoding binding relations were identified as ϕ -features and the [In]-feature, the latter feature being interpreted at the interface as an instruction to assign an index to the DP bearing it. The relevant syntactic dependency for encoding binding relations is ϕ -feature agreement, which must be established, as I have argued, in cases where the derivation starts with an uninterpretable [In]-feature. In English, the [*u*In] is then lexicalized by the morpheme *self*, which I assume bears this feature as part of its featural specification. Given certain assumptions regarding how the mapping from syntactic to semantic structure proceeds, the analysis ensures that the established syntactic dependency translates into a bound variable dependency at the interface. The availability of a bound variable interpretation on the current proposal is not however dependent on the presence of a syntactic feature dependency. Consequently, pronominals, which I have argued bear an interpretable [In]-feature and do not enter into an Agree relation with their antecedents, can nevertheless be interpreted as bound variables.

Investigation of Dutch, a language which unlike English distinguishes morphologically two kinds of anaphors, prompted us to further refine the featural composition of anaphoric expressions. Two types of local binding relations were identified, those where the anaphor and its antecedent are coarguments of the same predicate (labelled *reflexive* configurations), and those where they are not, but are still within the same local domain, namely the phase. The obligatory presence of SELF anaphors, or more precisely, of the SELF morpheme, in the former type of local binding configurations was tied to the obligatory presence of a particular feature in the syntactic structure. Differences between English and Dutch regarding when a SELF anaphor is required to surface stem, on the current view, from different featural specifications of the SELF morphemes in these two languages.

The final part of this chapter explores the consequences of the proposed analysis for the distribution of pronominals. As we have seen, the (un)availability of anaphors and pronominals in particular syntactic environments follows in most cases directly from their featural specification. An exception in this respect is local non-reflexive binding configurations, where without additional assumptions, the proposed analysis

predicts that both anaphors and pronominals should be licit. I have suggested that the distribution of anaphoric expressions in this type of syntactic environment might be captured by relying on economy considerations, leaving however a more careful and detailed analysis as a task for future research.

The goal of this dissertation was to offer a way of relating locality restrictions on movement and binding relations to a common set of syntactic factors. Since the empirical domains involved are extremely broad, many intriguing questions had to be ignored or have received only scant attention. Many aspects of the analysis also remain to be clarified. Though I have to leave these tasks for future work, I hope to have shown that the current proposal opens some promising avenues for deeper investigation of these phenomena.

Bibliography

- Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Doctoral Dissertation, University of Connecticut, Storrs.
- Abels, Klaus. 2007. Some implications of improper movement for Cartography. Ms., University of Tromsø and University College London.
- Abels, Klaus. 2008a. Templates vs. the grabbag theory of the left periphery. Ms., University of Tromsø and University College London.
- Abels, Klaus. 2008b. Towards a restrictive theory of (remnant) movement. *Linguistic Variation Yearbook* 7:53–120.
- Abels, Klaus, and Peter K. Muriungi. 2008. The focus particle in Kĩĩtharaka: Syntax and semantics. *Lingua* 118:687–731.
- Abney, Steven. 1987. The English Noun Phrase in its Sentential Aspect. Doctoral Dissertation, MIT.
- Adger, David, and Gillian Ramchand. 2005. Merge vs Move: wh-dependencies revisited. *Linguistic Inquiry* 36:161–193.
- Bach, Emmon, and Barbara Partee. 1980. Anaphora and semantic structure. In *Papers from the parasession on pronouns and anaphora*, 1–29. University of Chicago, Chicago: Chicago Linguistic Society.
- Baker, Mark. 1988. Theta theory and the syntax of applicatives in chichewa. *Natural Language and Linguistic Theory* 6:353–389.
- Baker, Mark C. 2008. *The syntax of agreement and concord*. Cambridge Studies in Linguistics. Cambridge: Cambridge University Press.
- Baltin, Mark R. 1995. Floating Quantifiers, pro, and Predication. *Linguistic Inquiry* 26:199–248.
- Barss, Andrew. 1986. Chains and Anaphoric Dependence. Doctoral Dissertation, MIT.

- Barss, Andrew, and Howard Lasnik. 1986. A Note on Anaphora and Double Objects. *Linguistic Inquiry* 17:347–354.
- Beck, Sigrid. 1996. Quantified structures as barriers for LF movement. *Natural Language Semantics* 4:1–56.
- Belletti, Adriana, and Luigi Rizzi. 1988. Psych-verbs and θ -theory. *Natural Language and Linguistic Theory* 6:291–352.
- Benedicto, Elena. 1991. Latin Long-distance Anaphora. In *Long-distance Anaphora*, ed. Jan Koster and Eric Reuland, 171–184. Cambridge, Massachusetts: Cambridge University Press.
- Bobaljik, Jonathan. 1993. Floating Quantifiers: Handle with Care. In *The Second Glot International State-of-the-Article Book*, ed. Lisa Cheng and Rint Sybesma. Berlin: Mouton de Gruyter.
- Bobaljik, Jonathan, and Susi Wurmbrand. 1999. Modals, raising and A-reconstruction. A talk given at Leiden University and University of Salzburg.
- Boeckx, Cedric. 1999. Conflicting c-command requirements. *Studia Linguistica* 53:227–250.
- Boeckx, Cedric. 2001. Scope Reconstruction and A-movement. *Natural Language and Linguistic Theory* 19:503–548.
- Bouchard, Denis. 1982. On the content of empty categories. Doctoral Dissertation, MIT.
- Bouchard, Denis. 1985. The Binding Theory and the notion of accessible SUBJECT. *Linguistic Inquiry* 16:117–133.
- Bošković, Željko. 1997. *The syntax of nonfinite complementation: An economy approach*. Cambridge, Massachusetts: MIT Press.
- Bošković, Željko. 2002. A-movement and the EPP. *Syntax* 5:167–218.
- Bošković, Željko. 2004. Be Careful Where You Float Your Quantifiers. *Natural Language and Linguistic Theory* 22:681–742.
- Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59:1–45.
- Bošković, Željko. 2007. On the Locality and Motivation of Move and Agree: An Even More Minimal Theory. *Natural Language and Linguistic Theory* 38:589–644.

- Broekhuis, Hans. 2005. Extraction from subjects: Some remarks on Chomsky's 'On Phases'. In *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk studies in honor of henk van riemsdijk*, ed. Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinherz, and Jan Koster. Berlin: Mouton de Gruyter.
- Bruening, Benjamin. 2001. Raising to object and proper movement. Ms., University of Delaware.
- Büring, Daniel. 2005a. *Binding Theory*. Cambridge Textbooks in Linguistics. Cambridge: CUP.
- Büring, Daniel. 2005b. Bound to bind. *Linguistic Inquiry* 36:259–274.
- Burzio, Luigi. 1991. The Morphological Basis of Anaphora. *Journal of Linguistics* 27:81–105.
- Caha, Pavel. 2009. The nanosyntax of case. Doctoral Dissertation, University of Tromsø.
- Cantrall, William. 1974. *View Point, Reflexives, and the Nature of Noun Phrases*. The Hague: Mouton.
- Carstens, Vicki. 2010. Implications of grammatical gender for the theory of uninterpretable features. In *Exploring crash-proof grammars*, ed. Michael T. Putnam, 31–59. Amsterdam/Philadelphia: John Benjamins.
- Chomsky, Noam. 1970. Remarks on Nominalizations. In *Readings in English Transformational Grammar*, ed. Roderic Jacobs and Peter Rosenbaum, 184–221. Waltham, MA: Blaisdell.
- Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, ed. Stephen R. Anderson and Paul Kiparsky, 232–286. New York: Holt, Rinehart and Winston.
- 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, Massachusetts: MIT Press.
- Chomsky, Noam. 1986a. *Barriers*. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 1986b. *Knowledge of Language: Its Nature, Origin and Use*. New York: Preager.
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel J. Keyser, 1–52. Cambridge MA: The MIT Press.

- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by Step: Essays in Minimalist Syntax in Honour of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge MA: The MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 133–166. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 2008. On phases. In *Foundational Issues in Linguistic Theory: Essays in honor of Jean-Roger Vergnaud*, ed. Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 133–166. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam, and Howard Lasnik. 1993. The theory of principles and parameters. In *Syntax: An International Handbook of Contemporary Research*, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 506–569. Berlin: Walter de Gruyter.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective*. New York: OUP.
- Citko, Barbara. 2006. The interaction between across-the-board *wh*-movement and left-branch extraction. *Syntax* 9:225–247.
- Clements, George. 1975. The logophoric pronoun in Ewe: Its role in discourse. *Journal of West African Languages* 10:141–177.
- Cole, Peter, and Gabriella Hermon. 1998. Long-distance reflexives in Singapore Malay: An apparent typological anomaly. *Linguistic Typology* 2:57–79.
- Collins, Chris. 2005. A smuggling approach to raising in English. *Linguistic Inquiry* 36:289–298.
- Collins, Chris, and Höskuldur Thráinsson. 1996. VP-internal structure and object shift in Icelandic. *Linguistic Inquiry* 27:391–444.
- van Craenenbroeck, Jeroen, and Marjo van Koppen. 2007. Feature inheritance and multiple phase boundaries. Handout from a talk given at GLOW, University of Tromsø, April 2007.
- Cresswell, M.J. 1973. *Logics and Language*. London: Methuen & Co.
- Culicover, Peter W. 1992. Topicalization, inversion and complementizers in English. In *Going Romance and beyond*, ed. D et al Delfitto. Utrecht: University of Utrecht.
- Culicover, Peter W. 1996. On distinguishing A'-movements. *Linguistic Inquiry* 27:445–463.

- Déchaine, Rose-Marie, Teun Hoekstra, and Johan Rooryck. 1995. Augmented and non-augmented HAVE. In *Proceedings of Langues et Grammaire I*, 85–101.
- Dékány, Éva. 2009. The nanosyntax of Hungarian postpositions. In *Tromsø working papers on language and linguistics: Nordlyd 36.1*, Special issue on Nanosyntax, ed. Peter Svenonius, Gillian Ramchand, Michal Starke, and Knut Tarald Taraldsen, 41–75. Tromsø: University of Tromsø. Available at <http://www.ub.uit.no/baser/nordlyd/>.
- Doetjes, Jenny. 1997. Quantifiers and Selection: On the Distribution of Quantifying Expressions in French, Dutch and English. Doctoral Dissertation, University of Leiden, Leiden.
- Dowty, David R., and Belinda Brodie. 1984. The semantics of “floated” quantifiers in a transformationless grammar. In *Proceedings of the 3rd West Coast Conference on Formal Linguistics*, ed. Mark Cobler, Susannah MacKaye, and Michael T. Westcoat, 75–90. Stanford, CA: Stanford Linguistics Association.
- Epstein, Samuel D., Hisatsugu Kitahara, and T. Daniel Seely. 2010. Uninterpretable features: What are they and what do they do? In *Exploring crash-proof grammars*, ed. Michael T. Putnam, 125–143. Amsterdam/Philadelphia: John Benjamins.
- Epstein, Samuel David, and T. Daniel Seely. 2002. Rule applications as cycles in a level-free syntax. In *Derivation and Explanation in the Minimalist Program*, ed. Samuel David Epstein and T. Daniel Seely, 65–90. Blackwell Publishing.
- Epstein, Samuel David, and T. Daniel Seely. 2006. *Derivations in Minimalism*. Cambridge Studies in Linguistics. Cambridge University Press.
- Everaert, Martin. 1986. *The Syntax of Reflexivization*. Dordrecht: Foris.
- Everaert, Martin. 1991a. Contextual determination of the anaphor/pronominal distinction. In *Long-distance Anaphora*, ed. Jan Koster and Eric Reuland, 77–118. Berlin: Mouton de Gruyter.
- Everaert, Martin. 1991b. Nominative anaphors in Icelandic: morphology or syntax? In *Issues in Germanic syntax*, ed. Werner Abraham, Wim Kosmeijer, and Eric Reuland, 277–306. Cambridge: Mouton de Gruyter.
- Falk, Cecilia. 1990. On double object constructions. *Working papers in Scandinavian syntax* 43:53–100.
- Faltz, Leonard. 1985. *Reflexivization: A Study in Universal Syntax*. New York: Garland.

- Fanselow, Gisbert. 1990. Scrambling as NP-movement. In *Scrambling and barriers*, ed. Günther Grewendorf and Wolfgang Sternefeld, 113–140. Amsterdam: John Benjamins.
- Fanselow, Gisbert. 2001. Features, theta-roles, and free constituent order. *Linguistic Inquiry* 32:405–437.
- Fiengo, Robert. 1977. On trace theory. *Linguistic Inquiry* 8:35–62.
- Folli, Raffaella, and Heidi Harley. 2002. Flavours of v: Consuming Results in Italian and English. In *Aspectual enquiries*, ed. Paula Kempchinsky and Roumyana Slabakova. Dordrecht: Springer.
- Fox, Danny. 1999. Reconstruction, binding theory and the interpretation of chains. *Linguistic Inquiry* 30:157–196.
- Fox, Danny. 2000. *Economy and Semantic Interpretation*. Cambridge, MA: MIT Press.
- Gill, Kook-Hee. 1999. The Long-Distance Anaphora Conspiracy: The Case of Korean. *U. Penn Working Papers in Linguistics* 6.
- Giorgi, Alessandra. 1984. Towards a theory of long-distance anaphors: a GB approach. *The Linguistic Review* 3:307–359.
- Grewendorf, Günther. 1989. *Ergativity in German*. Dordrecht: Foris.
- Grewendorf, Günther. 1994. Variable Binding and Reconstruction. Ms., Universität Frankfurt/Main.
- Grewendorf, Günther. 2003. Improper remnant movement. *Gengo Kenkyu- Journal of the Linguistic Society of Japan* 123:47–94.
- Grewendorf, Günther. 2005. The asymmetry of short and long *wh*-extraction in German. *Recherches linguistiques de Vincennes* 33:35–54. Available at <http://rlv.revues.org/1253>.
- Grewendorf, Günther, and Joachim Sabel. 1999. Scrambling in German and Japanese: Adjunction versus multiple specifiers. *Natural Language and Linguistic Theory* 17:1–65.
- Grimshaw, Jane. 1993. Minimal Projection, Heads and Optimality. Ms., Rutgers University.
- Grohmann, Kleanthes K. 2000. Prolific Peripheries: A Radical View from the Left. Doctoral Dissertation, University of Maryland, College Park.

- Grohmann, Kleanthes K., John Drury, and Juan Carlos Castillo. 2000. No more EPP. In *Proceedings of the 19 West Coast Conference on Formal Linguistics*, ed. Roger Billerey and Brook Danielle Lillehaugen, 153–166. Somerville, MA: Cascadilla Proceedings Project.
- Hale, Ken, and Samuel Jay Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. Linguistic Inquiry Monograph Thirty-Nine. Cambridge, Massachusetts: MIT Press.
- Halle, Morris. 1997. Distributed morphology: Impoverishment and fission. In *MIT Working Papers in Linguistics*, ed. Yoonjung Kang Benjamin Bruening and Martha McGinnis, volume 30, 425–449. MIT Press.
- Halle, Morris, and Alec Marantz. 1993. Distributed Morphology. In *The view from building 20.: Essays in honor of sylvain bromberger*, ed. Kenneth Hale and Samuel J. Keyser, 111–176. Cambridge/London: MIT Press.
- Harley, Heidi. 1995. Subjects, Events and Licensing. Doctoral Dissertation, MIT.
- Harley, Heidi. 2003. Possession and the Double Object Construction. In *Linguistic Variation Yearbook 2*, ed. Pierre Pica and Johan Rooryck, 31–70. Amsterdam: John Benjamins.
- Heim, Irene. 1998. Anaphora and Semantic Interpretation: A Reinterpretation of Reinhart's Approach. In *The Interpretative Tract*, ed. Uli Sauerland and Orin Percus, MIT Working Papers in Linguistics 25, 205–246. Cambridge, Mass.: MIT Press.
- Heim, Irene, and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Blackwell textbooks in linguistics. Oxford: Blackwell.
- Hestvik, Arild. 1991. Subjectless binding domains. *Natural Language and Linguistic Theory* 9:455–496.
- Hicks, Glyn. 2005. Deriving Condition A. *York Papers in Linguistics* 3:91–131.
- Hicks, Glyn. 2009. *The Derivation of Anaphoric Relations*. Amsterdam/Philadelphia: John Benjamins.
- Hinterhölzl, Roland. 2006. *Scrambling, remnant movement, and restructuring in West Germanic*. Oxford: UOP.
- Hiraiwa, Ken. 2002. Movement and derivation: Eliminating the PBC. In *UPenn Working Papers in Linguistics*, ed. Elsi Kaiser and Sudha Arunachalam, volume 9. Philadelphia, PA: University of Pennsylvania.

- Holmberg, Anders, and Christer Platzack. 1995. *The role of inflection in Scandinavian syntax*. New York: Oxford University Press.
- Hornstein, Norbert. 2001. *Move! A minimalist theory of construal*. Cambridge, USA: Blackwell.
- Hornstein, Norbert, Howard Lasnik, and Juan Uriagereka. 2007. The dynamics of islands: Speculations on the locality of movement. *Linguistic analysis* 33:149–175.
- Hornstein, Norbert, and Andrew Nevins. 2005. The review of Representation Theory by Edwin Williams. *Language* 81:757–761.
- Hornstein, Norbert, Jairo Nunes, and Kleanthes K. Grohmann. 2005. *Understanding Minimalism*. Cambridge Textbooks in Linguistics. Cambridge, New York: Cambridge University Press.
- Huang, C.T.James. 1983. A Note on the Binding Theory. *Linguistic Inquiry* 14:554–561.
- Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, Mass.: MIT Press.
- Jackendoff, Ray. 1990. On Larson's treatment of the double object construction. *Linguistic Inquiry* 21:427–456.
- Jackendoff, Ray. 1992. Mme. Tussaud meets the binding theory. *Natural language and linguistic theory* 10:1–31.
- Jaeggli, O. 1980. Remarks on *to* contraction. *Linguistic Inquiry* 11:239–246.
- Johnson, Kyle. 1991. Object positions. *Natural Language and Linguistic Theory* 9:577–636.
- Johnson, Kyle. 2009. Why movement? A talk presented at Chicago Linguistic Society meeting, April 2009.
- Kayne, Richard. 1975. *French Syntax*. Cambridge, Massachusetts: MIT Press.
- Kayne, Richard. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.
- Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Massachusetts: MIT Press.
- Kayne, Richard S. 2002. Pronouns and their antecedents. In *Derivation and Explanation in the Minimalist Program*, ed. Samuel D. Epstein and T. Daniel Seely, 133–166. Oxford: Blackwell.

- Kiss, Katalin É. 1993. Wh-movement and specificity. *Natural Language and Linguistic Theory* 11:85–120.
- Kiss, Tibor. 2001. Anaphora and Exemptness: A comparative treatment of anaphoric binding in German and English. In *Proceedings of the 7th International HPSG Conference*, ed. Dan Flickinger and Andreas Kathol. Stanford: CSLI Publications.
- Koopman, Hilda, and Dominique Sportiche. 1982. Variables and the bijection principle. *The Linguistic Review* 2:139–160.
- Koster, Jan. 1987. *Domains and Dynasties: the Radical Autonomy of Syntax*. Dordrecht: Foris.
- Koster, Jan. 1997. Anaphora and the uniformity of grammar. In *Perspectives on Binding and Atomism*, ed. Hans Bennis, Pierre Pica, and Johan Rooryck, 235–250. Dordrecht: Foris.
- Kuno, Susumo. 1987. *Functional Syntax: Anaphora, Discourse, and Empathy*. Chicago: Chicago University Press.
- Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19:335–391.
- Larson, Richard K. 1990. Double objects revisited: reply to Jackendoff. *Linguistic Inquiry* 21:589–633.
- Lasnik, Howard. 1995. A note on pseudogapping. In *Papers on minimalist syntax*, MIT Working Papers in Linguistics 27, 143–163. Cambridge, Mass.: MIT Press.
- Lasnik, Howard. 1999. Chains of arguments. In *Working Minimalism*, ed. Samuel Epstein and Norbert Hornstein, 189–215. Cambridge, Mass.: MIT Press.
- Lasnik, Howard, and Mamoru Saito. 1991. On the subject of infinitives. In *Papers from the 27th Regional Meeting of the Chicago Linguistic Society Part One: The General Session*, ed. Lisa M. Dobrin, Lynn Nichols, and Rosa M. Rodriguez. Chicago, IL: Chicago Linguistic Society, University of Chicago.
- Lasnik, Howard, and Mamoru Saito. 1992. *Move α : Conditions on its application and output*. Cambridge, MA: MIT Press.
- Lebeaux, David. 1983. A Distributional Difference between Reciprocals and Reflexives. *Linguistic Inquiry* 14:723–730.
- Lebeaux, David. 1984. Locality and Anaphoric Binding. *Linguistic Review* 4:343–363.

- Lebeaux, David. 1991. Relative clauses, licensing, and the nature of the derivation. In *Syntax and Semantics 25: Perspectives on phrase structure*, ed. Susan Rothstein. New York: Academic Press.
- Lebeaux, David. 2009. *Where Does Binding Theory Apply?*. Linguistic Inquiry Monograph 50. Cambridge, Massachusetts: MIT Press.
- Lees, Robert, and Edward S. Klima. 1963. Rules for English pronominalization. *Language* 13:17–28.
- Legate, Julie Ann. 2003. Some interface properties of the phase. *Linguistic Inquiry* 34:506–516.
- Lekakou, Marika. 2005. Reflexives in contexts of reduced valency: German vs. Dutch. In *The function of function words and functional categories*, ed. Marcel den Dikken and Christina Tortora. Amsterdam: John Benjamins.
- Mahajan, Anoop. 1990. The A-/A-bar distinction and movement theory. Doctoral Dissertation, MIT.
- Marantz, Alec. 1984. *On the Nature of Grammatical Relations*. Cambridge, Mass.: MIT Press.
- May, Robert. 1977. The Grammar of Quantification. Doctoral Dissertation, MIT.
- May, Robert. 1979. Must COMP-to-COMP movement be stipulated? *Linguistic Inquiry* 10:719–725.
- McCloskey, James. 2000. Quantifier float and wh-movement in an Irish English. *Linguistic Inquiry* 31:57–84.
- Merchant, Jason. 1996. Object scrambling and Quantifier Float in German. In *Proceedings of NELS 26*, ed. Kiyomi Kusumoto, 179–193.
- ter Meulen, Alice. 2000. Optimal reflexivity in Dutch. *Journal of Semantics* 17:263–280.
- Moltmann, Friederike. 1991. Coordination and comparatives. Doctoral Dissertation, MIT.
- Müller, Gereon. 1995. *A-bar syntax: A study of movement types*. Berlin: Mouton de Gruyter.
- Müller, Gereon. 1998. *Incomplete Category Fronting*. Dordrecht: Kluwer.
- Müller, Gereon, and Wolfgang Sternefeld. 1993. Improper movement and unambiguous binding. *Linguistic Inquiry* 24:461–507.

- Müller, Gereon, and Wolfgang Sternefeld. 1996. A'-chain formation and the economy of derivation. *Linguistic Inquiry* 27:480–511.
- Neeleman, Ad. 1994. Scrambling as a D-structure phenomenon. In *Studies on Scrambling: Movement and non-movement approaches to free-word-order phenomena*, ed. Norbert Corver and Henk van Riemsdijk, 387–429. Berlin: Mouton de Gruyter.
- Neeleman, Ad, and Hans van de Koot. 2010. A local encoding of syntactic dependencies and its consequences for the theory of movement. *Syntax* 13:331–372.
- Nemoto, Naoko. 1999. Scrambling. In *The handbook of Japanese Linguistics*, ed. Natsuko Tsujimura. Oxford: Blackwell.
- Nevins, Andrew. 2004. Derivations without the Activity Condition. In *Perspectives on phases: Proceedings of the workshop on EPP and Phases*. MIT Working papers in linguistics.
- Obata, Miki, and David Epstein. 2008. Deducing Improper Movement from Phase-Based C-to-T Transfer: Feature-Splitting Internal Merge. In *Proceedings of the 27th West Coast Conference on Formal Linguistics*, ed. Abner Natasha and Jason Bishop, 353–360. Somerville, MA: Cascadilla Proceedings Project.
- Oehrle, Richard. 1976. The grammatical status of the English dative alternation. Doctoral Dissertation, MIT, Cambridge, Mass.
- O'Grady, William. 1987. The Interpretation of Korean Anaphora: The Role and Representation of Grammatical Relations. *Language* 63:251–277.
- Ottósson, Kjartan G. 1993. Double-object small clauses and reanalysis in Icelandic passives. In *Proceedings of the 11th West Coast Conference on Formal Linguistics*, Stanford, California, 371–387. CSLI Publications.
- Partee, Barbara. 1978. Bound variables and other anaphors. In *Theoretical Issues in Natural Language Processing 2*, ed. David L. Waltz, 79–85. Urbana, IL: University of Illinois.
- Pesetsky, David. 1987. Binding problems with experiencer verbs. *Linguistic Inquiry* 18:126–140.
- Pesetsky, David. 1995. *Zero Syntax: Experiencers and Cascades*. Cambridge, Mass: MIT Press.
- Pesetsky, David. 2000. *Phrasal movement and its kin*. Cambridge, Mass: MIT Press.
- Pesetsky, David, and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In *Phrasal and clausal architecture*, ed. Simin Karimi, Vida Samiiian, and Wendy K. Wilkins, *Linguistik Aktuell/Linguistics Today* 101, 262–294. Amsterdam: John Benjamins.

- Pollard, Carl, and Ivan A. Sag. 1992. Anaphors in English and the Scope of Binding Theory. *Linguistic Inquiry* 23:261–303.
- Postal, Paul. 1971. *Cross-over phenomena*. New York: Holt, Rinehart, and Winston.
- Postal, Paul. 1974. *On raising: One rule of English grammar and its theoretical implications*. Cambridge, Mass: MIT Press.
- Preminger, Omer. 2010. Failure to agree is not a failure: ϕ -agreement with post-verbal subjects in Hebrew. In *Linguistic Variations Yearbook 2009*, ed. Jeroen van Craenenbroeck, 241–278. Amsterdam: John Benjamins.
- Putnam, Michael T. 2007. *Scrambling and the survive principle*. Amsterdam: John Benjamins.
- Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge University Press.
- Reinhart, Tanya. 1976. The syntactic domain of anaphora. Doctoral Dissertation, MIT.
- Reinhart, Tanya. 1983a. *Anaphora and Semantic Interpretation*. Croom Helm Linguistics Series. Croom Helm Press.
- Reinhart, Tanya. 1983b. Coreference and bound anaphora: A restatement of the anaphora questions. *Linguistics and Philosophy* 6:47–88.
- Reinhart, Tanya. 1995. *Interface Strategies*. Utrecht: OTS Working Papers.
- Reinhart, Tanya. 2000. Strategies of anaphora resolution. In *Interface Strategies*, ed. Hans Bennis, Martin Everaert, and Eric Reuland, 295–324. Amsterdam: North Holland.
- Reinhart, Tanya, and Yosef Grodzinsky. 1993. The Inateness of Binding and Coreference. *Linguistic Inquiry* 24:61–101.
- Reinhart, Tanya, and Eric Reuland. 1991. Anaphors and logophors: An argument structure perspective. In *Long-distance Anaphora*, ed. Jan Koster and Eric Reuland, 283–321. Berlin: Mouton de Gruyter.
- Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657–720.
- Reuland, Eric. 1998. Structural conditions on chains and binding. In *Proceedings of NELS 28*, ed. Pius N. Tamanji and Kiyomi Kusumoto, 341–356. University of Massachusetts, Amherst.
- Reuland, Eric. 2001. Primitives of Binding. *Linguistic Inquiry* 32:439–492.

- Reuland, Eric. 2006. Agreeing to bind. In *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, ed. Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz, and Jan Koster, 505–513. Berlin: Mouton de Gruyter.
- Reuland, Eric. 2007. Binding Conditions: How can they be derived? Lectures on Binding given at St Petersburg University.
- Reuland, Eric, and Jan Koster. 1991. Long-distance anaphora: an overview. In *Long-distance anaphora*, ed. Eric Reuland and Jan Koster, 1–25. Cambridge: Cambridge University Press.
- Rezac, Milan. 2003. The fine structure of cyclic Agree. *Syntax* 6:156–182.
- Rezac, Milan. 2007. Escaping the person case constraint: Reference-set computation in the ϕ -system. *Linguistic Variation Yearbook* 6:97–138.
- Richards, Marc D. 2007. On Feature Inheritance: An argument from the Phase Impenetrability Condition. *Linguistic Inquiry* 38:563–572.
- Riemsdijk, Henk van, and Edwin Williams. 1981. NP-structure. *The Linguistic Review* 1:171–217.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. Liliane Haegeman, 281–337. Dordrecht: Kluwer.
- Rizzi, Luigi. 2002. Locality and Left Periphery. In *Structures and Beyond. The Cartography of Syntactic Structures*, ed. Adriana Belletti. OUP.
- Rizzi, Luigi. 2006. On the Form of Chains: Criterial Positions and ECP Effects. In *Wh-movement: Moving on*, ed. Lisa Lai Shen Cheng, 97–134. Cambridge, Mass.: MIT Press.
- Rochemont, Michael. 1989. Topic islands and the Subjacency Parameter. *Canadian Journal of Linguistics* 34:145–170.
- Rooryck, Johan, and Guido Vanden Wyngaerd. 1999. Puzzles of Identity: Binding at the Interface. In *Proceedings of NELS 29*, ed. P. Tamanji, M. Hirotani, and N. Hall. University of Massachusetts, Amherst: GLSA.
- Rooryck, Johan, and Guido Vanden Wyngaerd. 2007. The Syntax of Spatial Anaphora. In *Tromsø working papers on language and linguistics: Nordlyd 34.2*, Special issue on Space, Motion, and Result, ed. Monika Bašić, Marina Pantcheva, Minjeong Son, and Peter Svenonius, 33–85. Tromsø: University of Tromsø. Available at <http://www.ub.uit.no/baser/nordlyd/>.
- Rooryck, Johan, and Guido Vanden Wyngaerd. 2010. *From first principles: Anaphors without Binding theory*. HUB Research Paper. Available at <https://lirias.hubrussel.be/handle/123456789/3617>.

- Ross, John R. 1967. *Infinite syntax*. Norwood: Ablex Publishing.
- Ross, John R. 1970. On declarative sentences. In *Readings in English transformational grammar*, ed. Roderic Jacobs and Peter Rosenbaum. Washington: Georgetown University Press.
- Sag, Ivan. 1976. Deletion and Logical Form. Doctoral Dissertation, MIT, Cambridge, Mass.
- Saito, M. 1989. Scrambling as semantically vacuous A'-movement. In *Alternative conceptions of phrase structure*, ed. Mark Baltin and Anthony Kroch, 182–200. Chicago: University of Chicago Press.
- Saito, Mamoru. 1992. Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1:69–118.
- Saito, Mamoru. 1994. Improper adjunction. In *MIT Working Papers in Linguistics 24: Formal Approaches to Japanese Linguistics I*, ed. Masatoshi Koizumi and Hiroyuki Ura, 263–293. Cambridge, Massachusetts: MIT Press.
- Sauerland, Uli. 1996. The Interpretability of Scrambling. In *MIT Working Papers in Linguistics: Proceedings of Formal Approaches to Linguistic Theory*, ed. Masatoshi Koizumi, Masayuki Oishi, and Uli Sauerland, 213–234. Cambridge, Massachusetts: MIT Press.
- Sauerland, Uli. 2003. Intermediate adjunction with A-movement. *Linguistic Inquiry* 34:308–314.
- Schladt, Mathias. 2000. The typology and grammaticalization of reflexives. In *Reflexives: Forms and functions*, ed. Zygmunt Frajzyngier and Traci Curl, 103–124. Amsterdam: John Benjamins.
- Shlonsky, Ur. 1991. Quantifiers as functional heads: A study of Quantifier Float in Hebrew. *Lingua* 84:159–180.
- Son, Minjeong, and Peter Svenonius. 2008. Microparameters of Cross-Linguistic Variation: Directed Motion and Resultatives. In *Proceedings of the 27th West Coast Conference on Formal Linguistics*, ed. Natasha Abner and Jason Bishop, 388–396. Somerville, Mass.: Cascadilla Press.
- Sportiche, Dominique. 1986. A Theory of Floating Quantifiers and its Corollaries for Constituent Structure. *Linguistic Inquiry* 17:369–374.
- Sportiche, Dominique. 1988. Zibun. *Linguistic Inquiry* 19:425–449.
- Sportiche, Dominique. 2006. Reconstruction, binding and scope. In *The Blackwell companion to syntax*, ed. Everaert Martin and Henk van Riemsdijk, 35–93. Oxford: Blackwell Publishing.

- Starke, Michal. 2001. Merge Dissolves into Move. Doctoral Dissertation, University of Geneva.
- Starke, Michal. 2003. Movement: the size factor. Presented at the CASTL Kick-Off Conference, University of Tromsø.
- Starke, Michal. 2006. Nanosyntax lectures. Semester-long seminar at the University of Tromsø.
- Sternefeld, Wolfgang. 1993. Transformationstypologie und strukturelle hierarchie. Zürich.
- Svenonius, P. to appear. Case Alternations in the Icelandic Passive. In *Passive and impersonals in european languages*, ed. Satu Manninen, Diane Nelson, Katrin Hiietam, Elsi Kaiser, and Virve Vihman. John Benjamins.
- Svenonius, Peter. 2004. On the edge. In *Peripheries: Syntactic Edges and their Effects*, ed. David Adger, Cécile de Cat, and George Tsoulas, 259–287. Dordrecht: Kluwer.
- Svenonius, Peter. 2006. The emergence of axial parts. In *Nordlyd, Tromsø Working Papers on Language and Linguistics. Special issue on Adpositions*, ed. Peter Svenonius and Marina Pantcheva, 49–77.
- Taraldsen, Tarald. 1996. Reflexives, pronouns and subject/verb agreement in Icelandic. In *Microparametric Syntax and Dialect Variation*, ed. James Black and Virginia Motapanyane, 189–211. Amsterdam: John Benjamins.
- Taraldsen, Tarald. 2008. Locality, exceptional ECM and the *that-trace effect. A talk given at Tarald Taraldsen celebratory workshop.
- Thráinsson, Höskuldur. 2007. *The syntax of Icelandic*. Cambridge University Press.
- Uriagereka, Juan. 1999. Multiple Spell-Out. In *Working Minimalism*, ed. Samuel D. Epstein and Norbert Corver, 251–282. Cambridge, Massachusetts: MIT Press.
- Voskuil, Jan, and Pim Wehrmann. 1990. On SC-subjects and binding. Ms. University of Leiden.
- Webelhuth, Gert. 1989. Syntactic saturation phenomena and the modern Germanic languages. Doctoral Dissertation, University of Geneva.
- Williams, Edwin. 1977. Discourse and logical form. *Linguistic Inquiry* 8:101–139.
- Williams, Edwin. 2003. *Representation theory*. Cambridge, Mass.: MIT Press.

- Wiltschko, Martina. 1998. Superiority in German. In *Proceedings of the sixteenth West coast conference on formal linguistics*, ed. Emily Curtis, James Lyle, and Gabriel Webster, 431–445. Stanford, CA: CSLI Publication.
- Wurmbrand, Susanne. 2001. *Infinitives: Restructuring and clause structure*. Berlin, New York: Mouton de Gruyter.
- Zeijlstra, Hedde. 2010. There is only one way to agree. Handout from a talk given at GLOW 33, Wroclaw University, April 2010.
- Zribi-Hertz, Anne. 1989. A-type binding and narrative point of view. *Language* 65:695–727.
- Zwart, Jan-Wouter. 1996. ‘Shortest Move’ vs. ‘Fewest Steps’. In *Minimal ideas: Syntactic studies in the minimalist framework*, ed. Werner Abraham, Samuel David Epstein, Höskuldur Thráinsson, and Jan-Wouter Zwart, *Linguistik Aktuell/Linguistics Today* 12, 305–328. John Benjamins.
- Zwart, Jan-Wouter. 2002. Issues relating to a derivational theory of binding. In *Derivation and Explanation in the Minimalist Program*, ed. Samuel D. Epstein and T. Daniel Seely, 269–304. Oxford: Blackwell.



ISBN xxx-xx-xxxx-xxx-x