

Successive Cyclicity, Anti-locality, and Adposition Stranding

Klaus Abels

Diplomübersetzer, Humboldt Universität zu Berlin, 2001

M.A., University of Connecticut, 2002

A Dissertaion

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2003

Copyright by

Klaus Abels

2003


Doctor of Philosophy Dissertation

Successive Cyclicity, Anti-locality, and Adposition Stranding

Presented by

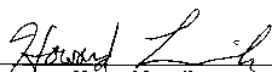
Klaus Abels, Diplomübersetzer, M.A.

Major Advisor



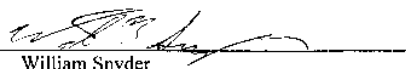
Željko Bošković

Associate Advisor



Howard Lasnik

Associate Advisor



William Snyder

University of Connecticut

2003

Acknowledgements

A dissertation cannot be written without the help of many people. And I mean a *lot* of people. Really.

I would like to thank my committee for carrying their share of the burden. Željko Bošković in five years never failed to support me. I surely won't forget the meetings in which I trashed my own proposals just to see Željko defend them. Željko also prodded me when prodding was necessary. Too often. Željko was always willing to provide insightful criticism share his knowledge of the literature – at least if one could figure out how to read his handwriting!. I thank Howard Lasnik for endless hours of discussion on every topic of syntax under the moon and the sun, but also for being an inspiring teacher and a humble human being. I might never have felt at home at UConn, had I not crawled around on the floor with Howard in the lab, trying to find spare computer parts to piece together 'Big Blue'. Finally, I'd like to thank William Snyder for his interest in even the wackiest proposals on P-stranding, long before this thesis was taking shape, for his ability to listen and ask just the right question to show me where I was going off-track or not done yet.

Many more people, of course, have heard parts of the story I tell here and have given me encouraging feedback, have answered my questions, criticized the ideas I had and sometimes still have despite their advise, or have provided judgments ... Among them are Andrew Nevins, Luisa Martí, Michal Starke, David Pesetsky, Misi Brody, Ad Neeleman, Ana Bastos, Anders Holmberg, Anna-Lena Wiklund, Anne Breitbarth, Anne Zribi-Hertz, Barbara Citko, Barbara Partee, Boban Arsenijević, Bum-Sik Park, Cedric

Boeckx, Danny Fox, Dave Braze, David Braun, Debbie Chen-Pichler, Diane Lillo-Martin, Dirk Bury, Dorian Roehrs, Doug Wharram, Esther Torrego, Gillian Ramchand, Hans den Besten, Hans van de Koot, Harry van der Hulst, Ivan Sag, Jairo Nunes, Jeff Parrott, Jim McCloskey, João Costa, John Frampton, Juan Uriagereka, Koji Sugisaki, Marcus Kracht, Marta Abrusan, Masashi Nomura, Mélanie Joutiteau, Michael Wagner, Miguel Rodriguez-Mondonedo, Milan Rezac, Mona Anderson, Natasha Rakhlin, Norbert Hornstein, Norman Clarke, Norvin Richards, Olga Tomić, Ora Matushansky, Peter Svenonius, Rajesh Bhatt, Sam Epstein, Sarah Felber, Serkan Sener, Sigrid Beck, Tjerk Hagemeyer, Tom Leu, Tony Kroch, Tsuyoshi Sawada, Yael Sharvit, Yakov Testele, and Yves Roberge. They all contributed to the thesis in one way or another. Thank you.

I'd also like to thank the other people who have helped me intellectually on my way towards becoming a linguist – roughly in order of appearance: Christine Maaßen – you must have seen something when you sent me to that first fateful summerschool ... thanks!, Daniel Buring, Zoli Graf, David Braun – not Brown, Lea Nash, Maaik Schoorlemmer, Chris Wilder, Marcel den Dikken, Peter Ludlow, Lyda Veselovska, Serkan Sener, Joanna Błaszczak, Michal Starke – how was I to know you don't speak German?, Emily McCoy – Absinthe and Krakow forever!, Ur Shlonsky, David Adger, Steven Franks, Luisa Martí – the woman who writes inviting letters and so much more..., the people in the fall 1998 first year classes at UConn: Norman Clarke – “[yü kō:n] [yü kō:n]” need I say more?, Isabel ‘joder’ Perez, Natasha Rakhlin – Vodka and politics for all!, Masashi Nomura – Don't buy another Vaio, please?!, Fumi Niinuma, Dorian Roehrs, Laura Sabadini, Bum-Sik Park, Sei-Rang Oh, and Emma Ticio.

After that my chronology breaks down. But I'd like to thank my more or less local linguist friends who made life in rural Connecticut bearable: Douglas Wharram – We're choosin' Susan, Doug., Debbie Chen-Pichler, Sarah Felber, Andrew Nevins, Nilüfer and Serkan Sener, Laurel LaPorte-Grimes – it don't mean a thing if it ain't got that swing, Eric Shortt, Ana Bastos, David Braun – Ecce homo. Ecco Walker., Marta Abrusan, Lydia Grebenyova, Elaine Grolla, Marcelo Ferreira, Dave Braze, Marcela Depiante, Mona Anderson, Eva Bar-Shalom, Jairo Nunes, Toshiko Oda, Nobu Miyoshi, Sachie Kotanie, Bosook Kang, Karlos Arregi, Cristiano Chesi – the VISION-ary.

I'd also like to thank Judy Marcus, the departmental secretary. You have managed to stay human despite everything – that's a bigger achievement than you might think.

I have more friends who have helped me see the light at various times or who have just made life a more bearable place: Olav Müller-Reichau, Dorothea Dette, Ebru Caba, Steve Martin, Marin Pichler, Bakul Badhnaga, the Willi food coop, Zsolt Nyiri, my family in Spain, Adam Potthast, Ryan Felber, Hedde Zeijlstra – I still can't spell your name. Sorry man!, Sylvia Blaho, Anna-Lena Wiklund, Mélanie Joutiteau, Milan Rezac, Dirk Bury, Anne Breitbarth, Jonathan Kaye, Mike Sukop, Yacov Kapiluto, and all the people on the Willimantic Peace Corner – Whose streets? Our streets!

Vor allem möchte ich mich aber bei den in Deutschland gebliebenen bedanken. Bei meiner Familie Elisabeth, Herbert, Monika, Edith, Georg, Regine, Simon, Antonia und jetzt auch Nuria. Und bei meinen Freunden, die mir die Treue gehalten habe, obwohl ich statt einem Jahr fünf Jahre weg war: Ulli Jensen, Friederike, 'Lieblingskommilitonin'

von Meltendorf, Ernst-Friedrich Kiel, Claudia, Thomas, Lennie und Elias Boike und
Alexandra Richter-Rethwisch.

Well, it's easy to list a bunch of people and say thanks. But how to thank Luisa
for sharing my life, I just don't know. Enkosi kakhulu, Luisa.

Table of Contents

Chapter 1: Introduction.....	9
Chapter 2: From Subjacency to Anti-locality.....	15
2.1 The theory of movement and punctuated paths	16
2.1.1 Uniform versus Punctuated Paths.....	18
2.1.2 Reconstruction: Is it nowhere, everywhere, or just somewhere along the path?22	
2.1.3 Punctuated Paths!	26
2.1.4 Conclusion.....	33
2.2 Escape Hatches.....	34
2.2.1 Punctuated Paths and the Null Hypothesis.....	37
2.2.2 Escape Hatches and Locality	38
2.2.3 Phases: Cyclic Spell Out vs. the Present View.....	62
2.2.4 Conclusion.....	75
2.3 On the Nature of C-command	76
2.3.1 C-command.....	78
2.3.2 Movement.....	84
2.3.3 Closer	88
2.3.4 Conclusion.....	90
2.4 Anti-locality	91
2.4.1 Feature Satisfaction	93
2.4.2 Last Resort and beyond	102
2.4.3 Conclusion.....	111
2.5 Pulling Together the Threads.....	111
Chapter 3: C°, v°, and anti-locality.....	114
3.1 TP immobility under C°	116
3.1.1 TPs do not strand C°.....	116
3.1.2 TPs are mobile in principle	121
3.1.3 Some non-cases of TP-movement.....	123
3.1.4 Consequences.....	131
3.1.5 Conclusion.....	138
3.2 VP immobility under v°	139
3.2.1 VPs are mobile.....	140
3.2.2 VPs do not strand v°	144
3.2.3 Conclusion.....	148
3.3 Conclusion.....	149
Appendix: Why do TPs move so little?	149

Chapter 4: Anti-locality and the ban against adposition stranding 157

4.1	The Stranding Generalization in non-P-stranding languages.....	159
4.1.1	Subextraction in Russian	160
4.1.2	'Left Branch Extraction' in Slavic.....	161
4.1.3	Combien-split in French	174
4.1.4	Subcomparatives into PP	176
4.1.5	Conclusion	177
4.2	On the diagnosis of P-stranding	178
4.2.1	On How to Distinguish Traces from (null) Resumptive Pronouns	179
4.2.2	Frisian	186
4.2.3	Conclusion	192
4.3	No P-stranding in German and Dutch!.....	192
4.3.1	R-R-pronouns-pronouns?.....	196
4.3.2	Are R-words P ^o -complements	201
4.3.3	So what are R-words? And where?	205
4.3.4	Extraction out of PP in German	210
4.4	Pronominal Clitics as the complement of P ^o	217
4.5	On the nature of P-stranding	230
4.5.1	The P-stranding generalizations	230
4.5.2	The Current Approach	233
4.5.3	Previous Approaches	240
4.5.4	Conclusion	259
4.6	Conclusion	260

Chapter 5: Head Movement and Anti-locality 262

5.1	Head Movement	264
5.1.1	The The-higher-the-bigger-Theorem	264
5.1.2	Head Movement in Mirror Theory Brody 2000	266
5.2	*{Q/Adv} – E-site	271
5.2.1	Sag's Generalization	273
5.2.2	The Radical Emptiness Account of Sag's Generalization	274
5.3	The Account	276
5.4	Conclusion	283

Chapter 6: Conclusion 285

References..... 287

Chapter 1: Introduction

Surgeon General's Warning:

This thesis on the syntax of natural languages is written within the Principles and Parameters tradition of Generative Grammar. Familiarity with the main tenets of Government and Binding Theory and the Minimalist Program is presupposed. Chapter 2 is there to scare away people who don't like formalism and theory. Afterwards it gets much easier, though.

The main persona of this thesis is the Stranding Generalization. The Stranding

Generalization is an unfamiliar character, so it bears introducing the generalization here together with its siblings and its lineage.

The Stranding Generalization can be stated as in (1). What (1) says is that for a particular class of heads, the phase heads, some (modalized) general statement is true.

The diamond symbol (\diamond) is used here with its usual interpretation in modal logic. Thus a structure like (1a) is possibly acceptable, a structure like (1b) is never acceptable. The

Stranding Generalization is a new generalization. Deriving and defending it is the main goal of this thesis.

(1) Stranding Generalization

Given a phase head α^o and a constituent X in α^o 's c-command domain

- a. $\diamond\checkmark[X \dots [\alpha^o [\dots t_X \dots]] \dots]$ and
- b. $\neg\diamond\checkmark[X \dots [\alpha^o t_X] \dots]$

You can think of the phase heads in my system in a first approximation as the same heads that project a phase in Chomsky 2000. In other words v^o and C^o and maybe some others are 'phase heads'. The Stranding Generalization says that phase heads may – when all the right conditions come together, which often they don't – a phase head may allow a constituent to move out of its c-command domain. That's (1a). But what a phase

head will never - never ever! - allow to happen is for its own complement to escape and move away out from under it.¹ And that's (1b).

Let me illustrate with an example from Icelandic what I mean. The complementizer *að* in Icelandic allows extraction across it in principle (2a). Icelandic also has a process that topicalizes clauses. Thus a CP complement of a verb like *think* can be topicalized without any major problems (2b-2c). But the TP, which is embedded under the complementizer *að* cannot be so topicalized. This is shown in (2d), which is totally unacceptable.

- (2) Icelandic (Halldór Sigurðsson, p.c.)
- a. ✓Hver heldur þú að hafi lesið þessa bók?
who think you that has_{SUBJ} read this book
lit: 'Who do you think that has read this book?'
 - b. ✓Jón heldur að María sé að lesa.
Jon thinks that Maria is_{SUBJ} to read_{INF}
'Jon thinks that Maria is reading.'
 - c. ?Að María sé að lesa heldur Jón.
that Maria is_{SUBJ} to read_{INF} thinks Jon
'That Maria is reading, John thinks.'
 - d. *Maríasé að lesa heldur Jón að.
Maria is_{SUBJ} to read_{INF} believes Jon that
- (3)
- a. ✓What do you think that Mary has read?
 - b. ✓Nobody thought that anything would happen.
 - c. ✓That anything would happen, nobody thought.
 - d. *Anything would happen, nobody thought that.
 - e. *Who do you think that has read this book?

The same is also true in English. Extraction past the complementizer *that* is possible (3a), topicalization of a CP is allowed (3c), but similar topicalization of TP is not allowed (3d). But whereas English is subject to the that-trace effect (3e) which may (or

¹ Since theses require footnotes, I will acknowledge that the last sentence was inspired by Jackendoff 1973 – in its form not in its content.

may not – I take up this question briefly in chapter 3) rule out example (3d), Icelandic is not subject to the that-trace effect (Maling and Zaenen 1978) as (2a) illustrates. What rules out example (2d) is not the that-trace effect on any interpretation, but rather the second part of the Stranding Generalization (1b). Stranding of complementizers seems to be universally banned.

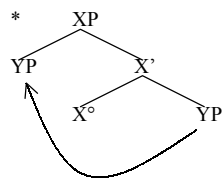
In chapter 3 I investigate v° and C° and spend a considerable effort to show that the Stranding Generalization is a surprising property of phase heads. It's surprising, because the complements of phase heads are in principle movable categories. It's the phase head above that immobilizes them!

In chapter 4 I pursue Emonds' (1985) idea that prepositions and complementizers are really the same thing, or at least almost, or at least the same kind of thing: phase heads. I claim that the ban against adposition-stranding, operative in most languages, is really just an instance of the Stranding Generalization. I also go through a lengthy detour in that chapter to call German and Dutch to order and bring them into the rank and file of non-P-stranding languages. Finally, I match up the Stranding Generalization against the most common theories of P-stranding. Any guesses who wins the match?

But I'm getting ahead of myself. I haven't told you yet about the other intriguing property that phase heads have. Well? – When a phase head allows movement across it as in (1a), the moving phrase must make a stop-over in the specifier projected by the phase head. If it is true that all phase heads require this stop over and obey the Stranding Generalization, then we would surely want to know why.

The answer is given in chapter 2, which provides the theoretical lineage of the Stranding Generalization. The idea is so simple that I wonder why chapter 2 got so long... Assume a condition, call it Last Resort, that says that movement is allowed only if there is a good reason for doing it – and the only thing that counts as a good reason is if you immediately establish of a new feature satisfaction (or ‘checking’) relation. (The Surgeon General warned you, didn’t he? Shouldn’t have read on!) So Last Resort is a local economy condition in the sense of Collins (1997). I claim that this categorically rules out movement of the complement of some head to the specifier of that very same head. The reason for this is that the Head-Complement Relation is the closest relation two things can be in in syntax. All features can be satisfied (or ‘checked’) in that relation. But if movement needs to give rise to some new feature satisfaction immediately, then there can never be any reason to move a phrase from the complement to the specifier position of the same head. This is really a corollary of Last Resort and some assumptions about what it means to ‘establish a feature satisfaction relation’, i.e. feature checking, but I call it the Anti-locality Constraint (4) anyway. The Anti-locality Constraint is totally general. It applies to all heads and their complements and is not limited to phase heads.

(4) Anti-locality Constraint



But let’s apply the Anti-locality Constraint to phase heads now. Movement out of the domain of a phase head requires movement through the specifier of that phase head. If you wanna get out, you absolutely must go through the specifier position. But since complements can never get to the specifier position in the phrase where they are complements, the complement of a phase head is frozen in place. Voilà, the Stranding Generalization: The complement of a phase head can never get out, because it cannot reach the specifier position, but that is an absolute requirement. The Stranding Generalization now is a corollary of the Last Resort Condition, some assumptions about feature satisfaction (‘checking’), and the funky property of phase heads that movement must pass through the specifier position they project.

But now come on already, why should phase heads have this funky property? – Well, ... I remember now why chapter 2 got so long... One answer is that phase heads simply have this property - period. Chomsky 2000 calls that the Phase Impenetrability Condition. But we want to do better, don’t we? So the idea is that movement through the specifier of a phase head is enforced by general considerations of Locality. Here’s the idea in a nutshell: nobody of the higher ups can see you (and if they can’t see you, hell, they can’t attract you!) if you are c-commanded by a phase head. Locality comes in like this: In syntax, if there’s a piece of blue cheese close to you, you can’t see let alone reach out for the piece of blue cheese further away. The bizarre property of phase heads – and this is my final version, I promise – is that they don’t just look like blue cheese, or [+wh], or something like that, they look like everything imaginable all at once! So if there’s a

glittery phase head in your way, you can't see anything at all beyond it. And everything the phase head c-commands is already beyond your reach...

Ah! The Anti-locality Constraint has a little brother: the ban against Head Movement. If movement from complement to specifier position of the same head is banned because it doesn't lead to feature satisfaction, then Head Movement should be banned on the same grounds. Head Movement is discussed in chapter 5.

And then there are a few roadblocks that I had to clear, like the idea that Right Node Raising might be movement and mess up the Stranding Generalization (chapter 3.1.3), or the problem of distinguishing when a prepositional phrase contains a trace and when it contains a resumptive pronoun as its complement (chapter 4.2), or indeed the idea that movement might just not work that way after all (chapter 2.1).

Let the fun begin!

Chapter 2: From Subjacency to Anti-locality

The bulk of this thesis is about impossible movement operations. In order to understand why certain movement operations are ruled out, we need to have a basic understanding of how movement works. Three aspects in particular have been the focus of study at various times.

- (5) What are the locality conditions on movement?
- (6) What are the triggers for movement? Why do particular things move? Why does movement exist as an operation to begin with?
- (7) How is long distance movement implemented in the grammar? Is long distance movement successive cyclic or is it one fell swoop movement?

Although questions (5) and (6) are certainly pertinent to the topic of this dissertation, it is really the answer to question (7) that matters. In section 2.1 of this chapter I point out a distinction between theories of long distance movement that postulate uniform paths and contrast them with theories that postulate punctuated paths. The distinction, which will be clarified below, is roughly this: some theories treat every node or every phrase along the path of movement uniformly. Theories where movement is unbounded are clear examples of uniform theories. All nodes along the path are uniformly unaffected by the movement past them.

The same holds for theories that postulate that successive cyclic movement proceeds by adjunction to every node along the path. Under this kind of theory every node along the path of movement is treated the same way: all get adjoined to. On the other hand are theories where movement stops over in a few select intermediate points along the path. This is true for example in the Barriers system (Chomsky 1986a) where stop-

overs are made in certain positions but not in others. Section 2.1 provides a novel kind of argument from reconstruction for the punctuated path approach.

Assuming punctuated paths, I turn to subadjacency in section 2.2. The main question here is how movement can be forced to pass through certain positions and why locality domains come with escape hatch positions. I explore what it would take to reduce escape hatch properties to the general locality constraint Attract Closest.

In section 2.3 I digress a little to explore the nature of c-command. This is necessary because of a paradox that arises under the theory developed in section 2.2. A novel approach to c-command is offered that is not only elegant but resolves the paradox.

In section 2.4 I explain my view of how movement is triggered. The point of this section is that if movement is too short, it is uneconomical and can therefore not be triggered. Section 2.4 offers the theoretical derivation of the stranding generalization from chapter 1. These results are crucial for the rest of the thesis.

2.1 The theory of movement and punctuated paths

For a long time, it has been a core concern of syntactic theory to characterize filler-gap relations in language adequately. The main question in this area concerns the locality of such filler-gap relations. One of the basic insights of locality theory is that, in characterizing the well-formedness conditions on filler-gap relations, only the path between the filler and the gap influences the well-formedness of the relation. Relativized Minimality interveners intervene with movement only if they lie along the path of movement but not otherwise, a barrier matters only as long as and in so far as it lies along the path of movement, etc. If we were to take a path to be a set of nodes in a tree, then a path construed in

terms of c-command and one construed in terms of domination are very different sets, but a notion of points along a path is available either way.² Broadly speaking, three notions are sufficient in characterizing locality theory (8): filler, path, and gap.

(8) $\text{filler} - \text{path} - \text{gap}$

This section concerns itself with the path, but it is not about locality proper. As far as the path is concerned, two types of theories are possible: (i) theories which treat every point along the path uniformly and (ii) theories according to which the path is not uniform but punctuated. Theories where movement is unbounded are clear examples of uniform theories as are theories where long distance movement is implemented as a kind of local feature percolation along the entire path. Theories according to which long distance movement proceeds in leaps and lands in only a few select intermediate positions would characterize the path as punctuated in my terminology.

This section provides an argument from reconstruction for the view that paths are punctuated. Unlike prior arguments, which cut no ice since they are based on the possibility of reconstruction to certain positions, the present argument comes from the impossibility of reconstruction to certain positions. To the extent that it is successful, the argument presented here offers much needed empirical support for theories that postulate punctuated paths.

² For concreteness, assume a standard constituency tree and a definition of c-command in terms of first branching node. Then, given a filler ϕ and a gap γ , a node v lies along the path between ϕ and γ iff (i) ϕ c-commands γ and ϕ c-commands v and (ii) v either c-commands (c-command path) or dominates γ (dominance path).

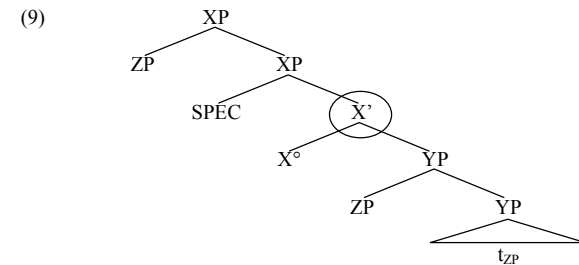
The section is organized as follows. Subsection 2.1 briefly reviews a number of syntactic theories and the position these theories take on the issue of uniform vs. punctuated paths. I then lay out the general logic of the issue. Subsection 2.2 concentrates on reconstruction for anaphor binding. I argue that the standard examples from the literature cannot decide the issue and offer a novel reconstruction paradigm that favors the punctuated paths approach.

2.1.1 Uniform versus Punctuated Paths

The question of whether the path between filler and gap is uniform or punctuated cuts across various theories of grammar. In fact every theory of grammar necessarily takes a stance on the question. Uniform paths are assumed for example in Head Driven Phrase Structure Grammar (HPSG, Pollard and Sag 1994), Tree Adjoining Grammar (TAG – Frank 2002; Joshi, Levi and Takahashi 1975; Kroch 1989; Kroch and Joshi 1985). In HPSG and other theories that invoke strictly local feature percolation (for a different example of this kind of theory see van de Koot and Neeleman, 2002), every node along the path between the filler and the gap inherits from its daughter the information that this node lies along the path of a filler gap relation. Early HPSG instantiated this by the so called slash feature. The path is treated uniformly in that every node along the path contains the information that this is a node along a movement path. In TAG the path is uniform as well but in a very different way. In TAG no node along the path of movement represents the information that this node is a node along the path of movement. The notion of path is only indirectly given as a byproduct of the filler-gap relation.

A question arises with respect to certain theories of movement within the frameworks of Government and Binding and Minimalism (Boeckx 2001; Bošković 2002b; Chomsky and Lasnik 1993; Fox and Lasnik 2003; Manzini 1992; Takahashi 1994 for A'-movement though not for A-movement). These theories postulate that movement (at least in the case of A'-movement) is very local. It proceeds by adjunction to every XP along the path. This makes movement not strictly speaking uniform since not every node along the path is affected. Where the difference matters I will call such theories quasi uniform.

The difference between quasi uniform and truly uniform theories is relatively small. Under the quasi uniform theories movement of ZP in (9) proceeds by adjunction to YP and to XP but not to X', which remains unaffected. Under a truly uniform theory there would also be adjunction to X', circled in (9). This opens the way for a differential behavior of XP vs. X' for some grammatical process under the influence of movement. Truly uniform theories do not predict such an asymmetry between XP and X'.



In order to tease apart the truly uniform from the quasi uniform theories, a test with extremely high local resolution would be needed. The test that I will actually use to tease apart uniform from punctuated theories is Binding Condition A (Chomsky 1986b),

which has a rather coarse local resolution. It just doesn't have the kind of local resolution required to distinguish uniform from quasi uniform theories, which I will therefore lump together in the discussion as uniform.

On the other side are theories like the Extended Standard Theory of the 1970s with its designated bounding nodes and escape hatches (Chomsky 1973; see for a precise formulation Riemsdijk's Head Constraint 1978a: 169, (40) below). Movement has to stop in these escape hatch positions but not elsewhere: movement leaps from one escape hatch to the next bypassing other positions. The jumps are fairly big giving us a good chance at detecting affected and non-affected points. Similarly, the theory of Barriers (Chomsky 1986a) is a theory with a non-uniform path since A'-movement must pass through [Spec, CP] and the VP-adjoined position, but need not, and in fact cannot, pass through [Spec, IP] or the IP-adjoined position. Finally, in his recent formulations of the Minimalist Program Chomsky (2000; 2001a; 2001b) suggests that long distance movement proceeds via the specifier positions at the edge of certain local domains, the phases. Movement need not pass through the positions intervening between two phase edges and, by economy, it must not. Movement through all intermediate XPs is barred technically because (i) all movement is driven by a feature on the target of movement and (ii) only certain heads, the ones that head a phase, can freely be assigned the relevant 'extra' EPP-feature enabling successive cyclic movement. The precise details of this instantiation are irrelevant, though, since the intuition is clear enough that movement must pass through all and only the edges of phases. Thus the phase theory again postulates punctuated paths.

The path between a filler and a gap is relevant for various grammatical processes. The textbook example of a process affected by the presence or absence of a movement path is the complementizer 'agreement' paradigm in Irish (McCloskey 1979; 1990 and much related work). Another process involving paths is reconstruction to an intermediate position (this line of research goes back at least to Barss 1986). We will deal with reconstruction in the next section. Further processes that come to mind are the possibility of quantifier float along the path of movement (Sportiche 1988; Bošković 2002b; McCloskey 2000 and focusing particles along the path of movement (Barbiers 1995 chapter 3.3). Regarding such processes, the two kinds of theories give rise to very different expectations. The theory of uniform paths leads to the expectation that in principle all nodes (or all XPs for the quasi uniform theories) along the path of movement ought to be able to be affected by such processes; all heads along the path of movement should in principle be able to show special agreement; all categories along the path should be potential reconstruction sites; everything else being equal, quantifiers should be able to be floated anywhere along the path; etc. The theory of punctuated paths on the other hand leads to the opposite expectation, namely that these processes should be highly selective; special agreement should be possible only on a select few heads along the path; reconstruction should be possible only to a limited range of positions; floated quantifiers and floated focus particles should be highly limited in their distribution. Moreover, the theory of punctuated paths predicts a clustering of properties: the heads that can undergo special agreement ought to be the very heads projecting the phrases to

which reconstruction is possible and in whose projections quantifiers and focus particles can be floated.

These considerations show that the issue is ultimately an empirical one. Whether paths are punctuated or uniform cannot be decided by fiat. The research program laid out in the preceding paragraph is not the main topic of this thesis however. In this section, I will simply demonstrate that there is evidence from reconstruction favoring the punctuated paths approach.

Both kinds of theories have been articulated within a syntactic idiom using the terminology of ‘movement’ and ‘traces’. For this reason, I will frame the discussion in this idiom. It allows the most direct comparison. Concretely, I compare a theory where A’-movement leaves a trace in every maximal projection along the path (Boeckx 2001; Chomsky and Lasnik 1993; Fox and Lasnik 2003; Manzini 1992; Takahashi 1994) with a theory where movement leaves a trace only in certain, select positions along the path as in the phase-based theory (Chomsky 2000). These theories simply exemplify the logic of uniform paths versus punctuated paths. As far as I can determine, nothing in the argument hinges on this particular choice of syntactic idiom. The argument I will present against leaving traces in every maximal projection along the path is, in fact, an argument against any theory of uniform paths.

With this as background, I now turn to reconstruction.

2.1.2 Reconstruction: Is it nowhere, everywhere, or just somewhere along the path?

As Barss (1986) has demonstrated, local anaphors can be licensed from intermediate positions along a path of movement. *Himself* in English is an anaphor that needs to be local-

ly bound. Lack of a local binder rules out example (10), where the binder, *John*, is too far away from the anaphor.³ Example (11) on the other hand is grammatical. The *wh*-phrase *which pictures of himself* moves to the embedded [Spec, CP] carrying *himself* to a position close enough to the antecedent, *John*.

- (10) * John wonders whether Mary likes some pictures of himself.
- (11) ✓ John wonders which pictures of himself Mary likes.

Likewise, in examples (12) and (13).

- (12) * John said that Mary likes this picture of himself.
- (13) ✓ John said that this picture of himself Mary likes.

In its base position *himself* is too far away from its antecedent to be licit, but when the DP *this picture of himself* is topicalized, the anaphor is brought into a local enough relation to its antecedent in the matrix clause. The question arises where exactly the topicalized DP in (13) is situated. Lasnik and Saito (1992) claim that topicalization targets the IP-adjoined position, however it is also possible that in example (13) the DP is situated in the specifier position of some dedicated topic phrase between IP and CP.

Chomsky (1981; 1986b:173) claims that anaphors and reciprocals located within [Spec, IP] can also be bound by the matrix subject. This is shown in (14) (see Manzini 1992 for discussion of variation on this).

- (14) a. ✓ The children thought that pictures of each other were on sale.
- b. ✓ The children thought that each other’s pictures were on sale.
- c. ✓ Peter thought that pictures of himself were on sale.

³ For concreteness assume the definition of Binding Condition A from Chomsky 1986b.

We thus have evidence here that anaphors can be bound by the matrix subject if they are located as low as within the embedded [Spec, IP].

The examples involving movement are, unsurprisingly, also compatible with anaphor binding by the embedded subject. Thus all of (15)-(18), which contrast minimally with (10)-(13), are acceptable.

- (15) ✓ John wonders whether Mary likes some pictures of herself.
- (16) ✓ John wonders which pictures of herself Mary likes.
- (17) ✓ John said that Mary likes this picture of herself.
- (18) ✓ John said that this picture of herself Mary likes.

What we have seen so far indicates that the binding conditions for anaphors may be met either in the base position of an element or in its surface position. Furthermore, we have seen that both the embedded [Spec, CP] position, the topic position (i.e. the position adjoined to an embedded IP or the specifier of some dedicated head between I° and C°), and a position within [Spec, IP] itself are local enough to the subject of the next higher clause to allow anaphor binding.

Interestingly, as Barss discovered, binding in the base position and binding in the surface position do not exhaust the range of possibilities. Thus example (19) contrasts with (20). Example (19) is expectedly ungrammatical, but (20) is somewhat surprising since *himself* is in the relevant local relation with its antecedent neither in its base position nor in its surface position.⁴

- (19) * Jane believes (that) John thinks (that) she likes some pictures of himself.
- (20) ✓ Which pictures of himself does Jane believe (that) John thinks (that) she likes?

⁴ It is often suggested that in examples like (20) *himself* is not a local anaphor but a logophor. I disregard this possibility here for reasons that will become clear shortly.

This indicates that the binding relation must have been established somewhere along the path of movement. Examples like these have sometimes been called pit-stop reflexives.

According to the uniform path theory we are considering the relevant part of the structure of example (20) would be as in (21). The alternative punctuated theory, would posit only two relevant intermediate traces as shown in (22).

- (21) [which picture of himself]_i [... John ... [_{VP} _{t_i} _v^o [_{VP} _{t_i} thinks [_{CP} _{t_i} that [_{TP} _{t_i} Mary [...]]]]]]]
- (22) [which picture of himself]_i [... John ... [_{VP} _{t_i} _v^o [_{VP} thinks [_{CP} _{t_i} that [_{TP} Mary [...]]]]]]]

The uniform path theory posits a trace in every maximal projection, the punctuated theory only at the edge of a select few. Nevertheless, both are equipped to explain why example (20) is grammatical, because both posit traces in positions that are sufficiently local to the binder, *John*.⁵ In particular, the traces in CP and in vP in (22) are close enough to the subject, *John*, to allow binding of the anaphor.

⁵ It is sometimes assumed that examples like these argue for the existence of traces or copies in the representation or their derivational equivalents: stages in the derivation where binding can apply locally. These arguments are mistaken, however, since the relevant statements of binding theory could be formulated in terms of paths rather than traces, copies, or intermediate derivational stages obviating the need to postulate them.

Since this is such a common misunderstanding, let me pursue the issue a bit further. A typical formulation of binding principle A says, roughly, that antecedent α may bind anaphor β if (i) α c-commands β and (ii) there is no SUBJECT intervening between α and β . A formulation that is generalized to filler-gap relationships enabling binding along the entire path but without reference to intermediate traces or similar devices would be as follows: α may bind anaphor β if (i') α c-commands β or a gap of an item containing β and (ii') there is no SUBJECT that c-commands β and does not c-command α . Si-

Examples like these have been used to establish the (relatively weak) claim that binding is possible from some intermediate positions along the path of movement (see Lebeaux 1995; and Nissenbaum 2001 for related discussion).

Obviously though, this does not answer the question whether the path is uniform or punctuated. To positively demonstrate that the path is uniform, we would have to show that intermediate binding is possible anywhere along the path. Given the fairly poor spatial resolution of the principles of Binding Theory, this is an impossible task.

Demonstrating that the path is punctuated is, in principle, much easier. What we would need to show is that binding into certain intermediate positions is impossible. The next section provides a paradigm intended to show exactly that.

2.1.3 Punctuated Paths!

We saw in the previous section that subjects of higher clauses can license anaphors which are not in their base position (11, 13, 20). Objects in higher clauses can also license pit-stop reflexives as (23-26) show.

- (23) * Mary told John that she liked these pictures of himself.
- (24) ✓ Which pictures of himself did Mary tell John that she liked?
- (25) * It seemed to John that Mary liked these pictures of himself.
- (26) ✓ Which pictures of himself did it seem to John that Mary liked?

Barss has shown that anaphor binding in such contexts is even possible when the relevant movement step violates a *wh*-island. This is illustrated in (27) and (28). Example (28) is somewhat degraded, not because the binding relation into the *wh*-phrase is illicit,

milar formulations for scope reconstruction effects are possible without reference to intermediate traces. I leave it as an exercise to the reader to formulate them.

but because the movement of *which pictures of himself* violates a *wh*-island as a comparison between (28) and (29) reveals.

- (27) a. * Mary asked John whether she should buy these pictures of himself.
b. * Mary asked John when she should buy these pictures of himself.
- (28) a. ? Which pictures of himself did Mary ask John whether she should buy?
b. ? Which pictures of himself did Mary ask John when she should buy?
- (29) a. ? Which pictures of Fred did Mary ask John whether she should buy?
b. ? Which pictures of Fred did Mary ask John when she should buy?

For the uniform path theory, these examples are of course as expected. The theory positing punctuated paths might run into problems with examples such as these. All depends on where exactly intermediate landing sites are posited.

An analysis of examples like (28b) in terms of Barriers (Chomsky 1986a) for example, is not entirely straightforward. In the Barriers framework, the *wh*-phrase would adjoin to the embedded VP. It would not adjoin to the embedded IP, since adjunction to IP is never possible. It would not be able to land in the embedded [Spec, CP], since [Spec, CP] is already filled by *when*. It could also not adjoin to the embedded CP, since this CP is an argument and adjunction to arguments is prohibited in Barriers. The next available landing site of the *wh*-phrase is then the matrix VP. This is illustrated in (30).

- (30) [which pictures of himself]_i Mary ask [_{VP} t_i'] [John t_{ask} [_{CP} when [_{IP} she should [_{VP} t_i'] buy t_i]]]]]

The problem is that none of the traces subscripted with 'i' are c-commanded by *John* and sufficiently local to it for anaphor binding to take place in (30).

One way out of this problem is to assume a richer structure in the matrix clause.

The matrix VP might have a shell structure along the lines shown in (31), where the trace in VP₂ is both c-commanded and local to the antecedent of the anaphor, *John*.⁶

- (31) [which pictures of himself]_i Mary [VP₁ t_i ask John [VP₂ t_i t_{ask} [CP when [IP she should [VP t_i buy t_i]]]]]

It could also be that *John* is in [Spec, VP] and t_i’ adjoined to VP as in (30) and that this is sufficient for John to c-command himself locally.⁷

Alternatively, we could reject the Barriers assumption that there is no intermediate trace in the projection of CP. In a theory like Chomsky’s recent phase theory (Chomsky 2000; 2001a; 2001b) for example nothing rules out an intermediate trace in the embedded CP as far as I can see. In fact, the phase impenetrability condition forces the presence of such an intermediate trace. The *wh*-island effect would then not stem from the fact that movement fails to skip a potential landing site, but because at some point Relativized Minimality (Rizzi 1990), here understood as a constraint against crossing, is violated.⁸ This

⁶ The discussion is sketchy here on purpose. It could be that *John* is base generated in the upper VP-shell as in (31) or it could be that *John* moves to the upper shell (or some other position above the intermediate trace of the *wh*-phrase, possibly Agr^o) for Case reasons, i.e. object shift.

⁷ This kind of c-command of a specifier into a higher adjunct or a higher second specifier has sometimes been suggested for certain kinds of multiple nominative constructions in Japanese. The argument is not affected by this as far as I can see.

⁸ The question is whether *wh*-island violations are violations of the Phase Impenetrability Condition in addition to a feature based version of Relativized Minimality, i.e. Attract Closest.

(i) wh₁ ... [CP (t_{wh1}’)] wh₂ ... t_{wh2} ... t_{wh1}]

is illustrated in (32). In (32) the movement from the embedded vP to the embedded [Spec, CP] crosses another *wh*-phrase violating Relativized Minimality. If this last possibility is on the right track, this would be evidence that multiple specifiers to the same head do not count as equidistant (contra e.g. Richards 1997) and for a more strictly geometric approach to distance. Given the stipulative nature of definitions of equidistance, this is a gain rather than a loss. Of course this requires a reanalysis of the facts that equidistance is standardly invoked for, i.e. the problem that arises if we assume that objects are generated below subjects but check their case features in an agreement projection above the base position of the subject. See Koizumi 1995 for a solution without equidistance and without crossing. See also Richards 1997; Starke 2001; Williams 2002 for general discussion of crossing dependencies and order preservation.

- (32) [which pictures of himself]_i Mary [vP t_i ask [John [t_{ask} [CP t_i when [IP she should [vP t_i buy t_i]]]]]]]

What all this shows is that the grammaticality of a particular example, like (28b) above, can be used to argue against a particular version of the theory of punctuated paths, namely that illustrated in (30). Example (28b) cannot, however, be used to argue against the theory of punctuated paths in general as (31) and (32) illustrate. Together with the considerations at the end of subsection 2.1.2 this goes to show that arguments coming from the grammaticality of particular binding and scope possibilities support only

Wh-island violations are relatively weak. They violate Attract Closest anyway because either the movement from t_{wh1} to t_{wh1}’ violates Attract Closest or if this step is not realized, then the direct movement t_{wh1} to wh₁ will violate Attract Closest. The question is thus whether the Phase Impenetrability Condition is violated in addition. Given the weakness of the violation this seems unlikely.

extremely weak conclusions as far as the question is concerned whether paths are uniform or punctuated.

We are now ready to look at the crucial set of examples. Recall from above that the experiencer of the verb *seem* can bind an anaphor contained in a *wh*-phrase in an intermediate position (26). This possibility disappears when *seem* is used as a raising verb as in (33). Example (33) should be contrasted with examples (34) and (35). Example (34) simply shows that binding of the anaphor by the experiencer is impossible from the base position of the object of *like*. Example (35) demonstrates that the fact that (33) is unacceptable can truly be traced to anaphor binding and not, for example, to some failure of *wh*-movement to be licit.

- (26) ✓ Which pictures of himself did it seem to John that Mary liked?
 (33) * Which pictures of himself did Mary seem to John to like?
 (34) * Mary seemed to John to like these pictures of himself.
 (35) ✓ Which pictures of Frank did John seem to Mary to like?

The ungrammaticality of example (33) supports a whole gamut of otherwise independent, non-trivial conclusions. The most obvious conclusion from (33) is that the theory of uniform paths cannot be correct. The theory of uniform paths predicts that binding at very point along a given path should be possible, but (33) illustrates that this is not so. These are sufficient grounds for rejecting the theory of uniform paths. On the theory of punctuated paths under consideration, the fact that (33) is ungrammatical can be captured by assuming that there is no intermediate trace along the path between *which pictures of himself* and its trace, which is both local in the sense of Binding Theory to and c-commanded by *to John*.

- (36) [Which pictures of himself]_i Mary [_{VP1} (t_i) seem [_{VP2} to John t_{seem} [_{TP} t_{Mary} to [_{VP3} t_i like t_i]]]]]

The most plausible representation for (33) is given in (36). There are two VP-shells in (36) to make space for the experiencer and to account for the word order. The crucial fact about (36) is that there cannot be an intermediate trace of *which pictures of himself* adjoined to the embedded TP.⁹ We saw above in example (13) that adjunction to an embedded TP allows binding from the matrix clause; hence, if there were an intermediate trace adjoined to TP, (33) would be grammatical counter to fact. Following Chomsky (1981; 2000), (36) treats the raising infinitive as a bare TP (or IP in the older terminology), not as a CP.¹⁰ Finally, there is no intermediate trace in (36) within VP₂, because such a trace would be local to and c-commanded by *to John*. The question whether ‘seem’ actually projects the kind of shell-structure shown in (36) and whether there is a trace in VP₁ is irrelevant to our present concerns – though see fn. (7).

The argument given here has an indirect bearing on how to deal with example (28), the example of a pit-stop reflexive coming out of a *wh*-island. The discussion here lends plausibility to (32) with an intermediate trace in [Spec, CP] or CP-adjoined position

⁹ In Chomsky’s (1986a) system adjunction to TP would be impossible in (36) because (a) adjunction to arguments is banned in general and (b) adjunction to IP (my TP) is banned in general. What is at issue here is the impossibility of adjunction to TP in this particular structure not the question why it is impossible.

Any theory that rules out adjunction to TP here is, by its nature, a theory of punctuated paths.

¹⁰ There have occasionally been arguments in the literature that raising infinitives are CPs (Ormazabal 1995; Pesetsky and Torrego to appear). Discussion of the arguments offered by these authors would lead too far afield here. The argument I am in the middle of presenting shows that raising infinitives are not CPs.

even in *wh*-island configurations. There is no knockdown argument for this structure here, however. The ultimate answer to these questions will hinge on the resolution of some questions surrounding object shift in the sense of Lasnik 1995c, 1999b.

The ungrammaticality of example (33) thus gives us good empirical grounds for rejecting the notion that raising infinitives are CPs. If the raising infinitive were a CP, we would have to ensure that there is no intermediate trace in CP as shown in (37) – see fn. 10.

- (37) [Which pictures of himself]_i Mary [_{VP1} (t_i) seem to John [_{VP2} t_{seem} [_{CP} ... [_{TP} t_{Mary} to [_{VP3} t_i like t_i]]]]]

The presence of such an intermediate trace in [Spec, CP] or in a position adjoined to CP would allow the anaphor to be bound by *John*. Since this binding relation is impossible, there cannot be a trace in the putative embedded CP. The absence of a trace in CP would be very surprising. There is converging evidence from complementizer agreement along the path of movement (McCloskey 1979, 1990), from quantifier float (McCloskey 2000), from partial *wh*-movement (Bruening 2001), and from copying *wh*-movement (Nunes to appear) that movement can pass through [Spec, CP].¹¹ This is true under any theory, uniform or punctuated. But again, if there was an intermediate trace in CP in (37), the account of the ungrammaticality of (33) would be lost.

Finally, consider again the contrast between examples (33) and (26). The fact that (26) is acceptable is sometimes attributed to the idea that *himself* in English might have a logophoric use alongside its anaphoric use (e.g. Reinhart and Reuland 1993). Since it is

¹¹ See also Chung 1998; Collins 1993 chapter 4.4; Tuller 1986 among others.

entirely unclear how the distinction between (26) and (33) could come about under the logophoric approach, (33) gives us grounds to reject the account of (26) in terms of logophors.

2.1.4 Conclusion

In this section I have discussed two approaches to the relation between fillers and gaps, termed here the uniform path hypothesis (exemplified by some Minimalist theories, HPSG, TAG) and the punctuated path hypothesis (exemplified by bounding theory, Barriers, phase theory within Minimalism). I showed that the literature is marked by a lack of empirical evidence bearing on the issue. In subsection 2.1.3 I offered novel data that argue for the punctuated path hypothesis and against the uniform path hypothesis.

The crucial cases also lend credence to the conclusion that raising infinitives are TPs rather than CPs, and raise considerable difficulties for explanations of pit-stop reflexives in terms of logophoricity.

Hopefully the above discussion will inspire the search for comparable paradigms that can serve to directly support the punctuated path hypothesis. This subsection note can be seen as a step in the larger research program outlined in subsection 2.1.1. The aim of this research project will be to establish that the punctuated path hypothesis is correct by showing that intermediate reconstruction, agreement along a path, quantifier float, and possibly other diagnostics occur only in a limited range of positions along a path and that all of them cluster together.

2.2 Escape Hatches

Over time several ways have been suggested to force movement paths to be punctuated. Chomsky's (Chomsky 1973) principle of subadjacency is a milestone as are van Riemsdijk's (1978a:169) Head Constraint (recently resurrected under the name of Phase Impenetrability Condition by Chomsky (2000:108)) and Chomsky's notion of Barriers (Chomsky 1986a).

What all of these different approaches have in common is that they seal off certain domains and make available designated escape hatches to allow movement out of the sealed domains after all.

These kinds of theories have met with quite a bit of resistance. First of all, the fact that movement paths are punctuated in the sense of the previous section is something of a surprise. It is not the null assumption. The heated debates in the 70s on whether movement is unbounded or successive cyclic bears witness to this fact as does the frequency with which theories of uniform paths but without unbounded movement are suggested (see the references in the last section).

The other major objection has to do with the properties and the location of the escape hatches. Under the standard view, the properties of escape hatches are stipulated. There is no answer to questions such as the following. Why is [Spec, XP] or the position adjoined to XP an escape hatch and not any other definable position?¹² If there is a

¹² This is not just a problem of theories of punctuated paths but of all theories that use escape hatches. TAG and theories that use feature percolation or unification do not face this objection.

locality boundary somewhere, shouldn't either the movement to or the movement from the escape hatch position cross that boundary and hence fail to be local? More concretely, consider Chomsky's recent implementation in terms of the 'edge' of a phase. In talking about Chomsky's proposal I will use the spelling EDGE to distinguish Chomsky's technical term from the intuitive notion of an edge. There is a straightforward geometric notion of the edge of an XP: the unique constituent c-commanding all others contained in that XP. But that is not the technical definition of the EDGE, which includes the edge, all other specifiers and the phase head. But why? And why isn't EDGE defined in some other way? The objection, in other words, is that the escape hatch property is a stipulation that does not follow from independent properties of the theory. It and its corollary successive cyclicity are simply grafted onto the theory. I try to remedy this as best I can in this section.

I have no answer to the question why C° for example projects a phase but T° does not. This difference remains stipulated. However, I will discuss the other two objections. Why have punctuated paths at all and how can the properties of escape hatches be made to follow from more general considerations. The second issue arises for quasi-uniform theories discussed above as well, at least to the extent that they use escape hatches. Afterwards I return to the issue of subadjacency. I arrive at a theory that has certain properties in common with Chomsky's theory of phases. Like Chomsky, I assume that there is a set of syntactic heads whose properties force any movement past those heads to pass through the specifier position. These heads are 'phase heads' both in my and in Chomsky's system. The set of heads that have this property is stipulated. In contrast to

Chomsky's system, where phase heads have an array of functions, they are used in my system to regulate only successive cyclicity. For Chomsky phases and their heads play a role for locality and successive cyclicity, too, but they also determine points of spell out and define relevant portions of the numeration as well as phonologically and semantically isolable chunks of structure. I do not follow Chomsky in these assumptions. For this reason, the theory pursued here bears even more similarity with traditional bounding theory than with phase theory. It is, in fact, a fairly direct successor of van Riemsdijk's (1978a) theory.¹³ The crucial difference between bounding theory and the present system is that in bounding theory, the action is in the bounding nodes, which are all XPs, whereas here the action is not in the XPs but in the X° .

Having clarified the issue here, I will continue to call the relevant heads 'phase heads'. Where confusion between the present theory and Chomsky's theory is likely to arise, 'phase heads in the sense of Chomsky' will be distinguished from 'phase heads in the sense of the present theory'.

In section 2.2.1 and 2.2.2 I develop the present system. In 2.2.3 I turn to some differences between Chomsky's theory of phases and the present theory. I motivate the decision to adopt the present system by way of criticizing Chomsky's notion of phase.

¹³ In a sense Chomsky's theory is a descendant of van Riemsdijk's theory as well, since Chomsky's Phase Impenetrability Condition is identical to van Riemsdijk's Head Constraint. This is shown in section 2.2.2.1.

2.2.1 Punctuated Paths and the Null Hypothesis

A theory of punctuated paths deviates from the null hypothesis. However, if the considerations of section 2.1 are correct, we are forced to accept the idea that movement paths are punctuated. Consider this result in a broader context.

We know that certain relations operate at a distance without being affected by the path and without the relation affecting the path. Examples of relations of this type are variable binding, scope, non-coreference under condition C of the Binding Theory, etc. Then there is a relation (i.e. projection) which is both sensitive to every node along the path and which affects every node along the path. Projection from one source is blocked by projection from another source (sensitivity) and every node along the projection line of X° is itself a projection of X° and inherits properties from X° (effect on the path). Finally there is movement, which is sensitive only selectively to the path (Relativized Minimality and phases) and which only selectively affects the path (reconstruction, cyclicity effects, float, etc.)

Given a theory of phrase structure, there are three possible ways in which a relation between two items α and β that are not in one of the local relations of sisterhood and immediate dominance can affect the path between α and β . The three options in (38) are the now familiar ones from the previous section.

- (38)
- | | | |
|----|---|---|
| a. | α β
 | uniform path, path unaffected by relation |
| b. | α β
 | uniform path, path fully affected by relation |
| c. | α β
 | punctuated path |

As just discussed, language appears to have three different types of non-local relations. Given their substantially different properties, a unification seems unlikely at the moment. But then, the three possibilities in (38) are a parsimonious way of expressing the three different types of relations found in language.

2.2.2 *Escape Hatches and Locality*

A second reason why punctuated path theories have been resisted is the stipulative nature of the escape hatch theories they utilize. This argument has been most forcefully made in the literature on Tree Adjoining Grammar (e.g. Kroch 1989).

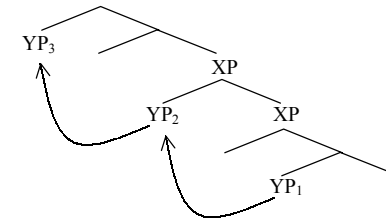
Take as an example the Barriers approach. Barriers defines locality domains in terms of a concept of dominance; under certain circumstances an XP is a barrier.¹⁴ This means that no phrase can move from a position that is dominated by the barrier to a position that is not dominated by the barrier. If a simple view of dominance is adopted, the given formulation does not allow any movement to reach a position outside of the Barrier, because some step in that movement necessarily goes from a position dominated by the barrier to a position not so dominated. To overcome this problem, the basic intuition is softened. Instead of using just the dominance relation – and its complement (called the ‘exclusion’ relation), Barriers makes use of three relations, namely the regular dominance relation (inclusion), the complement of the regular dominance relation (exclusion), and a relation misleadingly called ‘dominance’ in Barriers which I will

¹⁴ The conditions under which an XP counts as a Bounding Category and a Barrier need to be stipulated (see Frampton 1999 for relevant discussion). I concentrate here exclusively on the way the escape hatch property is implemented.

notate as $\text{dominance}_{\text{Barriers}}$. $\text{dominance}_{\text{Barriers}}$ applies only to categories (as opposed to nodes or to segments) such that α $\text{dominates}_{\text{Barriers}}$ β iff every segment of the category α dominates β .¹⁵

The basic intuition is then replaced by the following locality condition: no phrase can move from a position that is $\text{dominated}_{\text{Barriers}}$ by a barrier to a position that is not dominated by any segment of that barrier. Together with the stipulation that adjunction to a phrase creates a multi-segment category, these definitions allow escape hatches to be created through adjunction. This is illustrated in (39).¹⁶

(39)



Suppose XP in (39) is a barrier for YP. The movement step from YP₁ to YP₂ does not violate the locality condition. YP₁ is $\text{dominated}_{\text{Barriers}}$ by the barrier XP. YP₂ is

¹⁵ A ‘category’ is the collection of all nodes that are projected from the same head and that share the same bar level of projection.

Two distinct nodes are segments of the same category iff they are both members of the category. Thus, every node is a segment of some category, but not every node is a category.

¹⁶ Kroch’s criticism of the stipulative nature of such systems applies not only to systems with punctuated paths that use escape hatches but also to uniform theories that use escape hatches (e.g. Fox and Lasnik 2003). In other words, all of mainstream GB and Minimalism have to answer this challenge whether paths are construed as punctuated or not.

dominated by a segment of XP but not dominated_{Barriers} by it. Therefore the first step of movement in the example is allowed by the above condition. The second step is also allowed. Of course, YP₃ is not dominated by any segment of XP, but this does not matter, since the launching site of movement, i.e. YP₂ is not dominated_{Barriers} by XP.

This two step movement from YP₁ to YP₃ via YP₂ should be contrasted with one step movement from YP₁ to YP₃. Such one step movement would violate the locality condition because YP₁, as discussed, is dominated_{Barriers} by XP and YP₃ is not dominated by any segment of it. This is how escape hatches are created by adjunction in Barriers.

Unfortunately, all three key ingredients used in the account are quite stipulative. The notion of ‘segment’ of a category is stipulative and unsatisfactory (see Chametzky 1996, 2000 for extensive discussion).¹⁷ The notion of dominance_{Barriers} is a stipulation without independent support. Dominance is a natural relation in hierarchical structures, but the notion of dominance_{Barriers} relies in addition to the hierarchical structure on highly specific notions of projection and bar levels. Finally, the locality condition, which mixes the two notions of dominance and dominance_{Barriers} in an arbitrary way, is not particularly natural or satisfying. In a word, the escape hatch property is stipulated rather than

¹⁷ The category-segment distinction was introduced by May 1985 to account for certain quantifier scope ambiguities. Chametzky 1996, 2000 argues that even the original use of the distinction did not lead to true explanation since the relevant effect (free scope relations) does not follow in any way from the definition of a segment. Chomsky’s use in Barriers of the distinction is basically unrelated to May’s use of it. Finally Kayne 1994 adopts the distinction terminologically. However, his LCA rules out the kinds of structures that May and Chomsky used the distinction for. The category segment distinction has not given rise to a coherent or interesting body of doctrine.

derived. We do not gain an understanding of why escape hatches exist, why the XP adjoined position and not some other position is an escape hatch, etc.

The same problem arises for other formulations as well. Consider van Riemsdijk’s (1978a:169) Head Constraint (40) and Chomsky’s (2000:108) version of it, the Phase Impenetrability Condition (41). Contrary to current terminology, I will use van Riemsdijk’s name, ‘Head Constraint’, for the condition rather than Chomsky’s, ‘Phase Impenetrability Condition’.

- (40) *Head Constraint*
No rule may involve X_i (X_j) and Y in the structure ...X_i ... [α ... Y ...] ... X_j ... if Y is c-commanded by the head of α.
α ranges over V''', N''', A''', P'''
- (41) *Phase Impenetrability Condition*
In phase α with head H, the domain of H is not accessible to operations outside α, only H and its edge are accessible to such operations.

The only difference between the two conditions is that van Riemsdijk includes a list of categories to which the Head Constraint applies, whereas Chomsky does not consider this list (i.e. the specification of which XPs are phases) part of the definition of the Phase Impenetrability Condition. Otherwise, the two conditions are identical: No relation can be established between an element outside of (phase) XP and the complement of the phase head or anything properly contained in the complement of the phase head. These statements are less obfuscatory than the formulation of the escape hatch property given in Barriers, but they are in no way less stipulative.

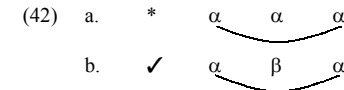
What we would like to do, of course, is to derive this strange condition from independently motivated principles. The following paragraphs offer what seems to me to be a first step in that direction.

The final solution I offer has the following form. Locality is taken to regulate attraction relations (Attract Closest) rather than chain links. The fact that specifiers are escape hatches follows from this assumption. Secondly I will assume that phase heads always bear all features that can enter into an attraction relation. This is to make sure that no movement relation can cross a phase head without passing through the [Spec, phase] position (2.2.2.1). Since in many cases these features on the phase heads are not checked, a theory of default feature values becomes necessary (2.2.2.2). Since such a theory is needed anyway, this is not a distressing property of the current proposal. Finally, a distinction needs to be drawn empirically between the phase head (X°) and the phase (XP). X° is a universal intervener, but XP is not. This leads me to postulate that the features of X° relevant for intervention are not projected to XP. This is an unsatisfying assumption, to be sure, but apparently one that is necessary also on Chomsky's approach to phases.

2.2.2.1 Escape Hatches and Attract Closest

Subadjacency, Barriers, or the Head Constraint are all statements about the locality of movement. There is one locality condition on movement that is fairly well understood and natural, Relativized Minimality (Rizzi 1990). I would like to suggest that we can gain some insight into subadjacency by trying to derive it from Relativized Minimality.

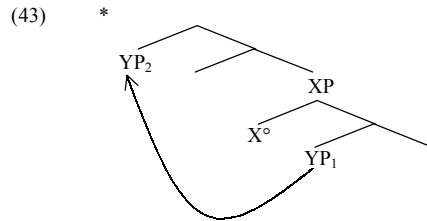
Relativized Minimality is a ban against crossing items. No movement of an item that belongs to a certain class α may cross over another item that belongs to the same class. Items that belong to some other class β may freely be crossed. This is schematized in (42). In Rizzi's original formulation the classes were positional classes (A-positions, A'-positions, heads), i.e. relational properties, but it has become standard practice to view them as intrinsic (i.e. projected) properties of syntactic objects, i.e. features.



If Relativized Minimality is to be relevant to subadjacency, i.e. if non-subadjacent movement is to be ruled out by Relativized Minimality, then there must be an intervener along the path of movement. The most plausible candidate for an intervener is the phase head. On the view just sketched, non-subadjacent movement of YP in (43) is illicit, because it crosses the intervener X° . This entails that there is a class α such that both YP (or Y°) and X° are members of that class – or, in other words, X° and YP have an α -feature. So far this is of course not a novel or idiosyncratic assumption.

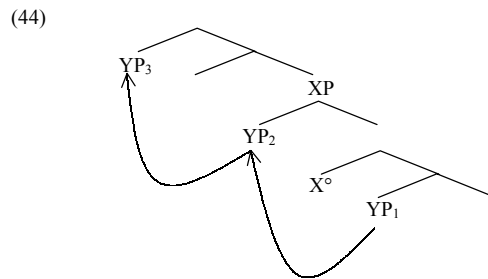
If we now take YP to be a *wh*-phrase for example and X° to be an intermediate complementizer, then α would be a *wh*-feature on C° . For C° to intervene, this feature must be present on C° whether C° is interrogative or not. I return to the questions raised by this assumption. On the standard Minimalist view of successive cyclic movement such a *wh*-feature is (at least optionally) present on C° anyway. My own proposal is that the relevant feature is not only optionally present, but that its presence is obligatory. The

feature is always present on every C° . Here I do differ from standard assumptions. Some implications of this view will be discussed in section 2.2.2.2.



Intervention is defined in terms of c-command and the licit derivation is given in

(44). At first blush it doesn't look as though this helps at all. The second step of movement (YP_2 to YP_3) is of course fine, but doesn't the initial movement step cross X° and shouldn't it, therefore, be ruled out on a par with (43)? In other words, do we really want to say that X° intervenes between YP_1 and YP_2 in (43) but not in (44)? Don't these structures simply recreate the Barriers problem, except that locality is now understood in terms of c-command rather than dominance? As we will see momentarily, they don't.



I believe that the outlook is a little less bleak than (44) would suggest. On the last two pages I have implicitly assumed that Relativized Minimality is a condition on chains, but this is not necessarily true. In fact Chomsky (1995b:297) offers the condition Attract

Closest as a reformulation of Relativized Minimality as a condition on attraction rather than a condition on movement chains. The details of the formulation of Chomsky's condition need not concern us here. I will use the simplified versions of Attract Closest given in (45) here and throughout.

- (45) *Attract Closest*
If a head K attracts feature F on X , X a feature bearer of F , no constituent that bears F is closer to K than X is.
- (46) *Closeness* (to be revised)
 Y is closer to K than X is if K c-commands Y and Y c-commands X .

The notion of closeness employed in (45) is to be understood as in (46). For reasons that will become clear immediately, I assume that heads act as attractors and phrases (i.e. maximal projections) act as attractees. The formulation in (45), taken almost verbatim from Pesetsky and Torrego 2001:362, is simplified from Chomsky's (1995b:194 ex. 184) original formulation, which refers to sublabels (i.e. features of complex head clusters) rather than heads as feature checkers. This complication only arises if head movement is a syntactic process and if various other assumptions are made. Below I reject head movement as a syntactic process, hence Chomsky's notion of sublabel does not play a role. It is replaced by the notion of feature bearer defined in (47).

- (47) *Feature Bearer*
 β is a feature bearer of feature F iff β is a constituent and the label of β contains F .

The formulation in (46), again taken from Pesetsky and Torrego 2001:362, is a simplification of Chomsky's (1995b:299 ex. 287) original formulation as well.

Chomsky's formulation allows the closest c-command requirement to be flouted

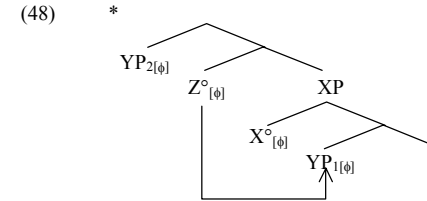
sometimes (equidistance).¹⁸ The stipulative nature of equidistance had long been noticed (Koizumi 1995). Vukić (2003) points out several empirical shortcomings, too, and shows how a simple version of Attract Closest strictly in terms of c-command gives better empirical results.¹⁹

Returning to the main thread, now consider the structure in (48). Structure (48) is just like (43), except that the head Z° , which ultimately attracts YP is shown. Suppose that Z° is trying to attract a feature ϕ and that both X° and YP have this feature. In the example discussed above X° would be the complementizer, which, by assumption, always contains a *wh*-feature. The following diagrams are to be read as follows. Square arrows represent attraction relations (downward); curved arrows stand for movement relations (upward).

¹⁸ In Chomsky 1995c equidistance was necessary to allow the direct object to cross of the subject on its way to AgrO and to allow the subject to cross over the moved object on its way to AgrS. The equidistance solution allows order preservation to emerge in this particular case ($SU > OB > t_{SU} > t_{OB}$), but it is not equipped to deal with order preservation more generally (e.g. $SU > DO > IO > t_{SU} > t_{DO} > t_{IO}$, etc.). A general solution to the order preservation problem should yield the empirical results of equidistance (Starke 2001; Williams 2002). A derivational formulation within the set of assumptions adopted here is still pending.

¹⁹ The formulation in (46) gives rise to an interesting case. If β and γ , two feature bearers of F , symmetrically c-command each other, then both β and γ are closer to a higher attractor α than the other. It follows from (45) that neither of them will be attracted (see also McGinnis 1998 and Pesetsky and Torrego 2001 for various consequences of this idea). This feature of the definition of closeness is behind the theoretical reduction of the generalization that no phase head may be stranded. The generalization was introduced in chapter 1 and will be defended at length empirically in chapters 3 and 4. The theoretical deduction is given in section 2.4 of this chapter. It would be easy, of course, to say that in such cases both β and γ are closest to α and that

The violation of locality, i.e. Attract Closest, occurs in (48) because X° intervenes between Z° and YP. Notice that XP does not intervene between Z° and X° and YP₁. Intervention requires c-command. Since XP does not c-command either X° or YP₁ it cannot intervene.

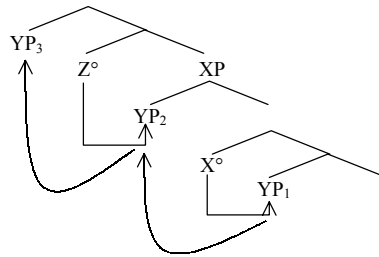


The corresponding adjusted structure for (44) is given in (49). We saw above in (44) how the movement chain (YP₁, YP₂) crosses over X° in apparent violation of Relativized Minimality. But notice that the attraction relations in (49) conform with Attract Closest. X° can attract YP₁ because there is no relevant c-commanding intervener. As a consequence of this attraction relation YP₁ moves to [Spec, XP]. As noted, the fact that the chain formed by YP₁ and YP₂ violates the non-crossing constraint is irrelevant since locality is taken not to be a constraint on chains but on attraction. Similarly Z° can attract YP₂ since Z° c-commands YP₂ and X° does not intervene with attraction of YP₂. The resulting chain (YP₂, YP₃) again irrelevantly violates no-crossing.²⁰

either one can be attracted. I refrain from doing this because of the Stranding Generalization.

²⁰ There is one wrinkle here. The sister of YP₂ has the relevant feature α in its label. It should therefore be eligible for attraction. I claimed above that if two attractees

(49)



If this is so and the relations in (49) comply with Attract Closest, then the fact that the movement chain (YP₁, YP₂) in (44) violates Relativized Minimality is irrelevant.²¹

If it is assumed that a feature *F* which triggers movement belongs to the same class as the corresponding (attracted) feature on the moved item, then this shift from a definition of Relativized Minimality over chains to a definition of Relativized Minimality over attraction relations is necessary anyway. A glance at the chain (YP₁, YP₂)

c-command each other, neither of the two counts as a closest attractee for the purposes of Attract Closest (45). As a temporary solution, simply assume that intermediate projections do not participate in c-command relations (Chomsky 1995b; Epstein and Seely 1999; Kayne 1994). I return to the issue below in section 2.3.2.

²¹ Epstein (2001) offers the same kind of argument with respect to head movement. A well known problem with head movement is that the head of a head chain does not c-command its tail – at least as long as we assume a definition of c-command in terms of the first branching node. Head chains thus violate the Proper Binding Condition going back to Fiengo 1974, which demands on a current formulation that every trace be c-commanded by its antecedent. Epstein claims that this is irrelevant, since the Proper Binding Condition, in the cases where it does hold, is epiphenomenal. What matters is not that the head of a chain c-command its tail, but that the attractor c-command the attractee before movement takes place. And this condition is fulfilled by head movement.

I believe that this kind of consideration is natural in a theory where syntactic representations are derived through the successive applications of structure building operations. My own views on head movement are discussed briefly in section 2.4 of this chapter and in some detail in chapter 5.

immediately reveals this. If locality is feature based (rather than position based as in Rizzi 1990) we have to formulate it in terms of the relation between attractor and attractee, since otherwise no movement can ever comply with it – or reintroduce unwanted concepts such as equidistance.

This means that the escape hatch property itself, i.e. the fact that the edge, the specifier of the phase, is able to be extracted but not the complement of the phase head or anything contained within the complement, this property follows straightforwardly from Attract Closest because the complement but not the specifier (or an adjunct) is c-commanded by the head. The properties of the phase heads themselves, to which I now turn, are less straightforward.

2.2.2.2 Phase Heads and Default Values

X° must share a feature α with YP in order to be able to attract YP in (49) above. We must make sure that [Spec, XP] in (49) is not eligible as the final landing site for YP. Furthermore, X° must bear that same feature in (48) in order to be able to intervene with YP. X° must always bear the relevant feature, whether movement through its specifier takes place or not. In contrast to Chomsky 2000, 2001a, b I cannot allow for the existence of a variant of X° that does not have the feature α , otherwise, in the absence of the Head Constraint, (48) would be grammatical after all. If X° in (48) didn't have to carry the α -feature, movement of YP should freely be allowed to cross X°. Thus α must be an intrinsic feature of X°. This intrinsic feature α may attract YP, but it need not attract anything at all.

To see this, consider a concrete case. Suppose X° is the complementizer C° . If *wh*-movement is to proceed from out of the *c*-command domain of *that*, it has to stop by in [Spec, CP] as in (50). Some feature of C° , call it [*clause type*] gets checked this way.^{22, 23} Nothing goes wrong, however, if there is no *wh*-element within the *c*-command domain of *that*, i.e. if *that* doesn't attract any *wh*-element (51) or if there is a *wh*-element but it remains in situ (52). In fact aborted movement to an intermediate position is ungrammatical (53).

- (50) ✓What did you claim [_{CP} ~~what~~ that [Peter bought ~~what~~]]?
 (51) ✓You claimed that Peter bought a chair.
 (52) ✓Who claimed that Peter bought what?
 (53) *Who claimed what (that) Peter bought ~~what~~?

From the perspective of the embedded complementizer, the two cases in (51) and (52) are not distinct.²⁴ If we assume that the [*clause type*] feature can be checked only under a local relation (for present purposes we can assume this relation to be the Spec-Head relation, for a slight reformulation see section 2.4.1), then the [*clause type*] feature is checked neither in (51) nor in (52), but it is checked in (50) and (53). This describes a basic opposition between (51) and (52) on the one hand and (50) and (53) on the other

²² My feature [*clause type*] is usually called a *wh*-feature. I do not follow this convention, but what I say can be reformulated in terms of a binary [+/- *wh*] if this is deemed more desirable. I do not know if my proposal can be reformulated in terms of a privative *wh*-feature.

²³ The [*clause type*] feature on the embedded C° gets checked in (50). Below I will distinguish feature checking from feature valuation. This distinction will make sure that the [*clause type*] feature on the embedded C° is not valued interrogative.

²⁴ I know of no cases where covert *wh*-movement would give rise to intermediate complementizer agreement.

hand. There is a further opposition between (50) and (53). In (50) the *wh*-phrase moves out of the local domain of the embedded complementizer whereas in (53) it remains there.²⁵

If only uninterpretable features are able to attract, then [*clause type*] must be uninterpretable on C° . But if all uninterpretable features must be checked/valued for convergence and [*clause type*] need not be checked/valued, then [*clause type*] cannot possibly be both uninterpretable and obligatorily present because it remains unchecked in (51) and (52). Finally, [*clause type*] on C° is not itself attractable, i.e. clauses do not automatically count as *wh*-phrases.²⁶ This shows that within the framework of assumption of Chomsky 1995c, 2000, 2001a, b concerning the interpretability of features and feature checking, [*clause type*] has bizarre and, in fact, contradictory properties. But these properties not altogether unprecedented. The odd behavior of [*clause type*] is simply the behavior of an uninterpretable feature with a default value. The point of a default value for an uninterpretable feature is that that feature may remain unchecked/unvalued. Rather than leading to derivational crash, the feature is assigned its particular default value.

I assume that there are uninterpretable features that do and some that do not have a default value. The [*clause type*] feature on C° has a default value, namely 'declarative'.

²⁵ This remark presupposes a view where *wh*-phrases that are not located in [Spec, CP] overtly do not undergo covert *wh*-movement but are interpreted in situ through some other mechanism like unselective binding.

²⁶ Cases of clausal pied-piping must presumably be treated differently since the clause to be pied-piped must contain a *wh*-phrase to be eligible for pied-piping. The [*clause type*] feature on C° alone is insufficient to trigger clausal pied-piping.

The Case feature on T° on the other hand has no default; it must be checked and valued.

This gives rise to EPP effects in tensed clauses (see Bošković 2002b for extensive discussion of the EPP). Assume furthermore that feature checking may require (weakly) closest c-command or (strongly) a spec-head relation (see section 2.4.1 of this chapter for discussion and the claim that the Spec-Head configuration is a stronger sub-case of closest c-command).

These two distinctions give rise to four classes of uninterpretable features: (i) features that have a default value and require (weakly) closest c-command for checking (ϕ -features on T° in Russian might be an example. There appears to be default agreement in the language, but if a nominative subject is present, agreement must take place. Yet this happens without obligatory movement to [Spec, TP].); (ii) features that have a default value and require (strongly) a Spec-Head relation for checking (e.g. [*clause type*] in languages with obligatory *wh*-movement like English); (iii) features that do not have a default value and require (weakly) closest c-command for checking (no example given since cases (i) and (iii) are very difficult to distinguish empirically); (iv) features that do not have a default value and require (strongly) a Spec-Head relation for checking (e.g. Case on (finite) T° , i.e. the EPP).

In this light let's consider the facts in (50) through (53) again. The issue is what happens with [*clause type*] on C° in case no movement through [Spec, CP] happens, in case movement passes through [Spec, CP] without that [Spec, CP] being the final landing site and what happens if, indeed [Spec, CP] is the final landing site.

- (50) ✓What did you claim [CP ~~what~~ that [Peter bought ~~what~~]]?
 (51) ✓You claimed that Peter bought a chair.

- (52) ✓Who claimed that Peter bought what?
 (53) *Who claimed what (that) Peter bought ~~what~~?

In the previous paragraph I assumed that the feature [*clause type*] has as its default value 'declarative'. Since [*clause type*] is not checked on the embedded complementizer in (51) and (52), it is assigned the default value 'declarative' in these cases. In (53), the feature is checked and is assigned the value 'interrogative'. This violates the subcategorization of the verb *claim*, accounting for the deviant status of (53). Example (50) is still unaccounted for. As it stands, we would have to conclude that the [*clause type*] feature of the embedded complementizer is checked and valued as 'interrogative'. Example (50) should be just as deviant as (53). This necessitates the distinction that I have already slipped into the discussion, that between checking and valuation. The [*clause type*] feature on the embedded complementizer is apparently checked in a local relation with a *wh*-phrase, but this does not entail that the feature is assigned the value 'interrogative'. Apparently the feature [*clause type*] is only valued 'interrogative' if the *wh*-phrase remains in [Spec, CP] and doesn't move on further.

Observe, that the *wh*-phrase is in a Spec-Head configuration in (50) both with the embedded and with the matrix complementizer, but the matrix complementizer is closer to the *wh*-phrase than the embedded complementizer is. This follows from our definition of closeness. Repeated here for convenience.

- (46) *Closeness* (to be revised)
 Y is closer to K than X is if K c-commands Y and Y c-commands X.

The matrix complementizer (Y) is closer to the *wh*-phrase (K) than the embedded complementizer (X) is because K c-commands Y and Y c-commands X. We can then

conjecture that feature valuation requires closest checking (relativized to particular features or feature classes, of course). Example (50) now comes out as desired. The copy of *what* in the embedded [Spec, CP] cannot value the [*clause type*] feature, which therefore defaults to ‘declarative’. Notice that checking is now construed derivationally as a local relation, whereas feature valuation requires the full representation. At the derivational step where the *wh*-phrase in (50) moves to what ends up being the embedded [Spec, CP], nothing distinguishes the structure in (50) from that in (53). The distinction can only be drawn later based on the final representation.

Apart from accounting for examples like (50), the suggested mechanism of feature valuation also accounts automatically for examples like (54), which were problematic for Chomsky 1995c. The verb *wonder* subcategorizes for an interrogative, but the complementizer in (54) defaults to ‘declarative’ since it is not the closest checker of *what*.

(54) *What does John wonder ~~what~~ Mary bought ~~what~~

Example (54) is an example of movement from a *wh*-position to a *wh*-position. It is ruled out because only the highest landing site is valued ‘interrogative’. The lower checkers of the *wh*-feature are assigned default values. The same reasoning applied here to example (54) can be used to rule out movement from Case position to Case position, by assuming that the lower of the putative Case assigners will be valued the wrong way, i.e. will default to the wrong value. Similarly for other case of illicit movement from X-position to X-position. In other words, although the distinction between checking and valuation I make seems slightly artificial and ad hoc, it captures effects that fall under

independent stipulations in other theories. The current formulation unifies (50), (53) and (54) using a familiar concept of closeness. The present theory, then has some advantages compared to more standard theories that do not distinguish checking from valuation in the way suggested here but that need to formulate a ban for (54) separately (and sometimes have great difficulty ruling out (53)).

The crucial point in all this is that feature valuation is defined globally over output representations. Since both the interpretation of a feature (‘interrogative’ vs. ‘declarative’) and the morphological shape of its bearer (*that* or \emptyset if the feature is valued as ‘declarative’ and \emptyset if it is valued as ‘interrogative’) depend on the value of the feature and since feature valuation requires the full representation to be available. It follows that spell out to the interfaces cannot be done on-line – at least if the generalization from *wh*-features to Case-features and beyond is accepted.²⁷

We have arrived at an important difference between Chomsky’s theory of phases and the present theory. The present theory offers a - possibly relativized - theory of escape hatches. It emulates the Head Constraint relying on Attract Closest together with a theory of feature valuation and default values. However, this theory is incompatible with the model of cyclic spell out. All putative properties of phases that follow from cyclic spell out cannot be linked to the concept of phase in the present theory. I merely note this

²⁷ Some qualifications apply. One could devise a system such as Chomsky 2001b where a strong phase is spelled out at the level of the next higher strong phase. In such a model the larger phase would presumably contain all the relevant information. I don’t see what is gained by doing things that way, though.

point here since discussion would distract from the main line of argument at this point. I discuss these issues in some detail in subsection 2.2.3, though.

To flesh out the proposal concerning default values for uninterpretable features further, consider the ϕ -features on T° . Chomsky takes them to be uninterpretable, but they are a prime candidate for uninterpretable features with a default value. In many languages, if a clause has a non-nominative subject, the ϕ -features of the verb cannot be checked. But this does not make the derivation crash, rather, the features are set to a default value (typically 3rd person, singular, neuter). The ϕ -features on T° are not attractable, i.e. TP is never attracted by a ϕ -attractor for checking. The parallelism with the [*clause type*] feature on C° is complete.

I leave aside the last issue, that of attractability. ϕ -features on T° give rise to the same contradictions as [*clause type*] on C° did. The features are present. On standard assumptions, they are uninterpretable; they cannot on those same assumptions be uninterpretable since they need not necessarily be checked. The fact that putative uninterpretable features cannot have default values in Chomsky's theory points to a weakness of that theory.²⁸

²⁸ Note that I have given the feature [*clause type*] on C° a clear interpretation. It is not entirely clear to me what the fate of the always elusive distinction between interpretable and uninterpretable features is under my proposal.

In any case, assume that the ϕ -features on T° require closest c-command for checking²⁹ and that checking is an automatic consequence of this structural relation (much like matching is for Chomsky 2001b). Then the ϕ -features on T° will be checked and valued iff there is a local ϕ -bearer. Otherwise they will be assigned the default value.

Obviously, these issues are complicated and deserve further attention. I believe though that the ideas sketched here may shed some light on the nature of default values.

If this speculation about default values is on the right track, then the Head Constraint need not be stated as a constraint independent of Attract Closest. Consider again examples (50) repeated here for reference and its counterpart with one-fell swoop movement (55).

- (50) ✓What did you claim [CP ~~what~~ that [Peter bought ~~what~~]]?
 (55) *What did you claim [CP that [Peter bought ~~what~~]]?

Example (50) is ruled in, because every instance of attraction accords with attract closest. (55) on the other hand is ruled out, because the [*clause type*]-feature (or the [*wh*]-feature, see fn. (22)) on the embedded C° intervenes. The representation in (50) is simplified in that possible intermediate traces in [Spec, vP] position are not shown.

If the Head Constraint is on the right track and CP intervenes with ANY type of movement that does not pass through [Spec, CP], the obvious conclusion is that C° bears

²⁹ We pick checking under closest c-command rather than spec-head agreement, because a spec-head relation is not required for checking as expletive constructions indicate. Thus in (i) matrix T° agrees with *three men* in ϕ -features without being in a spec-head configuration.

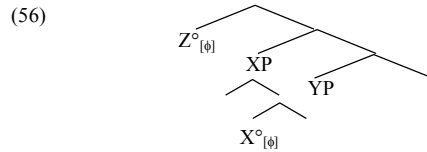
- (i) There seem to be three men in the room.

instances of all features or feature classes. This conclusion is forced in my system, but it is not unique to it. This view is implicit also in phase theory, where a phase head must optionally contain any number of different features to enable the phase head to trigger movement through its specifier position.

Thus far, the reduction of the Head Constraint to Attract Closest is promising.

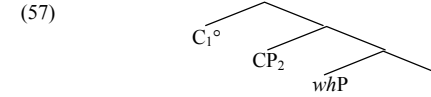
2.2.2.3 Differential intervention of X° and XP

There is a problem, though. I am claiming that phase head X° in (48) and (49) has feature $[\phi]$ which acts as an intervener with YP-movement in (48) and attracts YP in (49). If X° has feature $[\phi]$, then XP should have it, too. Consider the structure in (56). Suppose as before that Z° bears feature $[\phi]$ and that YP bears feature $[\phi]$ and that the two could enter into a checking relation. Suppose that XP is a phase. If XP bears $[\phi]$, then it will follow from the theory developed in this section that the presence of XP should block movement of YP in (56) on a par with the blocking by $X^\circ_{[\phi]}$ seen in (48).



Suppose that XP is a place-holder for a phase, i.e. XP is a CP or a vP, possibly a DP. I will argue later that PPs are also phases in the intended sense in most languages. In other words, I will claim that *wh*-movement passes through [Spec, CP], [Spec, vP], [Spec, PP], and possibly [Spec, DP]. If (56) were ruled out, as my theory currently

predicts, then every CP, vP, PP, and DP in a specifier position along the path of movement should lead to an intervention effect. To make the example concrete, assume that XP in (56) is a CP which, therefore, contains a [*clause type*] feature. Suppose Z° is also a complementizer and YP is a *wh*-phrase. This is shown in (57).



C_2° bears a [*clause type*] feature. I argued on the basis of (55) that this [*clause type*] feature is an intervener for *wh*-movement. If C_2° bears that feature, it is presumably percolated to CP_2 as well and this creates a problem. If CP_2 bears a *wh*-feature, then this feature should intervene with *wh*-movement of the *wh*-phrase.³⁰ In fact if the suggestion made earlier is correct, that C° (and other phase heads) are universal feature bearers, then they should intervene with any and all movement past them. In other words every CP, vP, DP, and PP in a specifier position should intervene with any movement. The most trivial examples show that this is false. One case is example (58), which is wrongly predicted to be ungrammatical.

(58) What does [_{DP} the man] like ~~what~~?

This is clearly an unwanted conclusion. The intuition we would like to capture is that the head of a phase, X° , can give rise to an intervention effect as in (48) when

³⁰ Howard Lasnik (p.c.) suggests that the prediction made here might actually be correct in the sense that clausal subjects and objects do seem to interfere with *wh*-movement past them. For the more general case, though, the prediction is certainly false.

movement occurs out of its own domain, but the phase itself, XP in (56), does not give rise to such an effect. In other words, while X° is a feature bearer of the α -feature driving movement, XP must be prevented from being a feature bearer of the α -feature. This means that the α -feature is not found in the label of XP. We can achieve this by stipulating that the α -feature is unable to project – a stipulation that might be needed under Chomsky's phase theory, too, for otherwise the features that are optional on the head of the phase would also systematically intervene whenever they are present.

We can ask whether there are other cases where a feature that is present on a head does not project to the level of the phrase. As is well known verbs do not agree in ϕ -features with quirky subjects. Masashi Nomura (p.c.) suggests that this can be captured if we make the following two assumptions: (i) agreement is a relation between a head, T° , and a maximal projection; and (ii) ϕ -features on quirky D° do not project up to DP. Given this suggestion even the property that certain features are blocked from projecting to the label might not be without precedent in the theory.

As far as I can see, the stipulation of differential projection of features poses the biggest challenge for the attempt made here to unify the Head Constraint with Attract Closest. Again I hope that a general theory of feature percolation, which is needed to solve problems having to do with pied piping and which is not currently available, will solve this problem.

2.2.2.4 Summary

The considerations of this subsection make clear how subadjacency must work under the current approach. I am suggesting that certain (classes of) heads like C° , v° , maybe D° , and, as we will see later also, P° (this is of course van Riemsdijk's list from (40)), and possibly others bear inherent features that can (but need not) attract *wh*- and other phrases.

In every syntactic environment that requires the presence of one of these heads movement of a *wh*-phrase has to pass through the specifier position of that head. If that position is skipped by movement, Attract Closest is violated. Wherever there is a C° , v° , P° , or D° , movement must be successive cyclic.³¹

³¹ It is inviting at this point to explore the idea that T° is for Case relations what C° is for A'-relations – in other words we could pursue the idea that C° is not a universal feature bearer but only a universal A'-feature bearer whereas T° could be treated as the corresponding universal A-feature bearer. This would immediately explain the ban against superraising; it would force A-movement to be successive cyclic (for an overview of arguments to this effect see Bošković 2002b) because to get out of the local domain of T° it would need to go through the [Spec, TP]-escape hatch position; finally if the relevant feature on non-finite T° is sufficiently similar to the [*clause type*] feature on C° it would not necessarily need to be checked and could be assigned the default value, i.e. we could do away with the EPP requirement. This idea could be used to explain the observation that expletives do not seem to undergo raising but are base merged in the top [Spec, TP] position according to Bošković 2002b. On the other hand many questions arise pertaining to possible long distance Case relations. Potentially problematic cases can be found in Icelandic, where raising verbs with an experiencer block raising but the subject of the embedded infinitival can still be marked nominative. We can solve this by assuming that the nominative on the embedded subject is assigned by the embedded T° -head (cf. Sigurðsson 1996) and that agreement is strictly local and that the matrix T° does not agree with the associate of the expletive directly, but rather with the embedded T° ,

2.2.3 Phases: Cyclic Spell Out vs. the Present View

In this subsection I highlight some differences between my own version of phase theory and Chomsky's and defend the choice. Suppose the thesis is correct that the locality effects of the Head Constraint, Chomsky's Phase Impenetrability Condition, can be unified with Attract Closest along the lines suggested in the previous subsection and that the associated theory of features is correct. As noted, this theory requires access to the full representation and is incompatible with cyclic spell out – at least if phases are spelled out right after completion. The Head Constraint only takes care of the bounding effects but not the other properties ascribed to phases by Chomsky.

In this subsection I investigate Chomsky's (2000) original empirical argument for phases. The argument comes from expletive constructions. Other putative properties of phases (operation of the Head Constraint, phonological isolability, spell out to PF and LF, being semantically propositional) do not in fact cluster empirically (Bošković 2002b; in press; Epstein and Seely 1999, 2002; Matushansky 2003; Rezac ??? ref???) and are connected theoretically at best in a tenuous way.

Before turning to expletive constructions, let me briefly substantiate the claim that the putative phase properties do not cluster and are not derived theoretically. The first assumption about phases is that the complements of phase heads are spelled out to the

which, in turn, agrees with the associate (see for Hindi overt agreement facts that suggest this kind of treatment Mahajan 1990). I leave this important issue for further research.

interfaces as soon as the phase is completed.³² If, as claimed, C° is a phase head, then TP gets spelled out to the interfaces. The main consequence of this assumption for syntax proper is the Phase Impenetrability Condition, i.e. the condition that the complement of the phase head and whatever it contains are no longer accessible to the syntactic computation. The main effect of the Phase Impenetrability Condition is to derive subjacency. If the ideas developed in this chapter so far are on the right track, then cyclic spell out is not needed to derive subjacency and, moreover, subjacency can be partially unified with Attract Closest/Relativized Minimality which, if successful, is clearly an advantage.

The second claim about phases in Chomsky's sense is that phases but not other categories are phonetically isolable. The presumed theoretical connection would have to lie in the fact that phonetically isolable objects are spelled out to PF together. First of all, it is not true that only phases are phonetically isolable. Very clear evidence for both CP and TP being isolable comes, for example, from Right Node Raising examples like (59) and (60) (this argument is due to Bošković 2002b:182 fn. 118 the examples are taken from Postal 1998). Right Node Raising typically requires a pause before the Right Node Raised material. If there is a clear case of phonetic isolability, then this is surely one.

- (59) ✓John believes and Peter claims – that Mary will get a job.
 (60) a. ✓John believes that and Peter claims that – Mary will get a job
 b. ✓I know when but I don't know where – Amanda met Steve

³² There is a stipulation here to the effect that the complement of the phase head is spelled out and not some other definable chunk. Let's set this issue aside.

Second of all, the theoretical motivation for the claim that phases ought to be phonetically isolable is a non-sequitur. If what is spelled out together should be isolable and what is not spelled out together should not be isolable, then phases should not be isolable, since they are not spelled out together. The complement of the phase head and the rest of the phase (the edge) are crucially spelled out at different times. The complement of the head of a phase is spelled out as one unit and the rest is spelled out as a separate unit at the level of the next higher phase, i.e. TPs are the expected case of phonetic isolability and CPs the surprising case. Since neither the generalization nor its theoretical motivation hold water, I believe that we can safely ignore the argument from phonetic isolability for phases.

The next property Chomsky attributes to phases is that all and only phases are semantically propositional objects. The idea here is that propositions are semantically complete entities whereas non-propositions are not. This property of phases is therefore the semantic equivalent to phonetic isolability. But just as in the case of phonetic isolability, the theoretical claim is limping. What the semantic interface is presented with are not the complete propositional objects vP and CP themselves but rather their proper parts TP and VP. The purported generalization follows in no way from the theory. Furthermore, the generalization itself is dubious unless a coherent argument is offered why vP and CP are more propositional than TP. To illustrate this point consider the examples in (61) from (Bošković 2002b:182 fn. 118). It is at best unclear in which sense the embedded TP in (61a) is more propositional than the CPs in (61b) and (61c).

- (61) a. There seems [_{TP}to have arrived someone]
 b. It seems [_{CP} there had arrived someone]
 c. It seems [_{CP} someone had arrived]

Moreover, as pointed out by Epstein and Seely (2002:78), when vP and CP are extracted from, the resulting object is an open proposition.

To be fair, the objections that there is a mismatch between what is spelled out (TP and VP) and what is a phase (CP and vP) are met in Chomsky (2001b) who assumes that a phase “PH₁ is interpreted/evaluated at the level of the next higher phase PH₂” (p. 13). This may solve the problem that cyclic spell out picks out the wrong categories, but it does not solve the problem that the generalizations (differential phonetic isolability, propositionality) do not work to begin with. Furthermore, if the idea that a phase is interpreted/evaluated at the next higher phase is correct, then the Phase Impenetrability Condition/Head Constraint is left dangling without independent theoretical motivation since the original motivation that the complement of the phase head is inaccessible because it is spelled out is gone. The Head Constraint then needs to be stipulated independently to enforce subadjacency. There is no way at present to derive from a uniform non-stipulative mechanism both the phase impenetrability condition (making TP impenetrable) and the purported generalization that CPs but not TPs are relevant chunks for phonology and for semantics.

It is sometimes assumed that phases are syntactically mobile but that non-phases are not syntactically mobile (e.g. Svenonius 2003). It is unclear what the theoretical connection between mobility and phasehood is, though. In any case, chapter 3 of this thesis argues that TPs and VPs alongside CPs and vPs are syntactically mobile. If the

arguments developed in chapter 3 are correct, then the mobility generalization, too, fails. Consider also the following examples (62) exemplifying Left Branch Extraction in Russian. If everything that moves is a phase, then almost every phrase must be considered a phase (Matushansky 2003 takes this argument to its logical conclusion), but this takes all empirical bite out of the phase theory. Based on these considerations and the evidence to be presented in chapter 3, I discard the generalization that all and only phases are syntactically mobile.

- (62) a. ✓Ivan kupil novuju mašinu.
 Ivan bought new_{acc.sg.f} car_{acc.sg.f}
 b. ✓Novuju mašinu Ivan kupil.
 new_{acc.sg.f} car_{acc.sg.f} Ivan bought
 c. ✓Novuju Ivan kupil mašinu.
 new_{acc.sg.f} Ivan bought car_{acc.sg.f}
 d. ✓Mašinu Ivan kupil novuju.
 car_{acc.sg.f} Ivan bought new_{acc.sg.f}
 ‘Ivan bought a new car.’

The final argument for phases comes from computational complexity. Chomsky claims that the introduction of phases into the system is conceptually welcome since phases help to reduce computational complexity. Johnson and Lappin (1997) had criticized Chomsky 1993 for suggesting a computationally intractable theory.³³ The thrust of the argument was that the kind of economy consideration entertained in Chomsky 1993 are global in that they require comparison of a large number of complete derivations. To the extent that economy conditions are abandoned altogether (Brody

³³ Some though not all of Johnson and Lappin's (1997) points carry over to Chomsky 1995b. For other much less precise work with the thrust to reduce computational complexity see Collins 1994, 1997; Kitahara 1997 and much related work since.

1995) and to the extent that any remaining economy considerations can be handled locally (Chomsky 2000, 2001a, b; Collins 1997; Frampton and Gutmann 1997 also Frampton and Gutmann 2002), the problem of explosion of computational burden that Johnson and Lappin criticized does not arise. The attempt to formulate local economy conditions is independent of the phase based approach to computation.

The remaining complexity considerations that Chomsky (2000; 2001a; 2001b) alludes to without making them precise are based on search space. Such considerations are of a very different nature than the problem that Johnson and Lappin were talking about. Johnson and Lappin were concerned with cases where computational complexity grows exponentially or worse. Search through finite objects such as phrase markers, numerations, and computational workspaces on the other hand is computationally simple. Complexity grows at worst polynomially. No serious argument from computational complexity can be derived here on purely conceptual grounds. Searching seven subconstituents is not significantly simpler in the sense of complexity theory than searching fifteen, five hundred, of fifty thousand subconstituents. The argument for phases cannot be conceptual; it has to be empirical, but the empirical reasons for assuming phase theory we have seen so far fail.

I now turn to expletive constructions, the main empirical reason why Chomsky 2000 introduced phases in the first place. In what follows I show (following a rather well trodden path, see Bošković 2002b section 5.1; Castillo, Drury and Grohmann 1999; Epstein and Seely 1999; Nunes and Uriagereka 2000; Shima 2000; Vukic 2003) that the introduction of phases is not necessary to solve the problem posed by expletive

constructions and that in fact it is theoretically a rather costly solution to a simple problem.

To see why phases were introduced, we need to review Chomsky's (1995b) account of expletive construction. In example (63) *a man* is base merged in the embedded clause. Later *there* is merged directly into the specifier position of the embedded TP from where it raises to the matrix [Spec, TP] position. Still later the ϕ - and Case-features of *a man* raise to matrix T° with which they enter into a checking relation.

- (63) ✓There seems to be a man in the garden.
 (64) *There seems a man to be in the garden.

In (64) *a man* is again base merged in the embedded clause. It then raises to the embedded [Spec, TP]. At the level of the matrix clause *there* is merged into [Spec, TP]. Later the ϕ - and Case-features of *a man* raise to matrix T° , with which they enter into a checking relation.

Chomsky assumes that (63) and (64) use the exact same inventories of lexical items (i.e. they are generated from the same *numerations* in Chomsky's terminology). Structures that are generated from the same numeration enter into a competition with each other as far as economy is concerned and more economical derivations that yield a well-formed output are preferred over less economical derivations that yield a well-formed output. (Ill-formed outputs do not compete.)

Example (63) is grammatical and (64) is not. Following the logic of economy just outlined, the fact that (63) is grammatical shows that its derivation is more economical than that of (64). Specifically, Chomsky suggests that the derivation of (64) is blocked. At the derivational stage where *a man* moves to the embedded [Spec, TP] in (64), the

numeration still contains the expletive *there*. Movement of *a man* in (64) competes with insertion of the expletive into the embedded [Spec, TP] as in (63). Movement is blocked, Chomsky claims, because at any point in the derivation merging an item from the numeration is more economical than moving an element.³⁴

Consider now example (65), which is grammatical with the structure indicated. Crucially in this case, *a man* is able to raise to the embedded [Spec, TP] and later to the matrix [Spec, TP]. The only difference, according to Chomsky, between (63) and (65) is that the numeration that leads to the former contains *there* whereas the numeration that leads to the latter does not.

- (65) ✓[A man] seems [~~a man~~] to be [~~a man~~] in the garden.

Chomsky 2000 takes this account and in particular the Merge over Move preference as his starting point. He goes on to discuss examples similar to the following. (The examples I give are not exactly Chomsky's because I believe that there are a number of interfering factors in the original examples.) Given the preference of Merge over Move, the fact that (66) is grammatical comes as a surprise. Chomsky originally assumes that (66) and (67) are based on the same numeration. The trees are built up from

³⁴ In his writings since Chomsky 1995b, Chomsky has offered several different conceptual arguments why Merge might be less costly than Move. In Chomsky 1995b the claim is that given a particular numeration Merge is forced to apply, because Merge unlike Move depletes the numeration and numerations must be empty before the derivation can be terminated.

For Chomsky 2000 Move is a more complex operation than Merge because Merge is a proper suboperation of Move. He assumes that simple operations preempt complex ones.

The precise details involved here are of no concern to the argument in the text.

the bottom. When derivation reaches the stage in (68), the numeration still contains the elements in (69).

- (66) ✓ There exists evidence that a man was in the garden.
 (67) ✓ Evidence that there was a man in the garden exists.
 (68) $[T_{past}^{\circ} [[a\ man] [BE\ [in\ the\ garden]]]]$
 (69) $NUM = \{there, T_{pres}^{\circ}, that, evidence, exist\}$
 (70) $[there\ [T_{past}^{\circ} [[a\ man] [BE\ [in\ the\ garden]]]]]$
 (71) $[[a\ man] [T_{past}^{\circ} [~~a\ man~~] [BE\ [in\ the\ garden]]]]]$

Given this, there are two ways to continue the derivation. The two possibilities are shown in (70) and (71). In (70) the expletive is inserted into [Spec, TP], in (71) [a man] is moved to [Spec, TP]. According to the Merge over Move preference only one of these two options should be available: (70). Based on (70), only (67) can be generated from the remaining numeration. But this, obviously is an unwanted conclusion, since (66) – which comes from the blocked intermediate step (71) – is perfectly grammatical.

Clearly, one of the assumptions that leads to the prediction that (66) should be ungrammatical must be wrong. As far as I can see, there are three relevant assumptions at play in the examples. They are stated in (72)–(74). As we saw, assuming (73) prompts Chomsky to assume (74) and (74) clashes with (72).

- (72) The numerations underlying (66) and (67) are identical.
 (73) The only differences between the numerations leading to (63) and (65) is that one numeration contains *there* whereas the other does not.
 (74) Merge preempts Move.

Chomsky argues that (73) and (74) are correct. To resolve the contradiction, he then decides to give up (72). There are several ways in which (72) could be incorrect, of course, Chomsky, without discussing the alternatives, claims that numerations have structure. Instead of being unstructured sets of lexical items, numerations now are sets of

sets of lexical items. This can be given equivalent derivational or representational formulations. For reasons of clarity I stick with the representational formulation. The structure of the numeration underlying (66) is (75) and that underlying (67) is (76).

- (75) $NUM_{(66)} = \{\{that, T, a, man, be, in, the, garden\}, \{there, T, exist, proof\}\}$
 (76) $NUM_{(67)} = \{\{that, there, T, a, man, be, in, the, garden\}, \{T, exist, proof\}\}$

In other words, the numeration is divided into subsets (sub-numerations). Each of the sub-numerations determines a ‘phase’ of the derivation. Once the derivation starts to tap into one of the sub-numeration, it has to exhaust it before gaining access to a different sub-numeration; and within each subset the Merge over Move preference (74) is assumed to hold. This allows, correctly, for the derivation of both (66) and (67).

Example (64) must still be ruled out. The idea is that the numeration for (64) cannot be subdivided. If numerations could be subdivided at will, (64) could be generated simply by partitioning the initial numeration for (64) in the way illustrated in (77). Given the absence of an expletive in the sub-numeration for the embedded clause, *a man* could, and in fact would have to, move to the embedded [Spec, TP] position. In the matrix clause *there* would be merged into [Spec, TP]; the end result would be (64).

- (77) $NUM_{(64)} = \{\{there, T_{fin}, seem\}, \{T_{inf}, be, a, man, in, the, garden\}\}$

In other words introducing structure into the numeration and making non-trivial assumptions about how this structure is used (sub-numerations must be exhausted before a new sub-numeration can be tapped into) is insufficient. Further substantive conditions on the well formedness of sub-numeration are required. The one such substantive condition Chomsky offers is that there is a one to one relation between certain heads, phase heads, and sub-numerations.

Many questions arise at this point. Why can the computation draw on only one sub-numeration at any given time? Is there any evidence independent of examples (63), (64), (66), and (67) that the numeration has to be divided in this way? etc. Chomsky's way of rejection assumption (72) complicates the theory in non-trivial ways.

Phases and the associated problems are a high price to pay for the rather simple pattern in (63), (64), (66), and (67). We could easily avoid these troubling questions and unmotivated stipulations. Let's take another look at the assumptions that prompted Chomsky to suggest that the numeration is partitioned. As we saw, the facts in (63), (64), (66), and (67) are incompatible with the conjunction of (72)-(74). Suppose we agree with Chomsky in rejecting (72). Does this mean, we need to add structure to the numeration? No. There are much more pedestrian ways of making the numeration for (66) different from that for example (67).

Suppose the difference lies in the verbs. Adopting the well-known partitive Case hypothesis (Belletti 1988; Lasnik 1995a, b; Bošković 1997b; 2002b; Epstein and Seely 1999; Martin 1992 among others), we could say that the numeration for (66) is (78) and the numeration for (67) is (79), i.e. we give the analysis of these structures suggested in Shima 2000 (see also Bošković 2002b).

(78) NUM₍₆₆₎ = {that, T, a, man, be, in, the, garden, there, T, exist_{partitive}, proof}
 (79) NUM₍₆₇₎ = {that, T, a, man, be_{partitive}, in, the, garden, there, T, exist, proof}

We would then say that V_{partitive} assigns inherent partitive Case. In (66) *be* would assign partitive Case to its object. Since Case marked arguments usually do not undergo further A-movement (setting aside quirky Case in some languages), the object will stay put. The expletive is then required to appear in [Spec, TP] of the matrix clause either by

the EPP or, if it is assumed that *there* can be Case marked – the standard assumption for the proponents of the partitive Case account –, by the requirement that finite T° must assign nominative Case to something. *Be* in the embedded clause in (66) does not assign partitive Case to its object, which therefore has to move to the embedded [Spec, TP] to check Case.

Of course, this entails that we have to slightly change the numeration assumed for the original examples (63) and (64) that motivated the Merge over Move preference to begin with. If we are to allow the expletive in these examples at all, then *be* must be able to assign partitive Case, i.e. it must be *be_{partitive}*. This means that *a man* will be inherently Case marked by *be* and thus immobilized. But now we do not need the Merge over Move preference at all any more to explain why (64) is ungrammatical. A man simply doesn't have a reason to undergo Case driven movement.

In example (65) of course, the numeration would not contain expletive *there* nor would it contain *be_{partitive}*, rather, it would contain *be*. Rejecting (72) leads to abandoning (73), which allows us to abandon (74). A welcome result, since it is a simplification of the theory. Compare this gain in elegance and simplicity to the enormous machinery that Chomsky adds to the theory to account for (63), (64), (66), and (67).

Notice also that some kind of diacritic (like 'partitive' in the present system) is needed independently of phases and Merge over Move simply to explain why the following examples are ungrammatical:

- (80) *There seem there to be three men in the garden. (Chomsky's (2000:149 fn. 193) "perennial troublemaker" see also Bošković 1997b:98-99; Chomsky 1995b)
- (81) *There has a man read a book.
- (82) *There was told a long story (to Peter).

(83) *There boiled water on the stove.

If such a diacritic were not available, then to take just example (81), the subject *a man* could be assigned nominative case in situ (as in the standard expletive constructions on Chomsky's theory) and *there* would be introduced in [Spec, TP] for the usual reasons. Given the partitive diacritic in conjunction with the assumption that *there* needs to be assigned Case rules out examples like (81), because although *there* can be assigned nominative, *a man* remains caseless. Something must rule out such sentences. Examples of this kind abound. The partitive theory requires *there* to be structurally Case marked (nominative or accusative) ruling out example (81) trivially.

The core of the idea here is to differentiate the numerations for (66) and (67) in a rather pedestrian way: they contain different (though homophonous) lexical items.³⁵ If we distinguish the numerations for (66) and (67) this way, we are virtually forced to also

³⁵ This is not intended as a claim that the homophony is accidental. The feature *partitive* is presumable optionally added by a redundancy rule to a coherent subset of lexical items. See Levin and Rappaport Hovav 1995 for some useful ideas.

give up hypothesis (73). Without (73) there is no longer a need to assume (74), which we drop since the theory without it is simpler.³⁶

The discussion in this subsection shows that the original data motivating phase theory in Chomsky's sense, i.e. examples (63), (64), (66), and (67), really does not warrant the complexities of phase theory at all. In fact, phase theory in Chomsky's sense is more complicated than the data it is meant to account for and must therefore be rejected on grounds of parsimony.

2.2.4 Conclusion

In this section I have looked some more at punctuated paths. In subsection 2.2.1 I claimed that phrase structure theories offer punctuated paths as a way of mediating long distance relationships. There are three empirically distinct ways ways of mediating long distance relationships in language. These map in a straightforward way on the three ways long distance relations can be mediated in phrase structure. After discussing this, I attempted to reduce two properties of punctuated movement to each other: the influence the path

³⁶ Shima 2000 actually suggests to replace the Merge over Move preference be a Move over Merge preference. The main argument for the Move over Merge preference comes from superraising constructions like (i). The expletive *it* must be prevented from being merged into [Spec, TP] in the embedded sentence in (i). Otherwise later raising of *John* to the matrix [Spec, TP] might be possible, which would derive the ungrammatical (i). One way of ensuring that (i) cannot be derived is to assume that Move is preferred over Merge, which would block merger of *it*, in the embedded sentence since *John* is still there and available for movement. It might also be possible to deal with cases like (i) as a locality violation (see the speculation in fn. 31). There are more complex cases to consider like (ii), too. Both examples are from Shima 2000. See also Vukić 2003.

- (i) *John seems that it is likely t_{John} to win.
- (ii) Who was asked [_{CP} t_{who} C° [_{TP} it was told t_{who} [that Mary left]]]

has on the movement (intervention) and the influence the movement has on the path (intermediate traces).

In section 2.2.2 I showed what it would take to reduce the Head Constraint (and its empirical consequence subjacency) to locality, i.e. Attract Closest. This move is desirable, since the Head Constraint is otherwise an unmotivated, isolated stipulation. Several questions remained open in this discussion. The theory of features and feature values I developed led me to reject cyclic spell out (at least if it is narrowly construed as in Chomsky 2000 or as it was in Uriagereka 1999).

I ended up adopting a theory of phases that has one important similarity to Chomsky's theory of phases: Movement out of a phase must pass through the specifier of that phase. Since I derive this generalization from intervention rather than from cyclic spell out, the thesis of cyclic spell out becomes unnecessary for regulating successive cyclic movement. In the last part of this section I presented some arguments that giving up cyclic spell out and phases leads to a simplification of the grammar without empirical loss.

2.3 On the Nature of C-command

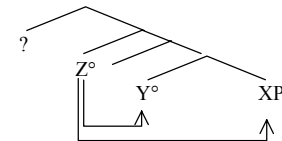
In subsection 2.2.2 I observed (p. 46 fn. 19) that the current formulation of locality entails that if two constituents c-command each other, neither of them can be attracted, because neither of them is a closest attractee. The relevant formulations are repeated here for convenience.

- (45) *Attract Closest*
If a head K attracts feature F on X, X a feature bearer of F, no constituent that bears F is closer to K than X is.

- (46) *Closeness* (to be revised)
Y is closer to K than X is if K c-commands Y and Y c-commands X.
(47) *Feature Bearer*
 β is a feature bearer of feature F iff β is a constituent and the label of β contains F.

This property is important. I argued in the preceding section that phase heads are universal feature bearers and thereby universal attractors but also universal interveners. In chapter 1 I introduced the generalization that the complements of phase heads never move stranding their embedding phase head. We now have a partial explanation for this generalization. Consider the structure in (84). Assume that Y° is a phase head and by virtue of this that it is a universal intervener. Assume further that Z° is attempting to attract XP in virtue of some feature F that XP bears. Although Y° cannot be attracted (for reasons that were left open in the previous section), Y° will intervene with attraction of XP iff sisters block attraction of each other, i.e. if the formulation of closeness given above in (46) is adopted. This is the first part of the explanation for the generalization that the complements of phase heads never strand those phase heads.³⁷

(84)



³⁷ This is only the first part of the explanation for the Stranding Generalization, since we also must rule out the derivation where XP first moves to the escape hatch position [Spec, YP] from where it could be moved out. This second part of the explanation is developed in section 2.4.

Given that sisters bearing the same feature block each other's movement, there is a problem also noted above (p. 48). The sister of a specifier, i.e. in traditional terms the X' -category, must be prevented from c-commanding its sister. Otherwise the present escape hatch theory would break down since X' would c-command [Spec, XP], the phrase in [Spec, XP] would c-command X' , and the two would intervene with each other's movement just like Y° intervenes with XPs movement in (84). In this section I show that a definition of c-command in terms of dominance alone (rather than in terms of sister containment) together with the assumption that movement chains are not constituted of copies but are multidomination structures yields the desired result. I do not review vast and quickly growing literature on multi-domination (see Blevins 1990; Bobaljik 1995a; Citko 2002; Epstein, Groat, Kawashima and Kitahara 1998; Gärtner 1999, 2002; Starke 2001; Wilder 1998, 1999). However, I do discuss briefly the debates surrounding c-command in the recent syntactic literature.

On the basis of the understanding of c-command reached here, a reformulation of closeness, i.e. a revision of (46), is offered.

The main claim of the first two sub-sections is that in a derivational, bottom up model for the generation of phrase markers, c-command is (simply) the elsewhere case of the fundamental containment (dominance) relation.

2.3.1 *C-command*

Recently there has been a debate about the nature and naturalness of the c-command relation. I will assume here without discussion that the c-command relation is a syntactically relevant relation, noting only that the definition of closeness for locality

constraints in the Relativized Minimality (Shortest Attract) family appear to require the c-command relation.

Chametzky (1996) suggests making two independent moves. Instead of asking whether node α c-commands node β (taking c-command as a relation between nodes), he suggests thinking of c-command as a relation between a node and a set of nodes: What is the set of nodes c-commanded by α ? Or what is the set of nodes that c-command α ? Chametzky's second move is to claim that the set of c-commanders of a node α form a simple class, whereas the set of c-commandees do not.

Chametzky's first move has also been made by several other authors Epstein 1999; Epstein et al. 1998, Chomsky (2000), whereas the second move is (implicitly) rejected. This is not the place to rehash the discussion (comments on Chametzky (1996) can be found in Epstein et al. (1998)).

Chametzky also claims that the c-command relation is a generalization of the sisterhood relation. Epstein and Chomsky again reach the same conclusion; both give definitions of c-command that rely on the sisterhood relation. In this section I dispute that this is a genuine insight.

Notice first that Epstein's theory cannot count as a true derivation, that is, theoretical reduction, of the c-command relation. For Epstein, syntactic structures are generated by successive applications of the operation Merge. Merge applies at the top of structures already created. According to Epstein an element α c-commands all items β with which α was merged and all items β contains. The strongly derivationalist approach Epstein takes allows an explanation of the fact that α does not c-command anything that

is higher up in the tree than α itself is; all items above α are introduced into the structure at a derivational point after α is introduced. α cannot enter into a c-command relation with items above it because these items are, literally, not part of the structure at the point of the derivation when α establishes its c-command relations.

The operation Merge creates sisterhood relations, making the sisterhood relation an entirely natural one. As noted above however, α c-commands not only all the items β with which α is merged in the course of the derivation, but also all items β contains. The second conjunct is, in the final analysis, a stipulation. Why does α c-command whatever β contains? Why are the members of β visible at all? And if they are, why isn't c-command established between α and whatever it contains and β and whatever it contains? None of this follows immediately from Epstein's account. The basic asymmetry underlying c-command remains a stipulated property.

The same is true of the comments in Chomsky (2000). Chomsky argues that the operation Merge gives rise to two natural relations: sisterhood and immediate containment (immediate dominance). We can define the reflexive, transitive closure of immediate containment, giving the containment (dominance) relation. C-command can then be defined as the composition of the two relations sister and containment. As Chomsky himself points out, this demonstrates that c-command can easily be defined in terms of natural relations, but it remains, ultimately, unexplained.

We might ask what the role of the two relations (sisterhood and containment) is in the definition of c-command. Interestingly, they do not play the same kind of role. As definition (85), a traditional definition of c-command, shows, the sisterhood relation is

unnecessary in the definition of c-command. The opposite is not true. C-command is not definable in terms of sisterhood alone. This is a strong indication that containment is more fundamental to c-command than sisterhood.

- (85) α c-commands β iff
- a. α does not contain β ,
 - b. β does not contain α , and
 - c. every node γ that contains α also contains β .

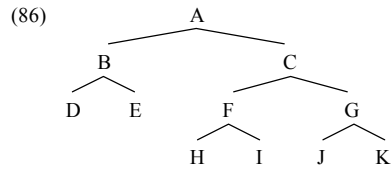
Phrase structure has long been interpreted primarily as a representation of part-whole (i.e. containment) relations.³⁸ With this in mind, consider node F in the phrase marker (86). F is related to A, C, F, H and I through reflexive containment.³⁹ If indeed containment is the most fundamental relation of the phrase marker, then F remains unrelated to B, D, E, G, J, and K in (86). Notice that the set of nodes that F c-commands is a subset of the nodes not in any containment relation with F. Put the other way around, the set containing B, D, E, G, J, and K is a proper superset of the set of nodes that F c-commands: G, J, and K.

³⁸ The question whether precedence is a primitive relation of phrase structures is discussed at length in Chametzky (1996). The issue is orthogonal to my concerns.

³⁹ It might be objected that the most natural relation is not reflexive containment but proper containment. It might also be objected that containment comes in two varieties: A contains B, and A is contained in B.

Neither of these objections changes the point made here if we can assume that the identity relation is available. The three relations 'proper contain', 'contained by', and 'identical to' are mutually exclusive as shown in (ia-c). These relations leave a number of nodes unrelated to F as shown in (id). The elsewhere relation is as in the text. I will use reflexive containment for simplicity in the text.

- (i) a. F contains H and I.
b. F is contained by A and C.



To go from the absence of a containment relation (85a-b) to the c-command relation, we have to add the seemingly arbitrary stipulation (85c). F c-commands node β only if the mother of F dominates β . This is restated in (87).

- (87) α c-commands β iff
- a. α and β do not enter into containment relations with each other,
 - and b. every node/the first node containing α contains β .

Condition (87a) can be understood as an elsewhere condition; condition (87b) remains a stipulation given the representation in (86). Note, however, that in a system where syntactic representations are built by merging at the root, (87b) becomes unnecessary. (For some comments on the theoretical merits of Chomsky's 'extension condition' see Chametzky (2000).)

Consider the step in the derivation when F and G are merged, that is, the moment the new part-whole relation between C and the tree rooted at F and between C and the tree rooted at G are established, the nodes in the new object (i.e. the constituent labeled C) are either in a dominance relation with F (C, F, H, I) or they are not (G, J, K). In other words, the moment when F is embedded into a larger structure, the c-command relation is identical to the elsewhere case of domination.

-
- c. F is identical to F.
 - d. F is in the elsewhere relation to B, D, E, G, J, and K.

Suppose that when F and G are merged forming C, this must result in F, G, and C entering into a significant relation with all parts of the new structure.⁴⁰ C is automatically related to all parts of the new structure by containment. F is in containment relations with some parts of the structure (F, H, I, and C). F will enter into the elsewhere relation with all remaining parts of the new structure (G, J, K). This elsewhere relation is what we call c-command. Similar comments apply to G. Since F and G (but not their proper parts) are the targets of the operation merge, it is natural, although not logically necessary, to single them out in this particular way: F and G enter into significant relations with all parts of the new structure, but their proper parts do not. For the structures considered so far, Epstein's and my definitions yield identical results as desired.

F forms a phrase marker with C, H, I, G, J, and K. That set of nodes is partitioned into the nodes that enter into dominance relations and the nodes that don't. The absence of dominance relations is called c-command.

In other words, the line of reasoning in Chametzky, Epstein et al, and Chomsky according to which c-command is a generalization of sisterhood is not the most fruitful way to look at c-command. C-command is better understood as the elsewhere case of domination.

⁴⁰ Chametzky 1996, 2000 discusses at some length that traditional theories of phrase structure have taken the view that every two nodes in a phrase marker are related by one of the primitive relations of precedence and dominance. He argues at length that precedence is not a primitive relation. On that view precedence was the elsewhere case of dominance. The view taken here claims this privilege for c-command.

The representational objects that are formed, phrase markers like (86), of course contain much more information than that. The claim here is that only those relations that are derivationally licensed are linguistically relevant. Only containment and the elsewhere relation of containment are thus licensed.

2.3.2 Movement

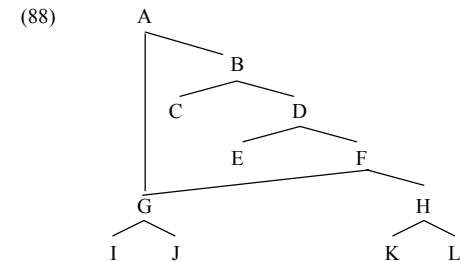
As argued in for example Bobaljik (1995a), Brody (1995), Chametzky (2000), Chomsky (2001a; 2001b), Epstein (1998), Gärtner (1999; 2002), Starke (2001), the analysis of movement in terms of copying and merging is problematic because of its redundancy. Movement is construed in Chomsky 1995b as a complex operation consisting of a copying step and a merger step. The merger step seems unavoidable. It is unavoidable to describe the derivation from the numeration to a connected phrase marker and it is also unavoidable in describing movement relations. But what about the copying step? The claim is that it is in fact unnecessary.

Two solutions have emerged. One can state movement in terms of re-merge (multidomination, occurrences) or generate the members of a chain separately. Without argument, I adopt the first option here. Apart from the redundancy problem noted above, Gärtner (1999; 2002) points out that introducing actual copies in the course of the syntactic derivation multiplies the unchecked uninterpretable features within these copies. If copies have a life of their own so to speak (Nunes 1995, 1996, 1999, to appear), then the number of operations needed to check these features is not reduced by movement (i.e. copy and merge) but actually increases. On this view it is unclear how any derivation can ever converge. To avoid this problem, a special mechanism IDENT can be invented to

make sure that copies remain identical throughout their lives. This complication can easily be avoided if we assume that movement is merge without copying, i.e. if we assume that instead of having the operations Merge, Copy, and IDENT, the grammar only has the operation Merge. Further arguments have been given in the literature cited above.

I would like to explore now the consequences that the above definition of c-command has in multidomination structures since this allows me to deduce without further stipulation what I need to deduce, namely that in certain cases X'-categories do not c-command their specifiers and vice versa. Given this lack of c-command, X'-categories do not intervene with attraction of the specifier since intervention requires c-command.

Consider the structure in (88). Suppose that at a certain point in the derivation, F has already been created and E is taken from the numeration. The next step is to create D from E and F, then B from C and D, and finally A from B and G.



When G first enters the structure in the creating of F, its c-command domain according to the regular definition and according to the definition arrived at above, is H,

K, and L. When G is targeted by an operation again, naturally, we ask again which nodes it stands in a domination relation with and which it doesn't. The result is now as in (89).

- (89) a. Domination: A, G, I, J, B, D, F
b. C-command: C, E, H, K, L

This, of course, does not correspond to the received wisdom according to which a moved item must c-command its own trace. However, no harm appears to be done, since the c-command requirement on traces largely follows from the extension condition.⁴¹

Interestingly, B and G do not enter into a c-command relation on the present definition. G has been moved, so on standard assumptions it is either in a specifier or in an adjoined position. Kayne (1994) stipulates (and Chomsky (1995b) follows him) that the sisters of specifiers/adjuncts do not participate in c-command relations. At least for the case of moved specifiers/adjuncts, the current proposal gives a principled reason why this is so. It follows from the elsewhere nature of c-command. In Kayne's and Chomsky's systems this property is stipulated.

There are obvious consequences to explore here, especially the connections between the present notion of c-command and Moro's (Moro 2000) theory of dynamic asymmetry. For Moro, movement can be driven by the need to break symmetries in the phrase marker, basically, to ensure that a linearization algorithm like the LCA that relies

⁴¹ A question arises with respect to the status of sideward movement in the present system. I will not pursue the issue here since nothing hinges on the resolution as far as I can see.

on asymmetric c-command can linearize it.⁴² In the present system it follows naturally that movement breaks symmetry. I will not pursue the issue here for the reason that the problem Moro's system solves strikes me as slightly artificial. The symmetries occur in terms of c-command under the exact terms of Kayne's (1994) formulation. The category segment distinction is crucial for this definition of c-command. Notice, that this distinction is defined over labeled trees and that labels cannot be completely eliminated from the definition, i.e. it is not a strictly geometric definition. But the labeled trees that Kayne and Moro use contain enough information to derive linear orders with the help of the labels. The particular restriction on how labels are used in the system, to distinguish categories from segments but not otherwise, doesn't strike me as particularly natural.

Note that this is exactly what is required for the theory of escape hatches developed in section 2.2. An element moves to a specifier position attracted by a feature of an attracting head. The sister of the moved element does not c-command the moved item itself and hence cannot intervene with it for the purposes of locality.⁴³

The same is not true of unmoved specifiers or, in fact, complements. Unmoved specifiers and complements will be in mutual c-command relations with their sisters.⁴⁴

⁴² In the present system we would have to reformulate the LCA in terms of the notion 'total c-command' discussed immediately below.

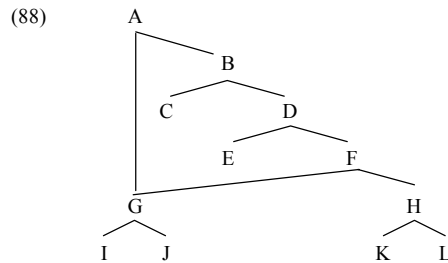
⁴³ In section 2.2.2.3 I speculated that the feature responsible for subjacency does not project up from the head. If that speculation is correct, then the intermediate projection could be allowed to c-command the specifier. It still would not intervene because it wouldn't bear the relevant feature.

⁴⁴ If asymmetric c-command (between occurrences) is necessary for linearization at PF (Chomsky 1995b) and if the present account of the asymmetry between a moved

The multidominance (occurrence) theory of movement does not yield any of these results when combined with a definition of c-command in terms of sister containment. If c-command is defined in terms of sister containment, G will c-command all of B, C, D, E, F, G, H, I, J, K, L and in (88). The asymmetry between the moved specifier and its sister does not fall out under the view of c-command as sister containment.

2.3.3 Closer

The approach to c-command developed in this section together with a multidomination approach to movement has solved an open problem: why does the intermediate projection, sister of a moved XP not c-command that XP. It's a pleasing result. But does it really allow the theory of escape hatches to work? Consider again the structure in (88), repeated here for reference.



specifier and its sister is correct, then there might be reasons to believe that every category has to move at least once to ensure linearizability. This conclusion, if true, is even more radical than the one drawn in Moro (2000), where only certain classes of configurations force movement for linearization. The issue warrants further investigation.

True, B would not intervene with further movement of G because it does not c-command it, however C still c-commands G and G c-commands C. So there is a symmetric c-command relation between the two. On the definitions of Attract Closest and Closeness offered above (and repeated here for reference), the presence of C should block further movement of G. We saw above in the discussion of (84) why we can't just give up the idea that two items that mutually c-command each other intervene with each other. So a different solution needs to be sought.

- (45) *Attract Closest*
If a head K attracts feature F on X, X a feature bearer of F, no constituent that bears F is closer to K than X is.
- (47) *Feature Bearer*
 β is a feature bearer of feature F iff β is a constituent and the label of β contains F.
- (46) *Closeness* (to be revised)
Y is closer to K than X is if K c-commands Y and Y c-commands X.

There is, fortunately, an asymmetry between C and G. Whereas C c-commands only one occurrence of G, G c-commands all occurrences of C. This would be true even if C had been moved from some lower position to its current position in (88). We can define occurrences of items as pairs of the item and its sister(s) or pairs of the item and its mother(s).

Closeness, to operate correctly, must be defined in terms of a notion of total c-command. Total c-command is defined as follows:

- (90) α *totally c-commands* β iff
- (i) α c-commands β and
 - (ii) there is a constituent γ immediately containing α , such that γ (reflexively) contains all constituents that immediately contain β .

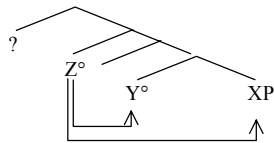
According to this definition G totally c-commands β , because there is a constituent, namely A , immediately containing G that contains all constituents immediately containing C . On the other hand, C does not totally c-command G since there is no constituent immediately containing C that contains all nodes immediately containing G . In particular, B does not contain A .

This suggests a formulation of closeness in terms of total c-command.

- (91) Closeness (final version)
 Y is closer to K than X is if K c-commands Y and Y totally c-commands X .

Notice that two sisters that symmetrically c-command each other can still symmetrically totally c-command each other, for example if both are base generated. Recall the structure from (84) above. Y° and XP not only c-command each other, they totally c-command each other. Thus Y° still intervenes with the attraction of XP , hence XP still intervenes with the attraction of Y° .

(84)



2.3.4 Conclusion

I have argued that in a model where phrase markers are constructed derivationally bottom up and the basic relation is the containment (dominance, part-whole) relation, c-com-

mand falls out naturally as the elsewhere case. This view of c-command, when combined with a view of movement as multidomination, entails that a specifier/adjunct created by movement and its sister (usually taken to be an intermediate, X' , projection – but the label or projection level have no influence on the result attained here) do not c-command each other. Nor does the sister of a specifier created by movement c-command into the specifier at all. For the purposes of this study it is necessary to block the sister of a moved constituent from c-commanding the moved constituent, otherwise the theory of escape hatches in terms of Attract Closest collapses.

As developed in subsections 2.3.1 and 2.3.2, the notion of c-command does not distinguish between different occurrences of an item. Put in traditional terms, my definition of c-command has the following property. If XP gives rise to chain CH_1 and YP gives rise to chain CH_2 and some member of CH_1 c-commands some member of CH_2 , then XP c-commands YP . Under some circumstances, this definition is not restrictive enough. In subsection 2.3.3 the notion of total c-command was defined. Total c-command has roughly the following effect in the above example: XP totally c-commands YP only if there is a member of CH_1 that c-commands all members of CH_2 .

2.4 Anti-locality

In this section I show how the assumptions developed in the preceding sections together with the assumption that movement operations are subject to a last resort condition conspire to make a specific prediction. We expect to find anti-locality effects, i.e. we predict that certain conceivable movement steps are actually impossible, because they are too short.

I will assume the Last Resort condition (92).⁴⁵

(92) Last Resort

A constituent α may only be merged, i.e. base-merged or re-merged, if that leads to the immediate satisfaction of a previously unsatisfiable feature.

The condition demands that every movement step must lead to the satisfaction of some feature. Obviously, we need to have a minimal understanding of feature satisfaction (checking, agree,...) in order to be able to state this condition. Subsection 2.4.1 explores the different ways in which features can be satisfied. Subsection 2.4.2 offers discusses the formulation of Last Resort given above and draws out the anti-locality prediction.

Subsection

⁴⁵ The formulation of Last Resort makes base-merger and re-merger subject to economy considerations. For base-merger see Chomsky 2001a, b and Collins 2002.

2.4.3 concludes.

2.4.1 *Feature Satisfaction*

Under Minimalist assumptions all movements must be triggered. The standard assumption is that movement, somehow, goes hand in hand with feature checking (agreement). There have been attempts to unify all feature checking (all checking requires a spec-head configuration (Chomsky 1993); checking happens uniformly in the checking domain (Chomsky 1995b); checking happens uniformly under closest c-command between probe and goal (Chomsky 2000). These attempts were driven by the desire to reduce the configurations in which features were checked (this was mainly Case assignment) in the Government and Binding theory. Case assignment there was viewed as heterogeneous as it happened under sisterhood (V – DO), under spec-head agreement (T – SU), and under government into the specifier of the complement (ECM). Once the raising analysis of ECM had been revived (Postal 1974; Chomsky and Lasnik 1993; Lasnik and Saito 1991), it was tempting to treat all Case assignment as instances of the spec-head configuration (Chomsky 1993). In this form, the unification encountered a number of problems since overt and covert movement are not adequately distinguished (see den Dikken 1995; Lasnik in press-a; Lasnik and Saito 1991 for A-movement, and the facts concerning Romanian parasitic gaps in Bošković 2001b for A'-movement). Said authors show that covert A-movement does not give rise to new binding (for which see already Mahajan 1990:23 fn 10) and scope configurations. The argument in its most simple form can be based on examples like (93) and (94).

(93) a. There isn't a man in the garden.

- (94) b. A man isn't in the garden.
 a. ✓Some applicants_i seem to each other_i to be eligible for the job.
 b. *There seem to each other_i some applicants_i to be eligible for the job.

In all of (93a) through (94b), the finite verb, or more precisely the head T°, agrees in person and number with *a man*. In (93a) and (94a), *a man* agrees without moving to T°, whereas in (93b) and (94b) agreement is accompanied by movement of *a man* to [Spec, TP]. Interestingly, in (93a) *a man* must take scope below negation, whereas in (93b) it may scope above negation. And in (94a) the reciprocal *each other* is not bound, but in (94b) it is. The generalization is that agreement coupled with overt movement creates new binding and scope configurations, but agreement without overt movement does not.

The difference between feature satisfaction with overt movement and feature satisfaction without overt movement must therefore be visible at the LF interface. This argues against Chomsky's (1993) system, where overt and covert movement are different at the PF interface but identical at LF.

In Chomsky 1995b the distinction is made in terms of feature movement vs. category movement. The price to pay for this is a somewhat baroque definition of checking domain. Checking domains are defined based on the notion of 'domain' – which designated the m-command domain of a head H in Chomsky 1993. (In his most recent writings Chomsky has taken to calling the c-command domain of an item its domain.) The m-command domain is then shrunk to the minimal m-command domain containing only those categories that are m-commanded by a head but not dominated by another category that is also m-commanded by that head. Finally, the minimal m-

command domain is partitioned into the complement domain (which is the subset of the minimal m-command domain containing, essentially, only the complement) and the rest. The rest is the checking domain of H. The checking domain is an elsewhere-category that contains specifiers of HP, adjuncts to H°, adjuncts to HP, adjuncts to specifiers of HP, and adjuncts to H's complement. What is crucial here is that the definition of checking domain comprises adjuncts to H° (for covert feature checking) and specifier of HP (for overt feature checking). It is unclear why head adjunction, which forms sisterhood relations between H° and whatever adjoins to it, and the complementation relation, which is also a sisterhood relation, should not form a natural class for example, but they do not under these definitions.

Chomsky 2000 abandons this approach and assumes that all feature satisfaction requires local c-command by the agreeing head. Some instances of feature satisfaction are then followed by movement. In other words, according to Chomsky agreement feeds movement. It is unlikely that this is a workable solution. Recall that we have had reasons above from complementizer agreement to reject this. In Irish complementizer agreement happens only if overt *wh*-movement through [Spec, CP] has taken place but not with an in situ *wh*-phrase. If agreement generally fed movement as on Chomsky's theory, complementizer agreement should be possible both with and without overt movement,

counter to fact. On the other hand, if we assume that complementizer agreement is fed by movement, the problem disappears.⁴⁶

In any case, it appears that three kinds of feature satisfaction have remained intact more or less through these various revisions of checking/agreement theory. There is feature satisfaction that only requires (closest) c-command (this is roughly Chomsky's AGREE), there is feature satisfaction that requires (closest) c-command and movement (roughly AGREE + (EP)P or, likewise roughly, spec-head agreement), and finally there are features that are only satisfied under base-merge (for Chomsky theta-features fall in this category).⁴⁷

Under standard formulations these three feature satisfaction relations have fairly little in common. I will now reconstruct these three relations within the framework developed here. My formulation has the advantage of bringing out the similarity between all three relations.

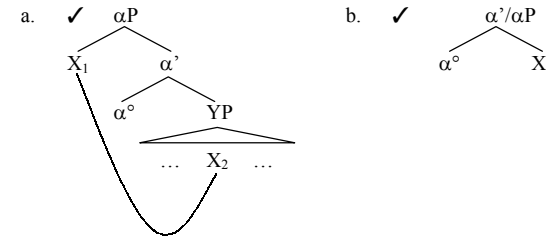
Given the definitions of c-command and total c-command arrived at above, I will distinguish (i) feature satisfaction under (closest) c-command, (ii) feature satisfaction under mutual (closest) c-command, and (iii) feature satisfaction under mutual total (closest) c-command. I ignore the locality issue in the discussion to follow. The three modes of feature satisfaction map in the following way onto the more familiar notions.

⁴⁶ See Boeckx 1999 for an A-movement paradigm establishing the same point. See Bošković 1997b:87 for facts from agreement with conjoined NPs which also indicate that movement feeds agreement.

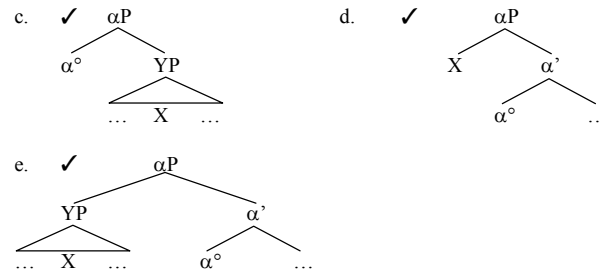
Feature satisfaction under c-command corresponds to AGREE. Feature satisfaction under mutual c-command corresponds roughly to AGREE plus Move (or spec-head agreement). Feature satisfaction under mutual total c-command corresponds to feature satisfaction under base-merger.

Let's consider these three relations in turn, starting with the weakest: Feature satisfaction under c-command. The c-command requirement is fulfilled just in case a feature bearer of ϕ c-commands another feature bearer of ϕ . In (95) the configurations in which the c-command condition is satisfied are marked with a checkmark, and the ones where it is violated are marked with an asterisk. For perspicuity the structures involve a copy notation rather than multidominance. Different occurrences of the same item are subscripted for ease of reference only. Since the theory assumes single copies of each item only, the subscripts on the copies in the structures, too, are there for expository purposes only.

(95) Feature satisfaction under c-command



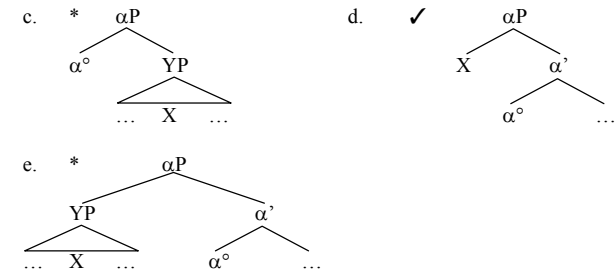
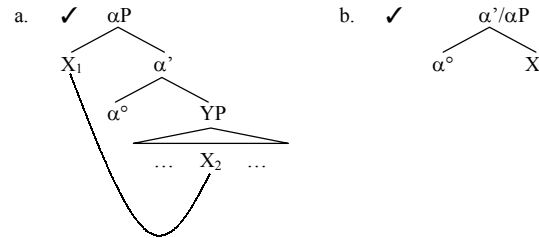
⁴⁷ Not all authors agree that theta-roles can only be assigned under base merge. See among others Bošković 1994a; Bošković and Takahashi 1998; Hornstein 1998, 1999; Hornstein and Nunes 1999. On the other hand Brody 1995:14 agrees with Chomsky.



In (95a) α^o c-commands X_2 , thereby satisfying the requirement. The fact that α^o does not totally c-command X^o is irrelevant. In (95b) and (95c) α^o trivially c-commands X . Thus the c-command requirement can be satisfied. Under the formulation given here, (95e) is a possible configuration of feature satisfaction under the formulation given. This seems to be wrong empirically, so we need to stipulate that this configuration is not a possible feature satisfaction relation.

I now turn to feature satisfaction under mutual c-command. Mutual c-command holds if a feature bearer of ϕ c-commands X chain and X c-commands that feature bearer of ϕ .

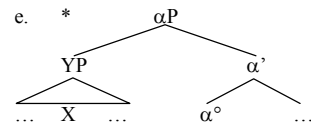
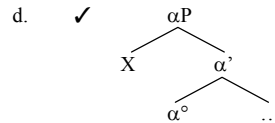
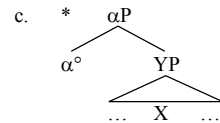
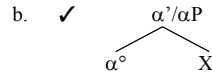
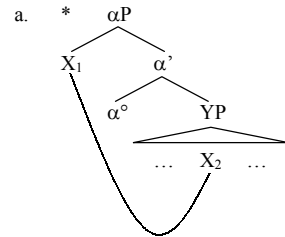
(96) Feature satisfaction under mutual c-command:



In (96a) α^o c-commands as before X_2 and X_1 c-commands α^o . Thus the mutual c-command requirement is satisfied. In (96b) α^o c-commands X and X c-commands α^o , again satisfying the mutual c-command requirement. In (96c) α^o c-commands X , but since X does not c-command α^o , mutual c-command is not satisfied. In (96d) the relevant mutual c-command relation obtains between X and α' . And in (96e) finally, α' c-commands X , but since X does not c-command α' , no mutual c-command relation is established. Notice about this, that what used to be called spec-head agreement, i.e. (96a) and (96d), fall under this condition as well as the head to complement relation (96b). In other words the early minimalist drive to unify the conditions under which Case assignment happens might have been right. But that it was a mistake to seek to make all Case assignment be conditional on spec-head agreement. Rather we see here that a uniform condition in terms of c-command captures both the spec-head and head-complement configurations. Under this formulation Case assignment is already unified.

Finally we turn to mutual total c-command. α and X mutually totally c-command iff α totally c-commands X and X totally c-commands α .

(97) Mutual total c-command



Notice that only base merged complements (97b) and base merged specifiers (97d), but not moved specifiers (97a) or XPs in any other configuration (97c) and (97e) satisfy the mutual total c-command requirement. This is a base merge requirement (as Chomsky invokes it for theta assignment) recast in terms of c-command.

The advantage of the present formulation is that it shows the theoretical cohesion of the various ways of feature satisfaction. They are all defined in terms of c-command. It also brings out the fact that there is a hierarchy of satisfaction relations, mutual c-command being stricter than c-command and mutual total c-command being stricter than mutual c-command. Finally, the spec-head configuration and the head-complement configurations are unified under the present formulation.

We can now see that ϕ -feature agreement of verbs in English require only c-command as (98) illustrates, whereas Case assignment appears to require mutual c-command (99) (see Epstein and Seely 1999 for pertinent discussion), and if theta roles are only assigned in a base merged position, theta assignment requires mutual total c-command.

(98) There T^o seem to have been [three men] in the room.

(99) Three men seem to have been arrested ~~three men~~.

As we have just seen, closest c-command_{cod} is the most permissive relation. I assume that it is the minimum condition on establishing a syntactic feature relation. In some cases, like ϕ -feature agreement, this minimum requirement is all there is to establish. In other cases, like structural Case, mutual c-command_{cod} is required. If it is not supported in the representation as is at a particular step of the derivation, then movement (re-merger) may apply as long as the weaker featural relation of closest c-command_{cod} is supported by the original representation. Another way of saying this is that ϕ -feature agreement and Attraction require local c-command, whereas Case agreement requires mutual c-command.

Certain details (like the fact that I take agreement to be a relation between feature bearers rather than heads) apart, the main operative assumptions in the present account that distinguish it from others are: (i) movement feeds agreement; and (ii) all feature relations are not equal, there is, in fact, a hierarchy of them (mutual total c-command is

stricter than mutual c-command which in turn is stricter than closest c-command).⁴⁸ It stands to reason that the agreement that we see realized morphologically is not a direct reflex of feature checking, but a reflex of feature valuation. As discussed above in section 2.2.2, valuation happens when all syntax is done. It is therefore not surprising at all that movement may feed agreement. In fact, given the considerations from section 2.2.2 it must. Recall that we concluded in section 2.2.2 that valuation requires closest checking. Recall now the definition of closeness.⁴⁹

- (91) Closeness (final version)
Y is closer to K than X is if K c-commands Y and Y totally c-commands X.

2.4.2 Last Resort and beyond

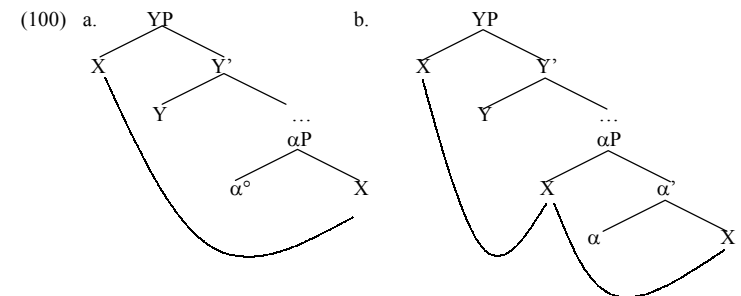
The ultimate aim of this chapter, recall, is to derive the stranding generalization, which I introduced in chapter 1. It is repeated here.

- (1) Stranding Generalization
Given a phase head α° and a constituent X in α° 's c-command domain
a. $\Diamond\checkmark[X \dots [\alpha^\circ [\dots t_X \dots]] \dots]$ and
b. $\neg\Diamond\checkmark[X \dots [\alpha^\circ t_X] \dots]$

⁴⁸ I have pretended so far that the more stringent requirement of mutual c-command_{ood} is fully co-extensive with overt movement, i.e. with feature strength or Chomsky's recent P-features. As always, there are serious questions whether this is a perfect fit for the data (see Pesetsky 2000) and if not how to deal with the problems.

⁴⁹ In Icelandic a dative experiencer in an expletive construction blocks agreement with the associate. However, when the experiencer is moved out of the way by A-movement, its trace no longer blocks agreement (Boeckx 1999; Holmberg and Hróarsdóttir 2003 among others). A'-traces on the other hand do block agreement. These facts require special assumptions about A-traces on my, or in fact on any account (see Holmberg and Hróarsdóttir 2003:1005 fn. 1004 and Lasnik 1999a for possible approaches).

There are two possible derivations for the illicit structure (1b). On one derivation, X in (1b) moves out of αP directly without stopping over in [Spec, αP]. The resulting structure for this derivation is given in (100a). This structure is ruled out as we saw. Movement of X is blocked by the intervening α° . Prior to movement, when the conditions on Attraction are applied, α° totally c-commands X and hence intervenes with its movement – this follows from the definition of Closeness just reviewed. However, there is another derivation, the resulting structure for which is sketched in (100b). If (1b) is indeed to be ruled out, the derivation of (100b) must also be blocked.



Blocking (100b) requires recourse to the Last Resort condition introduced above and repeated here.

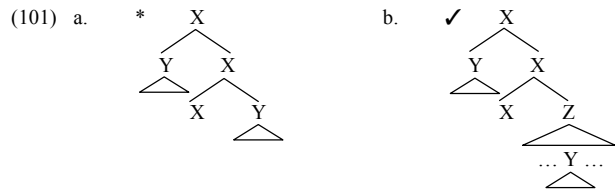
- (92) Last Resort
A constituent α may only be merged, i.e. base-merged or re-merged, if that leads to the immediate satisfaction of a previously unsatisfiable feature.

The initial step of movement, where X moves from the complement position to the specifier position of αP in (100b), is blocked by the Last Resort condition. This is so because it does not lead to a new feature satisfaction relation. X and α° are already in the feature satisfaction relations of mutual total c-command, mutual c-command and c-

command. Nothing further is achieved by moving X to [Spec, α P]. (100b) is then blocked by the Last Resort condition. This concludes the theoretical derivation of (1b).

The Last Resort condition has some further consequences. In what follows I make use of the more traditional (and more reader- and typist-friendly) copy notation rather than multidominance structures. This is not to be construed as a theoretical claim.

The first point to note was already implicitly illustrated above. When a constituent is base merged with another (either as its complement or as its specifier) the two are in the most stringent of all local conditions; they mutually totally c-command. This entails automatically that they are in the mutual c-command relation and in the closest c-command relation. It follows that if X and Y, bearers of some α -feature in need of satisfaction, are base merged, then the feature will be able to be satisfied.



The movement indicated in (101a) (which is essentially (100b) repeated) is ruled out. It can never allow feature satisfaction that was impossible before the movement took place. The same is true for movement of a specifier to another specifier position within the same phrase, of a specifier to an adjoined position within the same phrase or of a

complement to an adjunct position within the same phrase.⁵⁰ The asterisk in (101a) indicates that this configuration is categorically ruled out under all circumstances by the Last Resort condition.

The system developed here then provides a way of empirically teasing apart a system like Chomsky 2000, 2001a, b, where movement is driven by a (EP)P-feature on a head, which is satisfied as a rider on a checking relation, and an approach where the movement establishes the checking relation. Although under many circumstances these systems look like notational variants, there are cases where they are not. The structures in (101) exemplify in the abstract one such case, another case was discussed above under the heading that movement feeds agreement. The Stranding Generalization from chapter 1 rests crucially on the ungrammaticality of (101a). If the Stranding Generalization is correct, as I argue in chapters 3 and 4, then we have two arguments against the generalized (EP)P approach to movement.

Consider now (101b). The checkmark indicates a possible configuration. It does not indicate that all configurations of this type are grammatical. In fact they are not. X might not be able to attract Y because of locality or X might not bear a feature that requires mutual c-command to be satisfied. However, there are cases schematically like (101b) and that is what the checkmark indicates. Suppose that before movement takes place X may attract Y and X has a feature that requires mutual c-command for

⁵⁰ A condition that rules out movement from some specifier position to another specifier (or adjoined position) within the same XP has been invoked for example in Bošković 1994a; 1997b; Saito and Murasugi 1999. These authors also develop some of

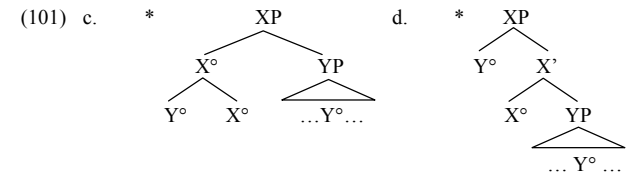
satisfaction. In this case, Y may move, since this establishes the required mutual c-command relation thus leading to an immediate feature satisfaction which was blocked before.

The discussion of these two cases shows again that movement from complement to specifier within the same phrase is systematically ruled out by the Last Resort condition. Longer movements may be allowed. Similar anti-locality conditions are also assumed in Bošković 1994a; 1997b; Grohmann 2000; Murasugi and Saito 1995; Ishii 1997, 1999. Unlike these authors I have tried to derive rather than stipulate the anti-locality condition. The Anti-locality conditions assumed by these authors have various degrees of overlap with the condition derived here, but I will not pursue the matter.

The same considerations that rule out (101a) also rule out head movement as a syntactic movement process, i.e. as a process that builds or alters syntactic structures. The structure in (101c) is a standard head movement configuration (bar levels are indicated strictly for legibility). On the assumption that labels are full copies of the terminals that project them (but see the discussion in section 2.2.2, which might weaken this point), (101c) is ruled out. Before Y° moves, a feature bearer of all features that Y° bears and a feature bearer of all feature that X° bears are already in the closest possible relation: X° and YP totally c-command each other. The movement of Y° can therefore serve no

the consequences of that ban. I concentrate here on movement from the complement to the specifier position.

feature satisfaction purpose and is therefore ruled out by the Last Resort Condition.⁵¹ The same is obviously true if head movement is construed as movement of a head to a specifier position as suggested for example in Matushansky 2002. The relevant configuration is depicted in (101d).



The remainder of this thesis explores some of the consequences of this general ban against movement that is too short in the well-defined sense given in this chapter. The ban against head movement that emerges from the present system, requires that traditional head movement analyses be recast in different terms. In chapter 5 I offer a treatment of head movement in the spirit of Brody 2000, where ‘head movement’ emerges as a result of word formation and linearization.

The second main prediction of the system is the Stranding Generalization from chapter 1. We have seen how it emerges from the assumptions about successive cyclicity from section 2.2 with anti-locality the results of this section. The two following chapters deal extensively with anti-locality effects of exactly this type.

⁵¹ Bošković (2001b) actually claims that heads can specify as a part of their lexical entry that checking for that particular lexical item requires head adjunction. The relevant cases involve very specific clitics in South Slavic in particular the interrogative particle *li*. See Abels 2000 for a reanalysis of these facts in terms that are close in spirit to the present work and do not rely on such a lexical specification.

A brief remark on split CPs, TPs, VPs, etc. (see Alexiadou 1997; Cinque 1999; Pollock 1989b; Rizzi 1997 and much related work) is in order here. The anti-locality result arrived at here has empirical bite only to the extent that these categories remain more or less unsplit. If CP for example is split into $CP_1 > CP_2 > CP_3$ it may be that CP_1 is a phase head and that by anti-locality its complement CP_2 is frozen in place, but this shows virtually nothing, because this leaves CP_3 as a potentially movable phrase embedded under CP. In particular the claim that TP is frozen in place because it is the complement of a phase head C° has little merit in a model that assumes such split categories.

The anticipated objection assumes that the intricate hierarchies discovered within the cartographic approach are expressed in the projection of phrase structure, i.e. that the observed hierarchies correspond to a fixed sequence of functional heads in a local syntactic relation with each other (selection). These assumptions give rise to several paradoxes. Nilsen (2001; 2003) discusses cases where the assumed transitivity in the ordering of adverbs fails. Bobaljik (Bobaljik 1999) observes that it is impossible to construct a consistent sequence of functional heads that will account for all the ordering facts observed, i.e. that will be able to handle (a) the fixed relative order of auxiliaries and main verbs, and (b) the fixed relative orders of several adjectives, and (c) the fixed relative orders of DP-related position with respect to each other, and (d) the fairly free interspersing of items from these three orderings with each other. See also Ernst 2002 for much relevant discussion.

Finally, there are many cases where adverb ordering clearly does not result from a local selectional requirement but where considerations of relative scope are paramount.

To give just a single example, Cinque (1999:6) argues that *already* must precede *no longer* (102). Cinque's original examples are Italian, but the ordering facts hold in English as well. Notice that both *already* and *no longer* can also appear in embedded clauses and in matrix clauses (103)-(104).

- (102) a. ✓He already no longer goes to school.
 b. *He no longer already goes to school.
 (103) a. ✓It is no longer the case that he goes to school.
 b. ✓It is already the case that he goes to school.
 (104) a. ✓It is the case that he already goes to school.
 b. ✓It is the case the he no longer goes to school.

Crucially, when one of the adverbs appears in the matrix clause and the other appears in the embedded clause, the relative order between the two adverbs must be preserved (105).⁵²

- (105) a. ✓It is already the case that he no longer goes to school.
 b. *It is no longer the case that he already goes to school.

The examples in (105) indicate that Cinque's explanation for the contrast in (102), relying as it does on local selectional requirements, is too narrow. In fact the example suggests a semantic explanation. Consideration of the temporal and quantificational semantics of these two adverbs suggests that this might be true and that (105b) is nonsensical (possibly a presupposition failure). The same is true for many more of the relative adverb orders considered by Cinque.

⁵² Nilsen 2003 would treat *already* as a positive polarity item and derive the relative order of adverbs in (102) that way. Example (i) is grammatical indicating that negative elements in the matrix do not rule out *already* in the embedded clause. Given this, Nilsen's solution will not explain why (105b) is unacceptable.

- (i) It is not the case that he already goes to school.

Together, all of this casts doubt on the hypothesis that Cinque's and related hierarchies should really be represented in the phrase structure. I will assume that they are not and that the phrase structure objects that are relevant are relatively simple, $C > T > v > V$ or something along those lines.⁵³

⁵³ Grohmann 2000 discusses a different kind of anti-locality effect. For him the clause is divided into several different zones. In general, no item may move within one and the same zone. Movement is possible only across zones. If Grohmann's zones can be (partially) identified with the projections that I am assuming (C – the discourse zone, T – the Case zone, v and V – the θ -zone), then Grohmann's ban would be a special case of condition pursued here. On the other hand, Grohmann does allow movement within a particular zone under certain conditions: Movement within a zone is allowed if it has a 'radical effect on output'. I will not discuss Grohmann's approach to anti-locality here noting only that his examples of short movement would be disallowed in the present system.

2.4.3 Conclusion

In this section I have laid out my assumptions about feature satisfaction and Last Resort. Together they give rise to anti-locality conditions. Movement that is too short in a well defined sense is ruled out by Last Resort, because it cannot possibly lead to new, immediate, feature satisfaction possibilities.

2.5 Pulling Together the Threads

This chapter has laid the theoretical groundwork for the exploration of the issue of anti-locality in the following chapters. The theoretical conclusion at the end of the chapter was that movement steps that are not long enough are ruled out.

Consider again the three different possibilities of how long distance relations can be mediated. This was illustrated in (38) repeated here for convenience. Let's assume that all three theories of movement paths are supplemented by the last resort condition on movement (92).

- (38) a. α _____ β uniform path, path unaffected by relation
 b. α | | | | | | | β uniform path, path fully affected by relation
 c. α | ... | ... | ... | ... β punctuated path

Anti-locality effects are unexpected if the movement path is uniform. Under the uniform path hypothesis depicted in (38a), we might find anti-locality effects under short distance movement, but they should disappear under long distance movement. This is so,

because short movement may be too short just in case it moves from complement to specifier. If the total distance traveled is longer the anti-locality effect would disappear, because (38a) does not require any intermediate traces or copies to be created. Discussion in the next chapter will show this expectation of (38a) is not to be borne out.

Similarly, if we assume the uniform path hypothesis (38b), anti-locality effects are also unexpected. This is so because either (i) the first step of each and every movement is too short and movement is never allowed or (ii) the first step of each and every movement is long enough to escape the effects of anti-locality. Option (i) can be discarded on empirical grounds. Making theory (38b) option (ii) sensitive to different launching sites of movement would be a complication of the theory as it stands. Uniform path theories do not lead to the expectation that anti-locality might exist.

The situation is slightly different under the punctuated path hypothesis (38c). Although there is no logical necessity that leads from punctuated paths straight to anti-locality effects since everything I just said about option (38b) could in principle carry over, but if paths are punctuated because of the intervention of universal interveners, my phase heads, then anti-locality effects follow as demonstrated in the previous section.

In this sense, anti-locality is the flip-side of subjacency. And that is what holds this chapter together. I showed in section 2.1 that movement paths are punctuated rather than uniform. In section 2.2 I developed a theory of phase heads and escape hatches which unifies the Head Constraint/Phase Impenetrability Condition with the independently necessary locality condition Attract Closest.

Given minimalist approaches to c-command (sister containment) the explanation was incomplete. A (moved) specifier/adjunct and its X'-sister should c-command each other. If they share a feature α relevant to a higher attracting head, then the moved specifier and its X'-sister are equally close to the attracting head and should therefore, by the definition of closest, block each other's movement. This issue was taken up in section 2.3. The assumptions made in section 2.3, i.e. the definition of c-command in terms of dominance together with the multidominance theory of movement, offers to the best of my knowledge the first solution to the puzzle why intermediate categories do not c-command into specifiers. The definition derives this for the special, but for our purposes crucial, case of moved specifiers only.

Section 2.4 then showed how the assumptions developed thus far conspire to give rise to the anti-locality effect. The next chapter is devoted to exemplifying the anti-locality effect first for simple cases (C° and v°). I will then in chapter 4 approach the main empirical issue of the thesis, the ban against preposition stranding, which is analyzed as an anti-locality effect.

Chapter 3: C°, v°, and anti-locality

In the previous chapter I made two claims that become important here.

The first claim was that the Last Resort condition rules out movement from complement to specifier position within the same phrase (and from specifier to specifier position within the same phrase). A complement is already in the closest possible relation with its selecting head (mutual total c-command). Movement from complement to specifier position never gives rise to new feature satisfaction relations and is therefore ruled out. The Last Resort condition rules out all objects of the general form $[X [\alpha X]]$ where α is a head and X some constituent.

The second claim was that given an attractor γ , γ can attract α if there is no closer potential attractee β , where β is closer to γ than α iff β totally c-commands α . Since we know that a head and its complement totally c-command each other, whenever there is a head and its complement, neither of them will ever be closer to any higher attractor than the other. If the head in question is an intervener for its complement, then this definition of closeness has the consequence that the head intervenes for movement of its complement and the complement intervenes for movement of the head.⁵⁴

⁵⁴ The same principle applies also to base merged specifiers under the definitions of the previous chapter. A base merged specifier is in a mutual total c-command relation with its X'-sister. The specifier and the X'-sister may, if both have the relevant feature, give rise to an intervention effect for each other. In such cases the X'-category is immobilized and also the specifier.

In the previous chapter the escape hatch property of specifiers (the edge condition of Chomsky's Phase Impenetrability Condition) was derived by assuming that phase heads are potentially universal attractors and that they are universal interveners. This puts the complements of phase heads in a particular fix. They cannot move into the redeeming specifier, a.k.a. escape hatch/edge, position because of the Last Resort condition. In their base position they cannot be attracted from the outside, because of the intervening phase head. The prediction is that the complements of phase heads should be frozen in place relative to the phase head itself.

When discussing the theory of punctuated movement paths in the previous chapter I noted that the landing sites of successive cyclic movement should have a cluster of properties. The ability to reconstruct to such positions was mentioned as was the potential for floating quantifiers and for having agreement with the relevant phase head. We now have a further prediction: the complements of phase heads should be frozen in place.

The present chapter examines this prediction.

On the assumption that certain *wh*-phrases like *how come* are base generated directly in [Spec, CP], this offers an immediate explanation for the total immobility of *how come* illustrated in (i) and (ii). See Starke 2001:29 for some discussion.

- (i) How come Peter arrived so late?
- (ii) How come you said that Peter arrived so late?
 - ✓What is the reason for your saying that Peter arrived so late
 - *What did you say is the reason for Peter's late arrival.

3.1 TP immobility under C°

In this section I look at the phase head C° and its complement TP. Assuming that C° is a phase head, we predict that TP is immobile as the complement of C°. Indeed, there do not seem to be cases where a TP is separated from its embedding complementizer in a way that would suggest movement. I first illustrate the claim that TPs never strand C°. I then argue that TPs are mobile in principle – if they weren't, the anti-locality constraint would simply be irrelevant. Finally I turn to the question why TPs seem never to be subject to movement processes that would allow them to strand C°. I also briefly discuss Right Node Raising.

3.1.1 TPs do not strand C°

To see that TPs do not strand their embedding complementizers, consider the examples in (106) through (108).

- (106) a. ✓Nobody believes that anything will happen.
 b. ✓That anything will happen, Nobody believes.
 c. *Anything will happen, nobody believes that.
- (107) a. ✓[I' homme que je crois [t qui [t viendra]]
 the man that I know that came
 'the man that I know arrived'
 b. ✓Jean a dit que Pierre a acheté une maison
 Jean has said that Pierre has bought a house
 c. ✓Que Jean a dit que Pierre a acheté ?
 what Jean has said that Pierre has bought?
 d. *Pierre a acheté une maison Jean a dit que/qui.
 Pierre has bought a house Jean has said that.
- (108) a. ✓Hver heldur þú að hafi lesið þessa bók?
 who think you that has_{SUBJ} read this book
 lit: 'Who do you think that has read this book?'
 b. ✓Jón heldur að María sé að lesa.
 Jon thinks that Maria is_{SUBJ} to read_{INF}
 'Jon thinks that Maria is reading.'

- c. ?Að María sé að lesa heldur Jón.
 that Maria is_{SUBJ} to read_{INF} thinks Jon
 'That Maria is reading, John thinks.'
 d. *Maríasé að lesa heldur Jón að.
 Maria is_{SUBJ} to read_{INF} believes Jon that

Examples (106a-b) show that the complement CP of the verbe *believe* may be topicalized, but the embedded TP cannot be topicalized stranding the complementizer *that*. This is shown by the ungrammatical example (106c). The NPI in the embedded clause was introduced to block the reading of *that* where it functions as a demonstrative rather than a complementizer. It might of course be objected that the *that-t* effect could independently rule out example (106c). It is unlikely that the *that-t* effect is the correct explanation for the status of this example, though. First of all, Sobin 1987 has argued that there are dialects of English that do not show *that-t* effects, still example (3d) needs to be ruled out even in those dialects. The problem with (106c) is therefore independent of the *that-t* effect. Second of all, the effect seen here holds in languages that lack *that-t* effects. Thus, if we accept the theory that French allows extraction of preverbal subjects for (107a) and thus does not show *that-t* effects, then the interfering factor noted for English is controlled for.⁵⁵ Icelandic is probably less controversial in not showing *that-t* effects (108a) (Maling and Zaenen 1978), yet Icelandic doesn't allow extraction of TP either, as the examples in (108) show. (The examples were provided by Halldór Sigurðsson, p.c.)

Under certain circumstances, English does not exhibit the *that-t* effect. That-t effects are absent for example if the null complementizer is used (109a). Example (109b)

⁵⁵ See for a treatment of the *que/qui*-alternation in French along these lines Rizzi 1990.

shows that the complementizer under the verb *think* can be either overt or null. Example (109c) shows that the embedded clause with the overt complementizer can be extraposed. Example (109d) then shows that extraposition of TP is impossible. Since this conclusion is not immediately obvious, consider the two possible structures underlying (109d) schematized in (110a) and (110b). The structure in (110a) is presumably ruled out by the conditions that govern the distribution of null complementizers in English (Stowell 1981; for a recent overview Bošković and Lasnik 2002). However, (110b) must also be ruled out. Notice that the *that-t* effect cannot be relevant here, since the null-complementizer does not give rise to that effect.

- (109) a. What do you think has happened?
 b. ✓Nobody thought (that) anything would happen yesterday.
 c. ✓Nobody thought yesterday that anything would happen.
 d. *Nobody thought yesterday anything would happen.
- (110) a. nobody thought t_{CP} yesterday [$_{CP} \emptyset_{C^0}$ anything would happen]
 b. nobody thought [$_{CP} \emptyset_{C^0}$ t_{TP}] yesterday [$_{TP}$ anything would happen]

It could be argued, however, that although the *that-t* effect does not rule out the structure in (110b), Ross's (1967) Right Roof Constraint might be violated in the structure, since the TP would have to be extraposed into the higher clause in violation of the Right Roof Constraint. The examples in (111) and (113) control for this factor. Again there are two structures to consider for the (c) examples. They are given in (112) and (114). The structures in (112a) and (114a) are again ruled out by the conditions that regulate the distribution of null complementizers in English. Clearly, (112b) and (114b) must also be ruled out – as far as I can see, they are only ruled out if we disallow TP from stranding its embedding complementizer.

- (111) a. ✓Nobody realized (that) anything would happen.
 b. ✓That anything would happen, nobody realized.
 c. *Anything would happen, nobody realized.
- (112) a. [$_{CP} \emptyset_{C^0}$ anything would happen] nobody realized t_{CP}
 b. [$_{TP}$ anything would happen] nobody realized [$_{CP} \emptyset_{C^0}$ t_{TP}]
- (113) a. ✓I believe (that) Luisa is a genius.
 b. ✓That Luisa is a genius is widely believed.
 c. *Luisa is a genius is widely believed.
- (114) a. [$_{CP} \emptyset_{C^0}$ Luisa is a genius] is widely believed t_{CP}
 b. [$_{TP}$ Luisa is a genius] is widely believed [$_{CP} \emptyset_{C^0}$ t_{TP}]

Extraposition of TP out of a relative clause is disallowed (115c) in contrast to CP extraposition (115b). The *that-t* effect is not operative in subject relative clauses as (115a) shows. The examples are not so clear cut evidence, however, since they might run afoul of the Right Roof Constraint and possibly involve extraction out of the adjunct formed by the relative clause. Moreover there are analyses (e.g. Kayne 1994) that derive relative clause extraposition without movement of the CP so that we might not even have evidence here for movement of CP in this environment.

- (115) a. ✓Frank saw a play that was long and boring yesterday.
 b. ✓Frank saw a play yesterday that was long and boring.
 c. *Frank saw a play that yesterday was long and boring.

TP cannot strand its embedding complementizer in other contexts in English either (e.g. 116), although there may again be interfering factors in cases like this.

- (116) a. ✓John made the claim that Mary will arrive yesterday.
 b. ✓John made the claim yesterday that Mary will arrive.
 c. *John made the claim that yesterday Mary will arrive.

There is an interesting contrast between TPs and smaller XPs with respect to movement in Yoruba. Dekydtspotter (1992) argues that predicate clefts in Yoruba involve remnant XP movement. In his discussion Dekydtspotter (1992:126) makes the observation that unambiguously inflectional elements (T^0 in my terminology) do not

participate in predicate clefting. In other words, remnant TPs do not front. This is shown in example (117) from Dekydtspotter (1992:127). The elements glossed as Asp are what Dekydtspotter calls unambiguously inflectional elements. On the other hand “[p]reverbs, which are aspectual, modals, and negation, have often been observed to be ambiguous between the inflectional and verbal paradigms. These elements may or may not participate in predicate clefts.” (Dekydtspotter 1992:126). This is shown in example (118) from Dekydtspotter (1992:127). The remnant XPs that are fronted in examples (118) are smaller than those that move in (117). The ones in (117) are TPs. They are immobile according to our generalization.

- (117) a. *Mímáa lə ni mo láa lə sǒjà.
 Asp go NI I Asp go to-market
 ‘I USUALLY go to market.’
 b. *Titi lə ni mo ti lə sǒjà.
 Asp go NI I Asp go to-market
 ‘I ALREADY went to market.’
 (118) a. ✓Títètè lə ni Tolú tètè lə sǒjà.
 quickly go NI Tolu quickly go to-market
 ‘What Tolu did is go quickly to market.’
 b. ✓Lílə ni Tolú tètè lə sǒjà.
 going NI Tolu quickly go to-market
 ‘What Tolu quickly did is go to market.’

If Dekydtspotter’s classification of the various elements is correct, then there is a clear example here where TP must be treated differently from other, smaller projections in exactly the way the present theory leads us to expect.

To sum up. There are some clear cases, where the ban against stranding C° needs to be invoked (106-108, 111, 113, 117) since the examples are not independently ruled out. The other cases discussed might run afoul of independent constraints. The generalization that TPs never strand their embedding complementizers is cross-

linguistically extremely stable. A clear counterexample would be a case where the TP and the complementizer are not adjacent to each other in the surface string, i.e. if they are separated by material that clearly belongs to a different clause. I know of no such clear counterexamples.

3.1.2 TPs are mobile in principle

Given the generalization that TPs never strand their embedding complementizer, it is legitimate to ask the question whether TPs ever move. I argue in this subsection that in fact they do. Consider example (119).

- (119) a. ✓How likely to win the race is John?
 b. ✓How likely is John to win the race?

The infinitival *to win the race* is a raising infinitive. I argued in the previous chapter that raising infinitives are TPs not CPs.

I assume further that *to win the race* is generated as the complement of *likely*. In Abels 2002b I provide some arguments for this assumption, which is the standard assumption in any case (but see Boeckx 2002; Kuno 2003; Rosenbaum 1967 for the opposite view and Abels 2002b for counterarguments). On this assumption it follows that (119b) can only be derived if the TP *to win the race* moves out of its position as the sister of *likely* at some point. This is so because the base generated constituent [*likely to win the race*] is disrupted in the surface order. The complement clause must move out to effect the separation of *likely* from its complement. This movement might either take the form of rightward extraposition or it might involve leftward movement of *to win the race* followed by remnant *wh*-movement of *how likely* t_{TP} .

Movement of the TP has syntactic ramifications as the contrast between (120a) (for which see Kroch and Joshi 1985; Lasnik and Saito 1992 among others) and (120b) indicates. The contrast between (120a) and (120b) in fact indicates that TP movement in these examples happens in the syntax rather than, say, at PF. PF movement (or scattered deletion of the chain *<how likely to be a riot, t_{how likely to be a riot}>* as in Nomura 2001) is unlikely to be able to account for the difference in grammaticality seen in (120a) and (120b). It seems that the expletive *there* cannot move out of the embedded clause if *how* intervenes.⁵⁶ This is illustrated in (121).

- (120) a. *How likely to be a riot is there?
 b. ✓How likely is there to be a riot?
 (121) a. ?*Who said that there is how likely to be a riot?
 b. ✓Who said that it is how likely that there will be a riot?

Example (121a) illustrates again the intervention effect caused by *how* – the example is perfectly grammatical without *how*, of course – and in (121b) the same meaning is expressed, but *there* does not have to cross *how* and the example becomes acceptable. In (120b) the embedded clause *there to be a riot* is moved above *how*. This voids the intervention effect thus accounting for the status of (120b). To account for the fact that (120a) is ruled out, we have to assume that string vacuous extraposition of the embedded infinitival is blocked in (120a). In Abels 2002b I give a principled account of this fact.

I conclude from this brief discussion that there are examples of TP movement.

⁵⁶ If expletives do not move, then some other way needs to be found to capture the intervention effect caused by *how*. For further discussion see Abels 2002b; Boeckx 2002; Kuno 2003. The latter reports that his informants did not find a significant difference between (121a) and (121b).

3.1.3 Some non-cases of TP-movement

I will review a few putative cases of TP movement in this subsection. I first discuss Right Node Raising and argue that it does not furnish examples of TP movement. I then turn to gerunds to infinitivals with PRO subjects. The situation here is less clear. A clearcut case for TP movement cannot be based on gerunds and infinitivals with PRO subjects.

Right Node Raising (RNR) appears to furnish *prima facie* evidence for TP movement. Nevertheless, I claim here that RNR does not involve TP movement. If RNR involved ATB movement, as is often assumed, then RNR would furnish examples not only of TP movement but in fact of C-stranding as examples (60), repeated here, and (123) from Postal 1998:177, 106 show. New evidence against the ATB analysis of RNR will be provided.

- (60) a. ✓John believes that and Peter claims that – Mary will get a job
 b. ✓I know when but I don't know where – Amanda met Steve
 (122) John believes that and Peter claims that – Mary will get a job
 (123) I know when but I don't know where Amanda met Steve

There is a large body of literature on RNR (e.g. Bošković no year; Bresnan 1974; Goodall 1987; Hudson 1976; Kayne 1994; Levine 1985; Maling 1972; McCawley 1987; McCloskey 1986b; Postal 1974, 1993, 1998; Ross 1967; Sabbagh 2003; Wexler and Culicover 1980; Wilder 1997b, 1999). According to the across the board (ATB) movement analysis of RNR (Ross 1967 also Bresnan 1974; Hudson 1976; Maling 1972; Postal 1974, 1993, 1998), the noun phrase *my best friend* in (124) ATB-moves out of both conjuncts as shown in (124a). The *in situ* analysis on the other hand (Wexler and Culicover 1980; Kayne 1994) assumes that *my best friend* does not ATB move in RNR contexts. I will argue for the latter position as elaborated in Bošković (no year), where it

is assumed that RNR involves backward ellipsis in the first conjunct as shown in (124b) (see also Wilder 1997b, 1998).

- (124) John admires and Mary hates my best friend.
 a. [[John admires t_{DP}] and [Mary hates t_{DP}]] [$_{DP}$ my best friend]
 b. John admires ~~my best friend~~ and Mary hates my best friend

Example (125) shows the structure for (60) under the ATB analysis of RNR.

- (125) [[John believes [$_{CP}$ that $_{TP}$]] and [Peter claims [$_{CP}$ that $_{TP}$]] [$_{TP}$ Mary will get a job]

Structure, if correct, would be a clear counterexample to the Stranding Generalization. Defending the Stranding Generalization commits me to assuming that RNR does not involve ATB-movement. Rather I have to side with Bošković (no year; Kayne 1994; Wexler and Culicover 1980; Wilder 1997b; 1998; 1999 and others in assuming that RNR involves some kind of ellipsis.

It might be worth pointing out that this does not amount to saying that RNR never involves any kind of movement. In fact it doesn't even amount to saying that RNR never involves ATB-movement. It only says that not all cases of RNR involve ATB-movement (Bošković no year).⁵⁷ I will not review the arguments for the deletion approach from the literature here. I will simply add two new arguments both of which lead to the conclusion that RNR involves ellipsis rather than ATB-extraction.

⁵⁷ It has been observed numerous times that RNR is subject to certain constraints on movement (Bošković no year; Postal 1998; Sabbagh 2003; Wilder 1998). What's at issue is not whether RNR may involve movement, but whether it involves ATB extraction. The alternative to ATB extraction being movement within each of the conjuncts. Postal 1998 does not seem to appreciate this point properly.

My first argument takes one of the strongest arguments for the ellipsis analysis and what is probably the single strongest argument for the ATB analysis and investigates how they interact with each other.⁵⁸

Proponents of the ellipsis analysis of RNR never fail to point out that RNR is island insensitive (Goodall 1987; Levine 1985; Wexler and Culicover 1980). This is illustrated in If RNR were movement, this would be extremely surprising. Example (126) illustrates the adjunct condition for *wh*-movement and (127) shows that RNR is exempt from it. This is even more surprising in light of the fact that RNR is rightward movement, which is otherwise clausebounded (Ross's 1967 Right Roof Constraint). Such examples are *prima facie* evidence against the ATB analysis of RNR. Proponents of the ATB analysis have to explain somehow why regular constraints on movement are irrelevant to RNR.

- (126) a. *What did Mary sell many diamonds before John had seen t_{what} ?
 b. *What did John eat lunch after Mary had prepared t_{what} ?
 (127) a. ✓Mary sold $_{TP}$ before John had seen $_{TP}$ - many diamonds from South Africa.
 b. ✓John ate $_{TP}$ after Mary had prepared $_{TP}$ - a five course meal in the Italian style.

On the other hand there are cases where ATB movement in RNR constructions seems unavoidable for semantic reasons. These examples form the strongest and most interesting *prima facie* argument for the ATB movement analysis of RNR. In such examples the constituent undergoing RNR cannot be adequately interpreted if RNR involves ellipsis without ATB movement as (128-130) show.

⁵⁸ I would like to thank Howard Lasnik for discussing and developing an earlier version of the idea reported here.

- (128) a. ✓John sang and Mary beat the drums – at equal volumes.
 b. *John sang at equal volumes and Mary beat the drums at equal volumes.
 (129) ✓John sang and Mary recorded {the same song | very different songs}
 (130) a. ✓John sold and Mary bought – gold rings and diamonds from South Africa respectively.
 b. *John sold gold rings and diamonds from South Africa respectively and Mary bought gold rings and diamonds from South Africa respectively.

Postal (1998) following others points out that such examples militate strongly against an analysis that never allows ATB movement to take place in RNR. The phrases *at equal volumes*, *the same song*, *different songs*, and *gold rings and diamonds from South Africa respectively* cannot be adequately interpreted in situ.⁵⁹ They must take scope outside of the conjunction. Clearly, there are semantic reasons to assume that ATB movement is involved here.

These examples argue for the ATB analysis of RNR, since that analysis directly provides interpretable structures for these examples, whereas the ellipsis analysis does not. In fact it is unclear what the source of ellipsis might even be. The ellipsis analysis, the further claim goes, doesn't make interpretable structures for such examples available at all. The last part of the argument is not strictly speaking true, though. Although it is true that the ellipsis analysis is forced to admit a certain amount of ATB movement in these cases – either overtly, prior to ellipsis, or covertly, after ellipsis.⁶⁰ If we admit the

⁵⁹ It is not entirely clear how these examples are interpreted at all. It seems possible to adapt ideas from Beck 2000 and Gawron and Kehler to appear to the purpose though.

⁶⁰ The discussion in the text does not commit me to a particular view as to the nature and timing of the movement involved in interpreting examples like (128a) and (130a). It could be that overt rightward ATB movement is involved, it could be that there is no overt movement even in these examples, the interpretive requirements of the

possibility that there might be ATB movement in just these cases (essentially driven by interpretive needs), then the two approaches make strikingly different predictions. On the ellipsis approach sentences that require ATB movement for interpretive reasons should show island sensitivity. The ATB approach makes no such prediction as no additional movements are necessary to interpret the structures and the putative movement involved in deriving RNR is not subject to islands.

We can test these predictions by modifying examples (128a-130a) minimally. This is done in examples (131-133). The examples are strongly ungrammatical. The facts thus favor the ellipsis approach over the movement approach. The facts in (128a) and (130a) when scrutinized carefully provide an argument against rather than for the ATB movement analysis of RNR.

- (131) *John sang __ after Mary beat the drums __ – at equal volumes.
 (132) *John sang __ after Mary had recorded __ {the same song | very different songs}
 (133) *John sold __ before Mary bought __ – gold rings and diamonds from South Africa respectively.

The same argument can be replicated with other types of islands as well. An example is the paradigm in (134) which involves a complex NP island.

- (134) a. ✓John knows a man who sells and Mary knows a woman who buys – gold rings and diamonds from South Africa.
 b. *John knows a man who sells and Mary knows a woman who buys – gold rings and diamonds from South Africa respectively.

examples being satisfied by covert movement. Given a theory of overt and covert movement these options are easily distinguishable (although the examples become very complex very fast), but nothing in this dissertation hinges on the answer to this question.

These examples are strongly incompatible with the assumption that RNR always involves ATB movement. They validate the traditional argument from island insensitivity against the ATB analysis of RNR.

My second argument relies even more directly on the properties of the ATB and the ellipsis analysis. Generally ATB constructions are compatible with ellipsis in the second conjunct. Example (135) is a regular example of ATB extraction. Example (136a) is different only in that VP ellipsis has applied in the second conjunct in (136a) but not in (135). The structure of (136a) is sketched in (136b).

- (135) ✓Who did you say that John had visited long ago but that Mary hadn't visited until yesterday?
 (136) a. ?Who did you say [[_{CP} that John had [_{VP} visited _{t_{who}}] long ago] but [_{CP} that Mary hadn't <[_{VP} visited _{t_{who}}] until yesterday]]?
 b. [WH ... [_{CP} ... [_{VP} ... _{t_{WH}}]] but [_{CP} ... <[_{VP} ... _{t_{WH}}] >]]

Consider now the structures in (137). The structure in (137a) shows that on the ATB analysis of RNR VP ellipsis in the second conjunct should leave the target of RNR unaffected. The target of RNR is located above conjunction, considerably above the ellipsis site. Structurally though not linearly, (137a) is exactly parallel to (136b). In (136b), too, the target of ATB movement, the *wh*-phrase, remains under ellipsis. The prediction of the ellipsis analysis of RNR is quite distinct, though. The structure is given in (137b). Since the target of ATB never moves out of its containing VP, the ATB target will be elided along with the VP. The question then is whether VP ellipsis can 'spare' a target of RNR.

- (137) a. [... [... [_{VP} ... _{t_{RNR}}]] and [... <[_{VP} ... _{t_{RNR}}] >]] RNR
 b. [... [... [_{VP} ... <RNR>]] and ... [... <[_{VP} ... RNR] >]]

The relevant examples are given in (138). Example (138a) simply is a control showing that RNR of *the students in the Wednesday section* is possible in this context. Example (138b) is an attempt to do VP ellipsis sparing the target of RNR. The sentence is sharply ungrammatical. Again the prediction of the ellipsis analysis is borne out and the prediction of the ATB analysis of RNR fails.

- (138) a. ✓Simon thinks that Peter is proud of, but Luisa doesn't think that Peter is proud of – the students in the Wednesday section.
 b. *Simon thinks that Peter is proud of, but Luisa doesn't – the students in the Wednesday section.

I conclude from these argument and the arguments from the prior literature that the ATB analysis of RNR is all but untenable. RNR therefore does not furnish examples of TP movement or, for that matter, of C°-stranding.

Consider now another potential case of TP movement. It has been claimed that gerunds in English are TPs (Munn 1991; Stowell 1982b claim that gerunds are IPs) and gerunds move quite freely, as the topicalized gerund in (139) illustrates. Thus, if gerunds are really TPs, then we have here cases of TP moving.

- (139) ✓Andrew talking about reduplication, I enjoyed most at the conference.

Stowell observes that gerunds never appear with an overt complementizer (140) and do not seem to have a position for relative operator movement (141).

- (140) I have heard about {✓∅ | *that | *for} him leaving office.
 (141) *the man {that | which | ∅} {him | PRO} seeing

Stowell takes such facts to indicate that gerunds are not CPs. Fair enough, but how do we know they are TPs? The two facts, namely the greater mobility of gerunds and the impossibility to occur directly under a complementizer, can be related if gerunds

are NPs/DPs rather than TPs as argued by many (Chomsky 1970, Abney 1987, Grimshaw 1991⁶¹, Corver 1990 and many others). A clear argument for TP movement can at the moment not be based on the mobility of gerunds.

Infinitives with PRO as the subject are another possible instance of mobile TPs. Bošković (1997b) argues for this position. Stowell (1981) had argued that the distribution of null complementizers (null *that* and null *for*) can be derived from the ECP. But in standard English infinitivals with PRO as the subject never occur with an overt complementizer, even in extraposed or clausal subject position where null *for* and null *that* are barred. It stands to reason, Bošković argues, that there are no ECP effects with the complementizer in control infinitivals simply because there are no complementizers in control infinitives, i.e. because control infinitivals are TPs.⁶²

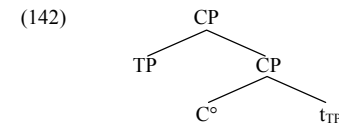
The argument is only as strong as the ECP based approach to the distribution of null complementizers that it rests on. In a recent manuscript Bošković and Lasnik (2002) collect a large amount of data showing that the ECP based approach to the distribution of null complementizers is inadequate. Again there is no overwhelming argument at the moment to the effect that infinitivals with PRO are TPs.

⁶¹ The discussion of gerunds is not included in the excerpt of Grimshaw 1991 published as Grimshaw 2000.

⁶² Bošković thus emulates an argument present already in Stowell 1982b. Stowell argues against the presence of C° in gerunds based on his (1981) ECP account of null complementizers, but because of the then current theories of government and PRO he couldn't extend it to control infinitives. The null Case theory of the distribution of PRO (Martin 1996) allows Bošković to generalize Stowell's argument.

3.1.4 Consequences

One of the main consequences of the approach pursued here has to do with the existence of the head parameter. Kayne (1994) suggested that there is no head parameter and that the order of constituents is uniformly Spec>Head>Complement. For clause final complementizers he suggested that they are derived through short movement of TP to [Spec, CP] as in (142). This sort of short movement is, of course, ruled out under the present approach.



Kayne (1994:53-54) offers three potential beneficial consequences of his approach. First, he claims that *that-trace* effects exist only in languages that have initial complementizers but are systematically absent in complementizer final languages. Kayne claims that this might have to do with the fact that on his approach C° c-commands a trace in subject position in complementizer initial languages but not in complementizer final languages. If the right theory of *that-trace* effects is indeed structural in the sense that *that-trace* effects require overt c-command by C°, then the absence of *that-t* effects in complementizer final languages would follow. Some facts concerning *that-trace* effects point towards a more linear explanation of the effect, however. Particularly telling is the observation going back to Bresnan (1977) that embedded topics void the *that-trace* effect. This is exemplified in (143). The examples are taken from Pesetsky and Torrego (2001:376)).

- (143) a. Sue met the man who Mary is claiming that [for all intents and purposes] t_{who} was the mayor of the city.
 b. Bill, who Sue said that [to the rest of us] t_{who} might seem a bit strange, turned out to be quite ordinary.

If Bošković and Lasnik (2002) are right, then the *that-trace* effect should receive an explanation in terms of linear adjacency. The lack of linear adjacency would then explain the lack of *that-trace* effect in example (143). But if the *that-trace* effect indeed relies on linear adjacency, then it follows immediately that head final languages will never show such effects independently of the existence of the head parameter.

Pesetsky and Torrego (2001) on the other hand offer a structural account of the *that-trace* effect. The account is formulated in terms of T-to-C movement, where TP is generally the complement of C° . Pesetsky and Torrego's account of the *that-t* effect is too complicated to recount it here. However, to block TP from moving to [Spec, CP] in English, they need to invoke (p. 363) the condition that whenever a head H attracts a feature of XP as part of a movement operation, then "if XP is the complement of H, copy the head of XP into the local domain of H". In other words, movement from complement to specifier position of the same head (as would be required by Kayne) is systematically excluded. Pesetsky and Torrego's structural account of the *that-t* effect is thus incompatible with Kayne's account of clause-final complementizers.

Ishii 1997, 1998 offers an account of the *that-t* effect in terms of anti-locality. Ishii's anti-locality, unlike the one pursued in the present thesis, is not based on primitive phrase structure relations but on the notion of minimal domain. Ishii suggests that subjects are adjoined to TP (as in Kayne 1994) and that they are thereby in the minimal domain of the embedding complementizer. Movement within a single minimal domain

can then be ruled out by last-resort. Ishii further assumes that there is no null-complementizer in English and analyzes the cases that are traditionally treated as involving null complementizers as bare TPs. This accounts for the *that-t* effect – but it crucially rules out movement from complement to specifier position.⁶³

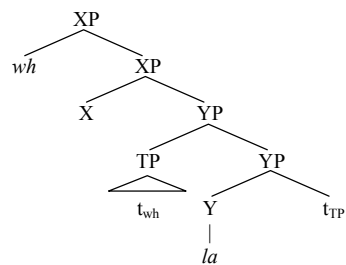
I know of no approach to *that-trace* effects that is both structural and compatible with Kayne's (1994) treatment of final complementizers. Lacking an account of the *that-trace* effect that relies on structural configurations and is compatible with Kayne's (1994) approach to final complementizers, we conclude that the first of Kayne's possible beneficial consequences is, at present, just that: a possible beneficial consequence. No support for Kayne's point of view can be drawn from it until Kayne's suggestion is cashed out.

The second possible beneficial consequence of Kayne's approach comes from the area of anaphors in subject position. Kayne offers the generalization "that nominative anaphors in subject position are excluded from languages (or constructions) in which complementizers are initial" (p. 54). Again there is no account of this purported generalization relying on c-command by the complementizer. Therefore, no strong support can be garnered from it. Moreover, the generalization seems to be wrong since, as noted by Kayne in the discussion leading up to the quote, Chinese allows subject anaphors, but at the same time Chinese has initial complementizers.

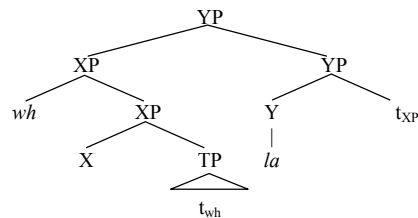
⁶³ Given the existence of languages like Icelandic that appear to lack *that-t* effects, it seems to me that the explanation should not be quite as deep.

Finally, Kayne notes that uniformly head final languages typically do not have obligatory overt *wh*-movement. This is supposed to be explained under the structure in (142), because the landing site of *wh*-movement is filled by TP, effectively blocking *wh*-movement to [Spec, CP] from taking place. In a footnote Kayne (1994:142 fn. 22) mentions Vata as a possible counterexample. “In root *wh*-questions in Vata (see Koopman 1984, pp.35, 89) the *wh*-phrase is initial; they also have a clause-final element *la* that the text proposal must analyze as not being the head of the projection in whose specifier the *wh*-phrase is found. The head whose specifier does contain the *wh*-phrase must have an unmoved complement.” The structure Kayne suggests for Vata then looks something like (144) or (145).

(144)



(145)



The assumption that such structures are available clearly undermines the prediction seemingly entailed by (142). In fact, just two pages before the discussion of final complementizers Kayne discusses German and Dutch and asserts that “[s]ince German and Dutch show person agreement, it would follow that in those languages [...] V raises to Agr_S” (p. 52). Furthermore, since in German and Dutch arguments of the verb appear to the left of the verb (in Agr_S), these arguments must all be higher than Agr_S. There is then a multitude of positions like XP and YP in (144) or (145) that are above Agr_SP (or TP) but below CP. The prediction derived from (142) rests on the assumption that there is only one unique [Spec, CP]-like position available, but Kayne, as we just saw cannot maintain this assumption, therefore the prediction is no longer made.

We see then that the independent motivation for the particular movement analysis of final complementizers given in (142) is rather weak. In fact, Kayne himself obviously shares this assessment. In subsequent work (Kayne 1999) he gives up the analysis in (142) without as much as mentioning the three putative benefits of the analysis just reviewed. Richard Kayne (p.c.) tells me that part of the reason for giving up the 1994 analysis was precisely the fact that short movement from complement to specifier of the same head is difficult to motivate.

There are also some genuine drawbacks of Kayne’s 1994 analysis of final complementizers. The first is the fact, driving this thesis, that movement from complement to specifier position is hard to motivate. The second drawback is that the analysis requires a kind of movement, movement of IP/TP to [Spec, CP] which is otherwise totally unattested (see section 3.1.3 and in particular the Appendix to this

chapter). Third, if we continue to assume that [Spec, CP] is somehow implicated as an escape hatch in successive cyclic movement, then Kayne's 1994 analysis of final complementizers begs the question why head final languages do not allow complementizer stranding by IP/TP, i.e. why is there no long distance scrambling of IP/TP in Japanese for example. After all, IP/TP routinely moves under that analysis and it moves not just anywhere but precisely to the escape hatch position.

Kayne 1999 offers an analysis of complementizers which is quite different from that given in Kayne 1994. Under Kayne's (1999) analysis a verb that under standard assumptions takes a CP complement, actually takes an IP complement (146a). The VP thus formed is merged with the complementizer (146b), which attracts the IP into its specifier position (146c). The order of elements is now IP>C°>V, the reverse of what it should be. To remedy this, a word order head W° is merged (146d). C° head-moves to W° (146e) and finally the remnant VP moves to [Spec, WP] yielding the correct word order V>C>IP (146f).

- (146) a. [_{VP} V IP_{VP}]
 b. [C° [_{VP} V IP_{VP}]]
 c. [_{CP} IP [C° [_{VP} V t_{IP} VP]] CP]
 d. [W° [_{CP} IP [C° [_{VP} V t_{IP} VP]] CP]]
 e. [[_{W°} C° W°_{W°}] [_{CP} IP [t_{C°} [_{VP} V t_{IP} VP]] CP]]
 f. [_{WP} [_{VP} V t_{IP} VP]] [[_{W°} C° W°_{W°}] [_{CP} IP [t_{C°} t_{VP}] CP]] WP]

Final complementizers in this system would presumably be derived by omitting the head movement from C° to W° (146e) and fully head final structures might be derived by also omitting movement of the remnant VP to [Spec, WP] (146f), i.e. such a derivation could stop at (146c) or (146d). (See Borsley 2001 for critical discussion.)

The prediction that Kayne 1994 was trying to make on the basis of (142), namely that *wh*-movement is blocked in languages with final complementizers because [Spec, CP] is filled, can no longer be even attempted in this system; [Spec, CP] is filled by IP in clauses with head final and in clauses with head initial complementizers. If there is an escape hatch position for successive cyclic movement it cannot be [Spec, CP]. In fact, an element on the path of long distance *wh*-movement no longer enters into a local relation with C° at all under this system. The system therefore leaves traditional cyclicity effects unexplained: Why do complementizers in Irish agree along the path of *wh*-movement but not otherwise if the moving *wh*-element never enters into a local relation with the complementizer?⁶⁴ Why do partial *wh*-movement patterns target [Spec, CP] across languages? Why can *all* be floated exactly in the traditional [Spec, CP] position in West Ulster English (McCloskey 2000)? Why should IP be frozen in place after its movement to [Spec, CP]? Why should we find a pattern like the one discussed in section 2.1, where the absence of a CP projection correlates with the absence of an intermediate landing site? In other words, to the extent that there is a clustering of properties indicating that CP is involved in mediating long distance movement relations, it is unclear whether or how Kayne's 1999 approach can capture these effects let alone relate them to each other in a natural way. The more traditional approach captures and relates these effects by assuming that long distance *wh*-movement passes through [Spec, CP]. The present thesis is an

⁶⁴ See Noonan 1997 for an LCA compatible approach to this particular example.

attempt to add the anti-locality effect to the list of known cyclicity phenomena and at the same time to relate all of them.

To the extent that the present project is successful, it provides an empirical argument against both of Kayne's accounts of complementizers. It is sometimes assumed that it should be possible to disprove the LCA in terms of undergeneration, but the real test for the research program that takes Kayne's LCA as its starting point is its predictive capacity, not its descriptive capacity. One of the claims of the present thesis, one implicit up to this point, is that the LCA fails as a predictive tool.

3.1.5 Conclusion

In this section I have argued that TPs are mobile in principle, but that they never move away from an embedding complementizer. Complementizer stranding appears to be universally disallowed. This follows from the combination of the Last Resort condition, which entails anti-locality effects, together with the present theory of escape hatches and closeness.

We know independently that movement from out of TP past the embedding complementizer is possible. Any example of long distance *wh*-movement will do to illustrate this: *Which man did John say [CP that [TP Mary likes t_{which man}]]?* We thus reach the generalization (147), which is intended to be read as follows. (147a) states that extraction from out of the complement domain of C° is in principle possible (but not always). (147b) states that extraction of the complement of C° is never possible. C° therefore shows exactly the expected behavior under the anti-locality prediction developed in the previous chapter.

- (147) a. $\Diamond\checkmark[C^\circ [\dots t \dots]]$
 b. $\neg\Diamond\checkmark[C^\circ t]$

I also discussed Kayne's 1994 theory of clause final complementizers and showed that it lacks strong motivation. I briefly discussed Kayne's 1999 theory of complementizers and claimed that it is ill equipped to predict cyclicity effects involving complementizers.

3.2 VP immobility under v°

Chomsky (2000) claims that vP is a phase on a par with CP. If this is correct, then we expect v° to give rise to anti-locality effects the same way that C° does. If v° heads a phase, there should be a cluster of properties associated with its status as a phase. Indeed (Nissenbaum 2001) argue that reconstruction to the edge of vP is possible. Bruening (2001) argues that in Passamaquoddy there is verbal agreement triggered by *wh*-movement, presumably through [Spec, vP]. Barbiers 1995 discusses data that suggest that focus particles can be 'floated' in [Spec, vP] under successive cyclic movement in Dutch. There is fairly good evidence then that vPs are phases on a par with CPs. Given that, we expect v° to give rise to anti-locality effects.

The structure of this section is straightforward. In subsection 3.2.1 I show, based on German data, that VPs are movable in principle. In subsection 3.2.2 I use English data to show that VP, when embedded immediately under v°, cannot move away stranding v°. This leads immediately to the conclusion that v° behaves exactly as expected under anti-locality.

3.2.1 VPs are mobile

Initial evidence for the claim that VPs move comes from the possibility of passive and unaccusative VPs to undergo VP-topicalization (148-151). Passive and unaccusative VPs are often, though not uncontroversially (see e.g. Legate 2003⁶⁵), assumed to be VPs rather than vPs. If they really are VPs, then (148-151) are *prima facie* evidence for the movability of VP.

- (148) John was [_{VP} arrested].
 (149) Arrested John certainly was _{VP}.
 (150) Read easily the book certainly does.
 (151) Freeze solid the river did.

Given the controversial status of the relevant constituents as VPs, I will turn to German for clearer evidence.

Consider a few German paradigms that illustrate the various properties of restructuring infinitivals in German. Wurmbrand 2001 argues that there are four different types of infinitivals: Lexical Full Restructuring, Functional Full Restructuring, Reduced Non-Restructuring, and Non-Restructuring. I ignore the distinction between Lexical and Functional Full Restructuring since it plays no role in my argument. These classes are distinguished by the size of infinitival constituent. The most important claim for the moment is that Full Restructuring infinitivals are extremely small: they are VPs; Reduced Non-Restructuring infinitives are vPs or TP; Non-Restructuring infinitivals are CPs. The argument for treating Full Restructuring infinitives as VPs comes from the fact that they

do not seem to contain their own accusative Case assigning vP. This makes them compatible with the long passive construction as illustrated in (152) and (153). In the long passive construction the matrix verb is passivized and this leads the object of the embedded verb to become the nominative subject of the whole clause. The lack of vP entails the absence of higher functional material such as negation and tense. It is then predicted that full restructuring infinitivals cannot be negated and cannot bear their own tense modification. This is illustrated in (154) and (155). Finally, Full Restructuring infinitivals allow fronting of an embedded pronoun across the matrix subject (156). All the data in (152) through (165) are taken from Wurmbrand (2001).

The argument for the mobility of VPs will obviously come from Full Restructuring infinitivals.

- (152) ✓...weil ihm Peter diesen Turm zu besichtigen empfohlen hat
 because him Peter this_{acc} tower to visit recommended has
 '...because Peter recommended that he visit this tower.'
 (153) ✓...weil ihm dieser Turm zu besichtigen empfohlen wurde
 because him this_{nom} tower to visit recommended_{pass} was
 '...because it was recommended to him that he visit this tower.'
 (154) *...weil ihm dieser Turm morgen zu besichtigen empfohlen wurde
 because him this tower tomorrow to visit recommended was
 intended: '...because it was recommended to him that he visit the tower
 tomorrow.'
 (155) weil ihm dieser Turm nicht zu besichtigen empfohlen wurde
 because him this tower not to visit recommended was
 ✓'...because it was not recommended to him that he visit this tower.'
 *'...because it was recommended to him that he not visit this tower.'
 (156) ✓weil es ihr der Hans zu essen empfahl.
 because it her the Hans to eat recommended
 '...because Hans recommended to her that she eat it.'

⁶⁵ Legate's main argument for the assumption that passive and unaccusative verb phrases are phases, i.e. vPs, comes from the possibility to do scope reconstruction. These facts can be handled without assuming intermediate traces as discussed in fn. 5.

Other classes of infinitivals behave differently. In reduced non-restructuring infinitives (157), long passive is impossible (158), but embedded temporal modification (159), embedded negation (160) temporal modification, and pronoun movement (159-160) out of the infinitival are possible.

- (157) ✓weil der Hans den Traktor zu reparieren beschlossen hat
 because the Hans the tractor to repair decided has
 '...because Hans decided to repair the tractor.'
- (158) *weil der Traktor zu reparieren beschlossen wurde
 because the_{nom} tractor to repair decided_{pass} was
 '...because it was decided to repair the tractor.'
- (159) ✓weil ihn der Hans morgen zu reparieren beschlossen hat.
 because it the Hans tomorrow to repair decided has
 '...because Hans decided to repair it tomorrow.'
- (160) weil ihn der Hans nicht zu reparieren beschlossen hat.
 because it the Hans not to repair decided has
 ✓'...because Hans didn't decide to repair it.
 ✓'...because Hans decided not to repair it.'

Finally in non-restructuring infinitives (161), long passive is impossible as before (162), embedded temporal modification (163) and negation (164) are possible, but pronoun fronting is impossible (165).

- (161) ✓dass der Hans bedauert den Traktor repariert zu haben.
 ✓that the Hans regrets the tractor repaired to have
 'that Hans regrets having repaired the tractor.'
- (162) *dass der Traktor repariert zu haben bedauert wurde
 that the tractor repaired to have regretted was
 intended: '...that they regretted that they had repaired the tractor'
- (163) ✓dass der Hans bedauerte ihn gestern repariert zu haben
 ✓that the Hans regretted it yesterday repaired to have
 '...that Hans regretted having repaired it yesterday'
- (164) ✓dass der Hans bedauerte ihn nicht repariert zu haben
 that the Hans regretted it not repaired to have
 '...that Hans regretted not having repaired it'
- (165) *dass ihn Hans repariert zu haben bedauerte
 that it Hans repaired to have regretted
 intended: '...that Hans regretted having repaired it'

We can summarize this rather complex pattern in the following table.

Type	Restructuring	Reduced Non-Restructuring	Non-Restructuring
Syntactic category	VP	TP	CP
Long passive	✓	*	*
Embedded tense	*	✓	✓
Embedded negation	*	✓	✓
Pronoun movement	✓	✓	*

These results are all taken from Wurmbrand's work. She argues that restructuring infinitives are VPs not vPs based on the fact that they do not assign structural Case to their object. Recall examples (162) and (163). Example (162) simply shows the verb *empfehlen* – 'recommend' taking an infinitival complement. Example (163) shows that when the matrix verb, *empfehlen*, is passivized but the embedded verb *besichtigen* – 'to visit' is not, the direct object of the embedded verb can become the matrix subject, marked nominative and inducing verb agreement. The fact that passivization of the matrix verb affects the Case of the embedded object, suggests that in the active sentence (162), the embedded object is assigned accusative Case by the matrix verb. Assuming that v° is the accusative Case assigner, we can account for this pattern, following Wurmbrand, by assuming that the embedded verb *besichtigen* is not associated with a v° in (162), i.e. the embedded verb heads a VP and there is no vP projection.⁶⁶

⁶⁶ See Wurmbrand (2001) for extensive discussion. The situation is actually a bit more complicated than the way I portray it here, because, as Wurmbrand argues, *empfehlen* may take a VP complement but it need not. Crucially, long passive as in (153) is possible only if *empfehlen* takes a VP complement.

Crucially, long passive is compatible with topicalization of the embedded verb phrase (166). But since long passive is incompatible with the embedded projection being bigger than VP, (166) shows that VPs can move.

- (166) ✓Zu besichtigen wurde ihm nur dieser Turm empfohlen.
 To visit was him only this tower recommended.
 ‘This is the only tower that he was recommended to visit.’

The same point is also made by example (167). If we find a fronted infinitival that is compatible with pronoun movement, this fronted infinitival can only be a VP. It cannot be a CP, since CPs are never compatible with pronoun fronting (see the table above). And it cannot be a TP since TPs never front (see discussion above in subsection Appendix).

- (167) ✓Zu besichtigen hat ihn ihm der Peter empfohlen.
 To visit has it him the Peter recommended
 ‘Peter recommended him to visit it.’

These examples thus provide evidence for the claim that VPs can move in principle.

3.2.2 VPs do not strand v°

I will now demonstrate the ban against moving VP away from under vP on the basis of English VP-topicalization. (Note that I use the term VP-topicalization as a cover term for topicalization of VPs as in the unaccusative and passive examples above and of vP as in the case of transitive verbs.) The crucial examples come from Huang 1993. The anaphor *himself* in (168a) cannot take the matrix subject *John* as its antecedent, but when

A question might be raised regarding the infinitival marker *zu*. This marker shows puzzling behavior since its presence or absence does not correlate with the other syntactic and semantic diagnostics developed by Wurmbrand. I will ignore it.

the object is topicalized as in (168b), the anaphor can take the matrix subject as its antecedent.

- (168) a. John_i said that Bill_j likes pictures of himself_{*i/j}.
 b. John_i said that pictures of himself_{i/j} Bill_j likes.

Huang observes that the pattern we find with VP-topicalization is strikingly different. Again, when the VP is in its base position, the anaphor *himself* cannot take the matrix subject as its antecedent (169a) and (170a), but this situation surprisingly does not change substantially under VP-topicalization (169b) and (170b). Huang argues that this can be explained under the VP-internal subject hypothesis, since under that hypothesis the fronted constituent contains a trace of the subject which will serve as the antecedent for the anaphor, whether the VP is topicalized or not.

- (169) a. John_i said that Bill would certainly wash himself_{*i/j}.
 b. John_i said that wash himself_{i/j} Bill_j certainly would.
 (170) a. *John_i said that Mary would like pictures of himself_i.
 b. ?*John_i said that like pictures of himself_i Mary would.

The current standard formulation of the VP-internal subject hypothesis claims that the subject is actually generated in the specifier position of a VP-external projection, [Spec, vP] (see Bowers 1993; Chomsky 1995c; Kratzer 1996 for proposals along these lines). For Huang’s argument to work, the structure of (169b) must be (171), where the topicalized constituent is vP, rather than (172), where it is VP.

- (171) John said that [[_{VP} t_{Bill} v^o [_{VP} wash himself]]] Bill certainly would t_{VP}
 (172) John said that [[_{VP} wash himself] Bill certainly would [_{VP} t_{Bill} v^o t_{VP}]]

A possible objection to an account of (169) in terms of anti-locality is to say that V^o undergoes overt head movement to v^o in English. On this view, if (169b) involved a fronted VP rather than a vP, the example would be (173) instead, which is

indistinguishable in relevant respects from (168b). On the assumption that head movement is syntactic, this is a perfectly valid objection, of course. However, it does not imply that topicalization of remnant VPs that have their head V° moved out can be allowed. This is illustrated in (174-178). In the (b) examples a lower verb phrase is moved. This time the verb itself is not pied-piped along. In (174b) for example it might be possible to move the VP containing [up Mary's phone number]. This kind of derivation must be ruled out as the examples indicate.

- (173) John said that pictures of himself Bill would certainly like.
 (174) a. ✓John looked up Mary's phone number.
 b. *... up Mary's phone number John looked.
 (175) a. ✓John looked Mary's phone number up.
 b. *...Mary's phone number up John looked.
 (176) a. ✓John gave Mary a pencil.
 b. *... Mary a pencil John gave.
 (177) a. ✓The D.A. accused the defendants during the trials.
 b. *... the defendants during each the trials the D.A. accused.
 (178) a. ✓John told Mary that she should leave.
 b. *...Mary that she should leave John told.

Clearly then, the remnant VP must be prevented from being topicalized. Anti-locality has no problem with these examples. The question is whether they are independently ruled out. There is in fact a constraint in the literature that might account for these examples independently of anti-locality. Takano (2000) claims that remnant movement of XP is generally impossible just in case the head of XP, X° , has moved out of XP.⁶⁷ If Takano were right, his generalization would cover examples (172) and (174-178).

⁶⁷ Notice that on certain assumptions, Takano's generalization can actually be made to follow from Attract Closest. If a definition of c-command is adopted where the head of a head-chain c-commands its tail, i.e. if moved X° c-command t_{X° , then the

The problem with this line of reasoning is that Takano's generalization appears to be wrong. German remnant VP-fronting allows examples like the following (Fanselow 1991; Lenerz 1995). The examples are taken from Müller (1998:260-261, 265).

- (179) a. ✓(Ich glaube) mit dem Ball ins Gesicht hat er ihr geworfen
 I believe with the ball into.the face has he her thrown
 'I believe that he threw the ball into her face.'
 b. ✓(Ich glaube) Kindern Bonbons gibt man besser nicht.
 I believe children sweets give one better not
 'I believe that one should rather not give children sweets.'
 c. ✓Die Borussia in Führung schoss dann Effenberg.
 the Borussia in lead kicked then Effenberg
 'Effenberg then kicked the Borussia into the lead.'

On standard assumptions, there is only a single constituent preceding the finite verb in V2 clauses.⁶⁸ Müller analyzes these examples as remnant VP movement as in (180). The idea is that first the verb (index '1') moves out of the VP (index '2'). At a later point the VP moves successive cyclically to the beginning of the clause. There is no plausible alternative explanation for such facts. They clearly violate Takano's putative generalization.⁶⁹

moved X° will also c-command XP. In fact it will asymmetrically c-command XP. Hence the head of a head chain will be closer to a higher attractor than XP, which would then prevent XP from being attracted. Takano does not seem to be aware of this possible reduction of his putative generalization to Attract Closest, since he explicitly reduces all other cases of illicit remnant movement to Attract Closest except for this one, which he stipulates as an independent requirement.

⁶⁸ There are rare cases of V3 like (i), but they require a clear intonational break after the initial constituent (for discussion see <refs??>).

- (i) mir fällt grad kein Beispiel ein.

⁶⁹ Further examples where head movement of X° conceived as syntactic movement must precede remnant XP movement can be found in Abels (2001b) for Russian, Koizumi (1995) for Japanese, Dekydtspotter (1992) for French and Yoruba, and

- (180) ✓<sub>[VP [NP Kindern] [NP Bonbons] t₁]₂ denke ich
 children_{dat} sweets think I
 [CP t₂ dass man besser nicht t₂ gibt t₁]
 that one_{nom} better not gives
 ‘I think that it is better not to give sweets to children.’</sub>

The explanation for the ungrammatical English examples in (174b), (175b), (176b), (177b), and (178b) ought therefore not to be sought in Takano’s generalization. For the moment, no known constraint except for anti-locality explains such facts.

3.2.3 Conclusion

In this section I have argued that v° gives rise to anti-locality effects. Based on data from Wurmbrand (2001) I showed that VPs can move in principle. I then showed that they are stuck when they occur under v° and that the reason for this is the anti-locality constraint. Again we find confirmation for the conjecture that anti-locality effects are in the cluster of properties that follow from successive cyclicity. Of course, extraction from within the complement of v° , i.e. extraction from out of VP is possible as any example of object movement indicates (*What did John buy?*). Thus v° conforms to the predicted anti-locality pattern as shown in (181).

- (181) a. $\Diamond\checkmark[v^\circ [\dots t \dots]]$
 b. $\neg\Diamond\checkmark[v^\circ t]$

Huang (1997) for Chinese. Movement in violation of Takano’s generalization is also routinely employed in work following Kayne (1998).

3.3 Conclusion

In this chapter I have looked at the properties of two well-established phase heads, C° and v° . In chapter 2 I predicted that phase heads should be associated with anti-locality effects, in particular I predicted that the complement of a phase head should be frozen in place but that extraction out of the complement of a phase head should in principle be possible. The facts discussed here support this prediction. In the next chapter I turn to prepositions, claiming that in languages that do not allow preposition stranding, P° also obeys anti-locality.

Appendix: Why do TPs move so little?

Examples (119b) and (120b) repeated here are bona fide examples of syntactic TP movement.

- (119) a. ✓How likely to win the race is John?
 b. ✓How likely is John to win the race?
 (120) a. *How likely to be a riot is there?
 b. ✓How likely is there to be a riot?

The anti-locality constraint predicts that C° s cannot be stranded. TPs not embedded directly by C° ought to be mobile. This prediction is borne out. Notice though that anti-locality is truly relevant only if movement of TP to [Spec, CP] is ever possible, since this is the precondition for long movement that could potentially strand C° . In other words, anti-locality is relevant only if TPs can ever reach the escape hatch position. It turns out that even raising TPs, i.e. those TPs that can move in principle, only extrapose or undergo other clause internal movements. They apparently do not move to [Spec, CP] and they cannot be topicalized either (182).

- (182) a. *To be a riot there is likely.
 b. *To win the race John is likely.

Could it be that TPs never move to [Spec, CP] for reasons that are independent of the anti-locality constraint? If TPs never move to [Spec, CP] this together with successive cyclicity would independently account for the fact that TPs never strand their embedding C°.

To resolve this question we need to look at some further examples. So far what we have shown is that TPs can move sometimes. But can they ever move up to the C°-domain? The answer appears to be 'no'.

Recall the German paradigms discussed above. I repeat here the table that summarized Wurmbbrand's 2001 conclusions.

Type	Restructuring	Reduced Non- Restructuring	Non-Restructuring
Syntactic category	VP	TP	CP
Long passive	✓	*	*
Embedded tense	*	✓	✓
Embedded negation	*	✓	✓
Pronoun movement	✓	✓	*

What is important about this table is that we can uniquely pick out reduced non-restructuring infinitives: they allow pronoun movement and embedded temporal modification. Thus if we can construct examples just like (159) involving both and that also involve movement of the infinitive, we would have a clear case of TP movement. Interestingly, the crucial example (183) is ungrammatical. When the infinitive is potentially ambiguous between a CP and a TP, i.e. when the pronoun has not moved out of the infinitival, the example is fully acceptable (184).

- (183) *morgen zu reparieren hat ihn der Hans beschlossen
 tomorrow to repair has it the Hans decided
 '...because Hans decided to repair it tomorrow.'
 (184) ✓ihn morgen zu reparieren hat der Hans beschlossen
 it tomorrow to repair has the Hans decided
 '...because Hans decided to repair it tomorrow.'

What this shows is that in German infinitives that are unambiguously TPs do not move to the topic position independently of the anti-locality constraint.

Similar cases can be constructed for other languages. Bošković (1997b:23) argues that control infinitivals may be TPs. One of the arguments comes from the observation that scrambling out of finite clauses (CPs) patterns with A'-movement in certain respects; it is subject to the Weak Crossover Effect. For this and other reasons, this kind of scrambling is often referred to as A'-scrambling. Scrambling out of control clauses on the other hand patterns with A-movement in certain respects; it is not subject to the Weak Crossover Effect. This kind of scrambling is often referred to as A-scrambling. Similar patterns can be found in many languages (e.g. Bošković 1997b for Serbo Croatian, Mahajan 1990 for Hindi, Nemoto 1991 for Japanese, Sugisaki 2002 and references cited there for German and a general overview, Abels 2002a observes that movement of *ni*-phrases out of control clauses in Russian has certain properties of A-movement).

If control clauses can indeed be TPs, then the generalization in these paradigms seems to be that movement becomes sensitive to Weak Crossover if it crosses a CP or if it crosses into the CP domain. Short, clause internal scrambling in Japanese and Hindi falls under the same generalization: Scrambling becomes sensitive to Weak Crossover if it crosses a C°. This, of course, does not preclude the possibility that there might be clause internal A'-movements. The generalization put forward here about A-scrambling

is itself part of a broader generalization to the effect that no A-relation can cross a CP boundary. For some potential counterexamples to the broader claim see for example the discussion of ECM in Japanese in Hiraiwa 2002 and Bruening 2001 (even in Belfast English ECM might turn out to be problematic for the generalization that no A-relations are possible across a CP boundary (Davis 1984)). If the generalization is true either in its restricted form where it applies only to A-scrambling or in its less restricted form where it applies to all A-relations, then control infinitivals that allow A-scrambling out of them must be TPs.

Consider the Serbo Croatian examples (185) and (186) from Bošković (1997b:27). Example (185) illustrates a case of movement out of a CP, which is therefore subject to weak crossover, whereas (186) is a case of movement out of a control infinitive, a TP, which is not subject to weak crossover.⁷⁰ The infinitive itself can be moved in a very local way as example (187) shows and it can be temporally modified. This is shown in example (188).⁷¹ Example (188a) represents the unmarked order, *danas* – ‘today’ can be interpreted to be modifying the verb *kazniti* – ‘to punish’. By Wurmbrand’s criteria independent temporal modification indicates that we are dealing with at least a TP.

⁷⁰ For those speakers that allow binding of the possessive in (186), the contrast between the A'-movement case with weak crossover in (185) and the putative A-movement case without weak crossover in (186) is rather weak. The judgments are reported here as given in Bošković (1997b).

⁷¹ Examples (187)-(189) are due to Boban Arsenijević (p.c.).

Examples (188b) and (188c) where the infinitival has moved and *nekoga* is scrambled out are available on the same reading where *danas* – ‘today’ modifies the embedded verb. The examples are more or less marked and require special discourse contexts and intonation to be acceptable. This is, of course, not surprising. It is important though that the examples are acceptable at all. The illustrate again the claim that TPs are mobile in principle; as we just saw the infinitivals must be at least TPs and the possibility of A-scrambling indicates that we are dealing with at most TPs.⁷²

- (185) Svakoga_i njegov_{j/*i} otac veruje da oni mrze t_i
 everyone his father believes that they hate
 ‘Everyone, his father believes that they hate.’
- (186) Svakoga_i njegov_{j,i} otac planira PRO kazniti t_i
 everyone his father is-planning to-punish
 ‘Everyone, his father is planning to punish.’
- (187) Svakoga_i njegov_{j,i} otac kazniti planira
 everyone his father to-punish is-planning
 ‘Everyone, his father is planning to punish.’
- (188) a. Svakoga_i je njegov_{j,i} otac planirao danas kazniti
 everyone is his father planned today to-punish
 ‘Everyone, his father planned to punish today.’
 b. Svakoga_i je njegov_{j,i} otac danas kazniti planirao
 everyone is his father today to-punish planned
 ‘Everyone, his father planned to punish today.’
 c. Svakoga_i je njegov_{j,i} otac kazniti danas planirao
 everyone is his father to-punish today planned
 ‘Everyone, his father planned to punish today.’

The crucial question is now whether the infinitival can retain its status as TP and be fronted. Now, it is impossible to A-scramble *nekoga* out of the infinitive and then remnant topicalize the infinitival to the sentence initial position (189).

⁷² All judgments here are Boban Arsenijević’s.

- (189) ??/*Kazniti je svakoga; njegov; otac planirao
 to-punish is everyone his father planned
 'Someone, his father was planning to punish.'

In other words, if we try to construct the examples in such a way as to guarantee that the infinitival is a TP, because A-scrumbling applies, short movement is possible but longer movement to [Spec, CP] is not.

The same general pattern holds in Japanese (Masashi Nomura, p.c.) where relevant examples can be constructed on the basis of Nemoto's (1991) discussion. Again as in Serbo Croatian, A-scrumbling out of the infinitival and fronting it are mutually exclusive, although both options exist independently. These facts are surprising and surely indicative of *something* in particular since Serbo Croatian and Japanese are notorious for their great freedom of constituent order.

But what exactly *do* these facts indicate? It appears that movement of TP to [Spec, CP] is ungrammatical independently of the anti-locality constraint.⁷³ This is in fact

⁷³ We would then have to analyze the fronted constituent in (184) as CP.

Independent evidence for the assumption that reduced non-restructuring infinitives can be CPs in German comes from clausal pied-piping under relative clause formation. Two kinds of infinitivals are mobile in German, full restructuring infinitives (VPs) and non-restructuring infinitives (CPs). Among the movable infinitives, clausal pied-piping is possible only with non-restructuring infinitives (CPs). In other words typical diagnostics for restructuring such as the long passive do co-occur with topicalization of a restructuring infinitive but never with clausal pied-piping. This suggests that only CPs allow clausal pied-piping. Crucially, pronoun fronting as in (183) does not cooccur with clausal pied-piping either (i), but a version of (184) with clausal pied-piping is acceptable (ii). This indicates that the verb *beschliessen* takes a CP complement in (ii), just as we assumed it did in (184). The relevant examples and generalizations are all discussed in Wurmbbrand 2001.

a plausible generalization. At first sight, it appears to weaken my explanation of the ban against stranding complementizers in terms of the anti-locality constraint somewhat.

Notice however that if TP could occur as the specifier of CP, the resulting structure would be the one in (190). If it is true TPs never move to [Spec, CP], this configuration must be ruled out in general. The structure in (190) bears an obvious resemblance to the structure ruled out under the Last Resort condition in the previous chapter (191), the difference between the two being that in (191) there is full identity between the complement and the specifier and in (190) there is only categorial identity (192). We might be able to unify both cases under a more general version of the Last Resort condition which would preclude merging C° with TP₂ in the presence of TP₁ if TP₁ has more features to check against C° than TP₂. Or if we assume that all T-related features on C° get checked or otherwise valued by the C-to-T₂ in which case the C-to-T₁ relation would be left without a job to do. This kind of application of Last Resort is not implausible I believe. In fact it is indeed conspicuous that, with the exception of coordination, no clear examples of (192) come to mind.

- (190) [TP₁ [C° TP₂]]
 (191) [X [α X]]
 (192) [X₁ [α X₂]]

-
- (i) *der Roman, [den zu geben] ihr der Hans seit langem beschlossen hat
 the novel which to give to-her the Hans since long decided has
 (ii) ✓der Roman, den ihr zu geben der Hans seit langem beschlossen hat
 the novel which to-her to give the Hans since long decided has
 'the novel that Hans decided to give to her a long time ago'

I leave the issue dangling in this somewhat unresolved state. I have shown that TPs can move in principle ((119b), (120b), (187), (188)), that they probably never move to [Spec, CP], whether they were born as the complement of C° or not, and that it is tempting to try to unify this ban with the anti-locality constraint from the previous chapter via some generalized version of Last Resort. Crucially, no counterevidence to the anti-locality prediction has turned up.

To conclude this appendix, it seems to be fair to say that while no convincing counterexamples to the generalization that TPs never strand their embedding complementizers have been found, the generalization itself may be part of a larger principle which makes TPs almost immobile in general. I offered some speculation that this ban and the anti-locality constraint have the Last Resort condition as their common source.

Chapter 4: Anti-locality and the ban against adposition stranding

This chapter applies the logic of anti-locality to the ban against adposition stranding, which is operative in all but a handful of the languages of the world. I will recognize a category of adpositions in a language if there is a morphosyntactically distinct class of words with a basic locative or temporal meaning that prototypically take DPs as their complements.⁷⁴ Adposition stranding occurs iff the entire complement of an adposition undergoes movement leaving the adposition behind. In slightly outdated terminology we can say that adposition stranding occurs when a pre- or postposition P° has a trace and nothing else in its complement position, [P t]. The point of this chapter is to show that PPs fall under the Stranding Generalization. The complement cannot move but subextraction out of PP is (sometimes) possible.

To account for the possibility of P-stranding in some languages I will suggest that the phase nature of P° is subject to parameterization (see Rizzi 1980). Despite significant conceptual similarities my proposal is crucially different from van Riemsdijk's (1978a) classic theory of preposition stranding. For van Riemsdijk PPs are bounding nodes in all languages and the parameter is whether they have an escape hatch position or not. Such a parameter cannot be formulated in the present system since every phase head has a full set of unvalued features. These features allow attraction, i.e. they allow creating

specifiers as escape hatches. Remember from chapter 2 though that these features are assigned a default value if they remain unchecked, i.e. if nothing moves through [Spec, PP] the features simply get their default value. This does not lead to a crash of the derivation. I return to the issue in section 4.5.

If P°s are phase heads in the sense of this thesis, they should obey the Stranding Generalization from chapter 1. In other words the complements of prepositions, although movable categories in principle, should be frozen in place, but sub-extraction from out of the complement of P° ought to be possible. This is schematized in (193).

- (193) a. $\Diamond\checkmark[P^\circ [\dots t \dots]]$
 b. $\neg\Diamond\checkmark[P^\circ t]$

Following the path laid down in the previous chapter, the plan of this chapter should be clear. I will have to show (i) that the complement of P° is immobile under P°; (ii) that the complement of P° is mobile in principle; and (iii) that extraction out of the complement of P° is possible in principle.

Since I am, at least initially, only looking at languages that do not allow P-stranding, (i) is true by assumption. The typical complements of prepositions are DPs (*on [DP the table]*), although PPs (*from [PP under the table]*) and CP (*the question of [CP where to go]*) are also possible (Jackendoff 1973; Wunderlich 1984). DPs, PPs, and CPs are all clearly movable categories so that I can assume (ii) without argument.

⁷⁴ Craig (1977) claims that the Mayan language Jacaltec allows preposition stranding. I refrain from discussing Jacaltec though, because it is unclear whether the putative prepositions are truly morphosyntactically distinct from relational nouns.

What I will do in section 4.1 then is to collect some examples from non-P-stranding languages that seem to involve movement out of PP. In section 4.2 I turn to some languages where it is not entirely clear whether they do or do not allow P-stranding and attempt to develop diagnostics that can help decide the issue. I claim that languages that allow comparatives on the complement of P° allow P-stranding. In section 4.3 I argue that German and Dutch do not allow P-stranding. In section 4.4 I defend and discuss the generalization that only P-stranding languages allow pronominal clitics as the complement of P° (Abels 2003b, in press). This generalization is plausibly viewed as an instance of the Stranding Generalization and comes as no surprise.

Finally in section 4.5 I discuss the main findings and generalizations concerning P-stranding. I argue that the present theory is the right kind of theory to capture these generalizations although much work remains to be done to fill in the details.

4.1 The Stranding Generalization in non-P-stranding languages

In this section I argue that the ban against P-stranding is an instance of the Stranding Generalization. In other words, in languages that disallow P-stranding, P° is a phase head. As such it requires all movements out of PP to pass through [Spec, PP]. The complement of P° cannot reach this escape hatch position because of Last Resort. This entails by the reasoning developed in chapter 2 that the complement of P° should be frozen in place.

The empirical burden, which I shoulder in this section is to provide cases of the sort in (193a), where extraction of the complement of P° is impossible, but extraction from out of the complement of P° is possible. I discuss four cases: (i) *wh*-extraction out

of PP in Russian; (ii) ‘left branch extraction’ in Slavic languages; (iii) *combien*-split in certain varieties of French; and (iv) a puzzle regarding subcomparatives.

4.1.1 Subextraction in Russian

Russian is a bona fide non-P-stranding language. This is illustrated in examples (194) and (195). However, subextraction out of PP is sometimes acceptable, or close to acceptable as shown in (196). Admittedly, these examples are rare. Most sub-extraction out of PP is ungrammatical. However, the existence of examples like (196) shows that PPs are not inherently barriers to movement. Moreover, the sharp contrast between (195) and (196) shows that subextraction out of PP and P-stranding are clearly different phenomena. Example (197) is comparable to (196) without *wh*-movement.⁷⁵ Importantly, the PP headed by *na* is a part of the NP headed by *pretenzija*. This is shown in (198), which demonstrates that the PP headed by *na* is not an independent argument of the verb and by (199), which shows that the presence of the PP headed by *na* is incompatible with pronominalization of the object of the preposition *ot*. I conclude from these facts that the PP headed by *na* truly originates within the PP headed by *ot*.⁷⁶ A similar example with the same verb is given in (200).

- (194) ✓*Ot čego sleduet otkazat'sja?*
 of what follows give up-self
 ‘What should one give up?’
 (195) **Čego sleduet otkazat'sja ot?*
 *what follows give up-self of

⁷⁵ Example (197) is slightly modified from one found in the Uppsala corpus (<http://heckel.sfb.uni-tuebingen.de/cgi-bin/korpdf.pl>).

⁷⁶ The judgments reported here are Natalja Rakhlin’s.

- (196) ?*Na čto sleduet otkazat'sja ot [vsjačeskih pretenzij t_{na} čto]*
 on what follows give up-self of whatsoever hopes
 ‘What should one rid oneself of any kind of hope for?’
 (197) ✓*Sleduet otkazat'sja ot vsjačeskih pretenzij*
 follows give up-self of whatsoever hopes
 na monopoliju istoričeskogo znanija
 on monopoly historical knowledge
 ‘One has to give up all hopes on a monopoly of historical knowledge.’
 (198) **Sleduet otkazat'sja na monopoliju istoričeskogo znanija*
 follows give up-self on monopoly historical knowledge
 (199) *Sleduet otkazat'sja ot nix (*na monopoliju istoričeskogo znanija)*
 follows give up-self of them on monopoly historical knowledge
 ‘One has to give them up.’
 (200) ✓*Za kakie prestuplenija on otkazalsja ot otvetsvennosti?*
 for which crimes he rid-self of responsibility?
 ‘Which crimes did he reject responsibility for?’

The following examples are similar. Example (201) is a bona fide example of extraction out of PP. Again, the PP headed by *protiv* cannot act as an independent argument as the control in (202) shows. The crucial contrast is again that between extraction in (201) and P-stranding in (203).

- (201) ?*Protiv kakoj točki zrenija ty ešče ne slyšal ob argumentah?*
 against which point view you yet not heard about arguments
 ‘Which point of view haven’t you heard about any arguments against?’
 (202) **Protiv kakoj točki zrenija ty ešče ne slyšal (o nih)?*
 against which point view you yet not heard about them
 ‘Which point of view haven’t you heard about any arguments against?’
 (203) **Kakih argumentah protiv etoj točki zrenija ty ešče ne slyšal o?*
 which arguments against this point view you yet not heard about

4.1.2 ‘Left Branch Extraction’ in Slavic

Consider Slavic languages. Most Slavic languages are notable for a great freedom in constituency orders. Thus in a simple transitive clause in Russian, Czech, or Serbo-Croatian for example, all six logically possible orders of verb, subject, and object are

possible under the right discourse conditions and given the right intonation. Yet preposition stranding is strictly prohibited.

- (204) a. ✓Prema velikoj kući je Jovan trčao
 towards big house is Jovan run
 'Jovan ran towards the big house.'
 b. *Prema je Jovan trčao (velikoj) kući.
 towards is Jovan run big house
 c. *(velikoj) kući je Jovan trčao prema.
 big house is Jovan run towards
 d. *Prema je velikoj kući Jovan trčao
 towards is big house Jovan ran

Neither the preposition alone can be fronted as in (204b) nor can the complement of the preposition as a whole be fronted leaving the preposition behind as in (204c). No permutation of these sentences that separates the preposition from its complement is allowed. Even the otherwise extremely intrusive second position clitic *je* cannot intervene between the preposition and its complement as shown in (204d). The same is true for all prepositions in the language. They cannot be stranded and second position clitics can never separate them from their complement.

It might be thought that the ban against P-stranding might be located in the phonology, at PF. In Abels 2003b, in press I explicitly argue against attempts to locate the ban against P-stranding or the ban against prepositions hosting clitics in Serbo-Croatian at PF. There are two main arguments. First of all, P-stranding is disallowed whether it is phonetically realized or not (Merchant 1999). This can be seen in the sluicing construction (205). The preposition in (205) must be pied-piped. It cannot be

stranded under sluicing.⁷⁷ If P-stranding were disallowed because of some PF requirement of the preposition, say a prosodic requirement or the clitic like behavior of prepositions in phonology, then we would expect sluicing to ameliorate P-stranding violations. The preposition is elided in the phonology, it can therefore not cause any phonological violations.

- (205) Ana *je* govorila sa nekim, ali ne znam *(sa) kim.
 Ana is spoken with someone but not I.know with who
 'Ana talked to someone, but I don't know who.'

Second, clitics can appear without overt complements. This happens under right node raising (206) or in P°-coordination (*sa ili bez* – 'with or without'; *iznad i ispod* – 'above and below'). If we make the counterfactual supposition that there is something wrong with pronouncing a preposition without its complement, then we wrongly predict all of these examples to be ungrammatical.⁷⁸ Together examples (205) and (206) already demonstrate the utter irrelevance of PF to the ban.⁷⁹

⁷⁷ This is true quite generally in non-P-stranding languages as Merchant 1999 discovered. I give an explanation of this fact in section 4.5.2.

⁷⁸ As we saw in chapter 3 section 3.1.3, RNR does not involve ATB-movement. Examples like of RNR 'out of' PP like (206) therefore do not counterexemplify the ban against P-stranding understood as a ban against moving the complement of P°. Example (206) comes from (Abels in press) and was inspired by similar examples from another non-P-stranding language: Irish (see McCloskey 1986b).

⁷⁹ In Abels 2003b I give a third argument involving example (i). The argument more involved and the details of the analysis given in Abels 2003b are probably wrong. The argument stands irrespective of the details, though. The crucial point is that *ispod* is either a preposition (as assumed in Abels 2003b) or it is an adjective homophonous to a preposition. In any case, it can host the clitics *mi* and *je* prosodically.

- (206) ✓Jovan je ostavio čekić ispod, a Petar ispred, Marijnih kola
 Jovan is put hammer below and Petar in-front-of Mary's car
 'Jovan put the hammer below, and Petar in front of, Mary's car.' (Bošković, p.c.)

Bošković (2001b) argues extensively that, if something is not base generated in front of a second position clitic in Serbo Croatian, the only way for it to get there is to move there in the syntax. Whatever is mobile in Serbo Croatian can get in front of the clitic and whatever is immobile cannot. This is true even where clitics appear to split constituents. We can now return to the examples in (204). Example (204c) shows that P-stranding is not a syntactic option in Serbo Croatian. Example (204d) indicates that there is no way for the preposition to break away from its complement. I further illustrate the ban against P-stranding in Serbo Croatian with the *wh*-question in example (207).

- (207) a. *Kim je govorila Ana sa? (Merchant, 1999)
 who is spoken Ana with
 b. ✓Sa kim je govorila Ana?
 with who is spoken Ana
 'Who did Ana talk to?'

Serbo Croatian is a textbook case of a non-P-stranding language. However, we do find examples like (208). The complement of the preposition *velikoj kući* does not form a contiguous unit any more in (208). Something that was contained within PP, namely *kući*, is contained in PP no more. It is plausible to assume following (Ćavar and Wilder 1994; Franks and King 2000; Franks and Progovac 1994 and Zabrocki 1984 for Polish) that the derivation of such sentences proceeds in two steps. First *kući* moves out of the containing

- (i) ✓ispod mi je dostojanstva da idem.
 below my is dignity that go.1sg
 'It is below my dignity to go.'

PP, then the remnant PP *prema velikoj t_{kući}* is fronted. This is schematised in (209).

Crucially the first step in (209), movement of *kući*, involves extraction out of PP. If this analysis is correct, we have a bona fide illustration of the Stranding Generalization here and support for the thesis that P° obeys the Stranding Generalization.⁸⁰

- (208) ✓Prema velikoj je Jovan kući trčao.
 toward big is Jovan house run
 'Jovan ran towards the big house.'

- (209) [prema velikoj t_{kući}] ... kući ... t_{prema velikoj kući}
-

Consider now the examples in (210).⁸¹ The examples lend immediate support to the idea that we are dealing with remnant movement as suggested in (209). The relevant examples are (210c) and (210d). On the remnant movement analysis *sobu* moves out of PP, then PP (*pravo u veliku t_{sobu}*) moves to the front of the clause. This predicts the word order in (210c). Example (210d) cannot be derived.

- (210) a. ✓U veliku on udje sobu.
 in big he entered room t_{PP}
 'He entered the big room.'
 b. ✓On udje pravo u veliku sobu.
 he entered straight in big room
 c. ✓Pravo u veliku on udje sobu.
 straight in big he entered room
 'He went straight into the big room.'

⁸⁰ Alternatively we can assume that the preposition + adjective is formed in some way by head movement (Borsley and Jaworska 1988; Bošković in press). After whatever is involved in forming this combination the preposition + adjective complex would move out. Even under this analysis we have to assume extraction out of PP. It will become clear immediately why I do not pursue this option.

⁸¹ In Abels in press (draft 04/01/2002) example (210c) is erroneously reported to be ungrammatical.

- d. *U veliku on udje pravo sobu.
in big he went straightroom

The alternative, direct extraction approach mentioned in footnote 80 gets the constituency wrong. On that approach preposition + adjective form a constituent. Thus if anything (210d) should be good. Whether (210d) is actually predicted to be acceptable or not depends to a large extent on detailed assumptions about how the preposition + adjective combination is derived. The grammatical (210c) is clearly ruled out on the direct extraction approach.⁸² On that approach, every constituent that includes the modifier *pravo* also includes the noun *kući* (except for the modifier itself of course). The modifier *pravo* should therefore be unable to move together with the preposition-adjective unit leaving the noun behind. The problem doesn't arise under the remnant movement approach sketched in (209) which is therefore supported by these data.

The Serbo Croatian situation concerning Left Branch Extraction with PPs is typical of the situation found more broadly in the Slavic languages (except for Bulgarian, which disallows Left Branch Extraction). See Borsley and Jaworska 1988; Siewierska

⁸² Given that in Serbo Croatian second position clitics follow the first constituent within a clause, example (i) with the clitic *je* is a telling probe into the constituency of the examples. Example (i) clearly shows that *pravo u veliku* may form a constituent. An alternative derivation where *pravo* is moved independently of *u veliku* is also possible as example (ii) indicates. The judgments reported are Boban Arsenijević's. Olga Tomić finds (i) acceptable and (ii) unacceptable.

- (i) ✓Pravo u veliku je ušao sobu.
directly in big is enteredroom
"He has entered directly into the big room."
(ii) ✓Pravo je u veliku ušao sobu.
directly je in big enteredroom
"He has entered directly into the big room."

1984; Corver 1990 for Polish Mehlhorn 2001; Sekerina 1997 for Russian; Siewierska and Uhlířová 1998 for an overview of the Slavic situation).

Bošković (in press) discusses the remnant movement analysis rejects it. I review the reasons for rejecting it below and show why I do not think that they are very strong. Let me point out at the outset that I agree fully with Bošković that the properties of regular adjectival Left Branch Extraction in Slavic are so similar to the properties of Left Branch Extraction with PPs that we have seen above that both should be given a uniform analysis. In other words, if I argue below for a remnant movement analysis of Left Branch Extraction with PPs, I would extend the analysis to other cases of left branch extraction.

The weakest of Bošković's arguments against the remnant movement analysis comes from the observation that in neutral contexts when a noun is modified by two adjectives that can be relatively easily reordered with respect to each other, Left Branch Extraction is disallowed (211) (Bošković's ex. 34b).

- (211) *Lijepe je on vidio visoke djevojke.
beautiful is he seen tall girls

But with a bit of contextualization such examples actually become acceptable as (212) (Bošković's ex. 72) shows. Similarly *wh*-words, which are morphologically adjectives in Serbo Croatian can be Left Branch Extracted out of an NP that contains another adjective (213).

- (212) A: I think that Marko said he saw ugly tall girls.
B: Ma, ne, lijepe je on vidio visoke djevojke.
no beautiful is he seen tall girls
'He saw BEAUTIFUL tall girls, not ugly tall girls.'

- (213) ✓Koje je Petar novo auto upropastio ?
 which is Peter new car bought
 'Which new car did Peter buy?'

In general, when two adjectives cannot be freely reordered within the NP such as *neozbiljnog* – 'not-serious' and *mašinskog* – 'mechanical' in (214) LBE becomes of the higher one, but not the lower one, becomes possible apparently even without a great deal of contextualization (215).

- (214) a. ✓neozbiljnog mašinskog tehničara
 not-serious mechanical technician
 'a non-serious mechanical technician'
 b. *mašinskog neozbiljnog tehničara
 mechanical not-serious technician
- (215) a. ?Neozbiljnog je on otpustio mašinskog tehničara
 non-serious is he fired mechanical technician
 'He fired a non-serious mechanical technician.'
 b. *Mašinskog je on otpustio neozbiljnog tehničara
 *mechanical is he fired non-serious technician

Bošković assumes without argument that the degraded status of (211) out of context should be accounted for syntactically. The fact that a bit of contextualization makes the examples acceptable, as we saw, casts doubt on this line of reasoning. A look at the context in (212) suggests that the constituent that is left behind describes an established class of entities which is further subclassified by the fronted adjective. When these sentences are given out of the blue, it is apparently difficult to construct a suitable situation that would make 'tall girls' an established set that is then further subdivided into beautiful, ugly, etc. tall girls. In other words there does not seem to be a reason to account for the status of (211) out of context in the syntax.

Notice that the fact that (215b) is ruled out follows immediately from the remnant movement analysis given that the order in (214b) is ruled out.

On the other hand, examples where two adjectives are subjected to Left Branch Extraction together are ruled out (Bošković in press ex. 38). The example improves if both of the adjectives receive independent focus.

- (216) *Visoke lijepe on gleda djevojke
 tall beautiful he watches girls

Bošković shows based on data like (217) that the particular version of the remnant movement analysis suggested in Franks and Progovac 1994 faces serious problems. Franks and Progovac had suggested that *auto* in the examples would move out of its containing DP and right adjoin to IP. Franks and Progovac would therefore expect (217a) and (217c) to be ill-formed and (217b) to be well formed. As noted by Bošković the facts suggest that movement of *auto* is to a much lower position and to the left. The relatively free ordering of the verb with respect to *auto* might derive from the fact that either *auto* has several possible landing sites or that the verb can move to different positions.

- (217) a. ✓Crveno je on kupio auto prije tri dana
 red is he bought car before three days
 b. ?*Crveno je on kupio prije tri dana auto
 red is he bought before three days car
 c. ✓Crveno je on auto kupio prije tri dana
 red is he car bought before three days
 'He bought a red car three days ago.'

Once we accept the conclusion that the initial step of moving *auto* out of its containing DP is to the left, the examples in (219), also mentioned by Bošković as a problem for the remnant movement analysis fall into place (Bošković in press ex. 20-23).

Adjectives appear to behave just like verbs in relevant respects. Again as in the case of verbs examples (218a) and (218b) can be derived by assuming that the adjective can be realized in a higher or in a lower position or that the noun phrase *novim poslom*

moves (218b) or fails to move (218a). In any case, to derive (218c) we can either assume that that there is enough space within the AP for *poslom* to move to, enabling remnant movement of [*novim t_{poslom}*] and that *poslom* in (218d) moves to a higher position, outside of AP prior to remnant movement of [*novim t_{poslom}*].⁸³

- (218) a. On je zadovaljan novim poslom
 he is content new job
 b. On je novim poslom zadovaljan
 he is new job content
 c. ✓Novim je on zadovaljan poslom
 new is he content job
 ‘He is content with his new job.’
 d. novim je on poslom zadovaljan
 new is he job content

The situation is markedly different for noun phrases within noun phrases.

Consider the examples in (219).

- (219) a. ✓On je vidio prijatelja njegov majke
 he is seen friend his mother
 ‘He has seen a friend of his mother.’

⁸³ Bošković construes these examples as a problem for certain versions of the remnant movement analysis. The problem is real since LBE of the lower adjective is possible in (218c) but impossible in (215b). It arises if the structure of a predicative adjective with a complement ‘content new job’ ([_{AP} [_{AP} [_{NP}]]]) is essentially the same as the structure of a noun phrase with two attributive adjectives ‘pretty tall girl’ ([_{AP} [_{AP} [_{NP}]]]). Their distinct syntactic behavior shows that predicative adjectives and attributive adjectives do not have identical structures. Whatever the exact details of the structures, it seems independently plausible to assume different structures for attributive and predicative APs. Thus even if we assume, for the sake of argument, that the constituency is basically as indicated above, there is a difference in behavior with respect to Case. In the attributive structure all the adjectives agree in Case with the head noun in the predicative structure the adjective assigns Case. A way to make the distinction between the two structures is in terms of a Case projection KP such that the attributive structure would be [_{KP} [_{AP} [_{AP} [_{NP}]]]] and the predicative structure [_{AP} [_{KP} [_{AP} [_{NP}]]]]. We could then postulate the remnant movement in these cases always targets KP. This would rule out (215b) and various variations on it.

- b. *Čije je on vidio prijatelja t_{čije} majke
 whose is he seen friend mother
 ‘Whose mother did he see a friend of?’
 c. (?)?on je njegov majke vidio prijatelja
 he is someones mother seen friend
 d. (?)?Čije je on majke vidio prijatelja
 whose is he mother seen friend
 e. *Čije je on prijatelja majke vidio
 whose is he friend mother seen

Example (219a) does not involve LBE. It simply indicates the base position of the possessive pronoun at the left edge of the noun phrase headed by *majke*. Example (219b) indicates that direct LBE from that position is impossible. If it is true that *majke* would first have to move to the left to enable remnant movement of *čije t_{majke}* this is of course not unexpected. Example (219c) shows that leftward movement out of the noun phrase headed by *prijatelja* is marginally possible. In that case LBE of *čije* becomes possible (219d). There are two possible derivations for example (219d). One possibility is that *majke* moves out of the noun phrase headed by *prijatelja* followed by remnant movement of [*čije t_{majke}*] to the clause initial position. An alternative would be to move the entire phrase [*čije majke*] out of the containing noun phrase first, followed by extraction of *majke* out of the moved phrase [*čije majke*] followed by remnant movement of [*čije t_{majke}*].⁸⁴ The fact that (219e) is ungrammatical follows again, since in the example the entire noun phrase [*prijatelja čije majke*] is first fronted. The example can therefore be ruled out by the remnant movement analysis on a par with (219b).

⁸⁴ If there are arguments for the second derivation sketched here, the fact that on this derivation movement takes place out of a moved constituent should not be counted as an argument against the remnant movement analysis, since the alternative direct

The most serious problem facing the remnant movement analysis comes from the following type of facts.

- (220) a. ✓Ljepe je on kuće vidio
 beautiful is he house seen
 'He saw a beautiful house.'
 b. *On je kuće vidio ljepe.
 he is house seen beautiful
 c. *Kuće je on vidio ljepe.
 house is he seen beautiful

If we postulate that (220a) is derived by first moving *kuće* to a preverbal position followed by remnant movement of *ljepe* *t_{kuće}*, the question arises why the first movement does not occur independently as in (220b) and why (220c) is impossible where *kuće* is fronted by itself.⁸⁵ The only approach I'm aware of that can handle these kinds of facts is an idea by Fox and Pesetsky (Fox and Pesetsky 2003) according to whom there are certain domains within which the linear precedence relations are fixed at a certain point during the derivation and cannot be altered later on in the course of the derivation. The relations that are fixed are precedence relations, not immediate precedence relations. This means that items from within that domain will be able to move with respect to other

extraction analysis needs to postulate movement out of a moved constituent for these cases, too.

⁸⁵ Even in Russian, where both of the orders shown in (220) are acceptable, the addition of a preposition shows that the generalization is true even there (i-iv). The preposition must always occur in the fronted constituent (Mehlhorn 2001:65; Sekerina 1997). The same is true for other Slavic languages.

- (i) ✓S bratom ona vstretilas' lenivym
 with brother she met-self lazy
 'She met with (her) lazy brother.'
 (ii) ✓S lenivym ona vstretilas' bratom.
 (iii) *Bratom ona vstretilas' s lenivym.

elements introduced into the derivation later on, but their relative linear orders must be preserved. Movements that do not preserve the relative linear order must proceed through the edge of the spell out domain, i.e. displacement must already occur domain internally. If we assume that the NP in SC is such a spell out domain, then the order preservation facts can be made to follow. Fox and Pesetsky use their system to account for Holmberg's generalization (Holmberg 1986, 1999) among a number of other phenomena.

The solution cannot be immediately imported into the present system, since it requires cyclic spellout, which we rejected above in chapter 2. I leave the issue in this unresolved state for the moment.

In any case, even if the remnant movement analysis is ultimately rejected, LBE in Slavic requires that PPs be transparent for movement operations in the sense that parts of the PP may be extracted – just not the entire NP.⁸⁶ The conclusion seems unavoidable

-
- (iv) *Lenivym ona vstretilas' s bratom.

⁸⁶ Bošković reports an observation by Sandra Stjepanović that LBE is often possible in contexts that are otherwise total islands to extraction. This is shown in (i), where extraction out of the phrase headed by 'because' is ungrammatical, presumably a violation of the adjunct condition. Example (ii) is a control, indicating that extraction of the dative argument of the noun is in principle possible. Finally, example (iii) illustrates the possibility to split the phrase headed by because through LBE despite the fact that this should give rise to an adjunct condition violation. At present I do not know how to handle these facts. They appear problematic under any movement analysis of LBE.

- (i) *Čime_i je on pobegao [zbog [pretnje t_i]].
 what.dat is he run-away because-of threat
 (ii) ✓Čime_i ga je [pretnja t_i] uplašila?
 what.dat him is threat scared
 'The threat of what scared him.'
 (iii) ✓Zbog čijih je došao studenta.
 because-of whose is arrived students
 'Because of whose students did he arrive?'

then that movement out of PP is allowed in some form or another, but that the complement of P° as a whole cannot strand the preposition.

4.1.3 Combien-split in French

We now turn to *combien*-split in French. We find the same kind of pattern that we have seen in the previous examples. As is well known, *combien* can under certain conditions break away from its containing noun phrase.

- (221) a. ✓Combien de livres as-tu lu ?
 how-many of books have you read
 'How many books have you read?'
 b. ✓Combien as-tu lu de livres?
 how-many have you read of books
 'How many books have you read?'

There are two plausible kinds of movement analysis for this pattern. Either *combien* is extracted directly (222a) or there is, again, a remnant movement derivation as in (222b).

- (222) a. combien [t_{combien} de livres]
 b. [combien t_{de livres}] ... [de livres] ... t_{combien de livres}

French does not allow P-stranding as (223) through (225) indicate (for discussion see Zribi-Hertz 1984; Kayne 1984; King and Roberge 1990; Roberge 1998; Roberge and Rosen 1999 and references cited there).⁸⁷

- (223) a. ✓À quels photographes est-ce que tu as parlé ?
 to which photographers is-it that you have talked
 'Which photographers have you talked to?'

⁸⁷ On the phenomenon of 'orphan prepositions' in French see Roberge 1998; Roberge and Rosen 1999; Zribi-Hertz 1984. I will briefly touch upon orphan prepositions below. The functional prepositions *à* and *de* used in the examples here, never occur as orphan prepositions anyway.

- b. *Quels photographes est-ce que tu as parlé à ?
 which photographers is-it that you have spoken to
 (224) a. ✓De combien de photos (est-ce que) tu as besoin ?
 of how-many of photos is-it that you have need
 'how many pictures do you need?'
 b. *Combien de photos (est-ce que) tu as besoin de ?
 how-many of photos is-it that you have need of
 (225) a. ✓Pour quel candidat as-tu voté ?
 for which candidate have-you voted
 'Which candidate did you vote for?'
 b. *Quel candidat as-tu voté pour ?
 which candidate have-you voted for

Kayne (1983 chapter 3 p. 51-52); Starke (2001:45) note that *combien*-split is possible with PPs. Thus compare examples (223) and (224) with (226) and (227).⁸⁸

- (226) a. ✓Tu as besoin de combien de photos ?
 you have need of how-many of photos
 'How many photos do you need?'
 b. ✓De combien (est-ce que) tu as besoin de photos ?
 of how-many is-it that you have need of photos
 'How many photos do you need?'
 (227) a. ✓Tu as parlé à combien de photographes ?
 you have talked to how-many of photographers
 'How many photographers have you talked to?'
 b. ✓À combien (est-ce que) tu as parlé de photographes ?
 to how-many is-it that you have talked of photographers
 'How many photographers have you talked to?'

Clearly the partitive *de photo(graphe)s* must somehow have escaped from the containing PP headed by *à* and *de* respectively. Given our discussion of Slavic above, the most plausible assumption is to follow Starke 2001 and adopt the remnant movement

⁸⁸ These sentences are not equally acceptable in all varieties of French. The judgments given here are Starke's (p. 45), they are confirmed by Éric Mathieu (p.c.).

analysis of *combien*-split (222b).⁸⁹ Crucially, the examples show that material can escape a PP – just not the complement of P° itself. Kayne 1983 remarks (citing Obenauer 1976:17) that not all prepositions are equally acceptable with the construction. Michal Starke (p.c.) informs me that the cline is along the functional-lexical axis.

4.1.4 Subcomparatives into PP

A final plausible case of extraction out of PP is furnished by sub-comparatives. As has been noted in the literature (Corver 1990; Kennedy 2002) subcomparatives are possible into PPs even in languages that do not allow P-stranding.

- (228) *Bydlel jsem ve více městech než ty jsi bydlel v.
live.past.1sg AUX in more cities.gen.pl than you AUX live.past.2sg in
'I have lived in more cities than you have lived in.'
- (229) ✓Chci bydlet ve více amerických městech
want.1sg.pres live.INF in more American city.gen.pl
než jsem bydlel v evropských městech.
than AUX lived.past.1sg in European cities
'I want to have lived in more American cities than I have lived in European cities.'

The Czech example (228) from (Kennedy 2002:560) illustrates that comparatives cannot be formed on the complement of a preposition in Czech. Presumably this is due to the fact that comparatives require movement of the (abstract) comparative operator, but since Czech does not allow P-stranding this movement is ungrammatical. Example (228) then merely shows that comparative movement patterns with other kinds of movement in

⁸⁹ We can assume that the first step of movement lands in a very low position, say AgrO and that the verb in French is obligatorily above this position. The distinction between French and Serbo Croatian (recall (217)) is that verb movement above the landing site of the initial movement is obligatory in French but optional in Serbo Croatian.

obeying the ban against P-stranding. Example (229) is somewhat surprising. It shows that subcomparison into a PP is possible. We can understand the pattern if full comparative movement is movement of the complement of PP and if subcomparatives are formed by moving the degree operator only. If this is true, comparative movement is of a categorically disallowed movement of the type of (193b), whereas movement in subcomparatives would be extraction out of the complement of P° (193a). The latter is but the former is not in principle allowed by the present theory. It is fairly common for a language that has clausal comparatives and subcomparatives to disallow comparatives involving the complement of PP but to allow subcomparatives into PP. This pattern is also found in Russian, German, Dutch (Corver 1990), Hebrew (Yael Sharvit, p.c.) to name but a few.⁹⁰

4.1.5 Conclusion

I believe it is safe to conclude at this point that even in non-P-stranding languages, extraction from PP is not categorically excluded. What is excluded is movement of the complement of the preposition. The question remains of course what elements can move out of PP and under what conditions. I will not concern myself with that question here, though. The purposes of this section has, I believe, been achieved. I have demonstrated that PPs obey the Stranding Generalization (193).

⁹⁰ French does not seem to allow the pattern (see Kayne 1984:53).

4.2 On the diagnosis of P-stranding

The business of detecting which languages allow P-stranding under the definition I am assuming is somewhat subtle. Remember: A language allows P-stranding iff the entire complement of the preposition can be moved out of its position as the complement of P. The subtlety arises for the following reasons. When we find an element that looks like an adposition with an empty complement and the empty complement position (the gap) appears to be linked to an antecedent somewhere else in the clause (the filler), we cannot be sure that we have detected a case of P-stranding. At least three different possible interpretations of the situation are possible: (i) we are dealing with a genuine case of P-stranding ([P t]); or (ii) the complement position of the adposition is occupied by a silent resumptive pronoun *pro* ([P *pro*]); or (iii) the putative adposition really belongs to a different morphosyntactic class in such environments (typically the contemplated alternative is ‘adverb’ – see for example Campos 1991. The examples discussed in Cinque 1990:50 should probably receive the same analysis.).⁹¹ The situation is complicated even further because Muysken (1977) has argued that in Papiamentu when a preposition is followed by an audible complement, the preposition might still be stranded,

⁹¹ There is a fourth possibility, namely that movement out of PP has taken place but that it was a masked case of subextraction (pattern (193a) rather than (193b)). Subextraction could be masked if the material that separates the preposition from the moved element is abstract. Lacking a coherent, sufficiently restrictive theory of abstract categories (see my comments on Cinque’s universal hierarchy of functional heads in chapter 2, the comments on restrictiveness in the Kaynean program in chapter 3, and the discussion following example (291) below), I will not attempt to discuss this fourth option seriously.

the audible morpheme being the spell out of a trace.⁹² Muysken’s theory, which has been adopted also for São Tomense by Hagemeijer (2000) and a number of other Portuguese based Creoles (Alexandre and Hagemeijer 2002), must be distinguished from the idea that resumptive pronouns are realizations of traces. If Muysken’s theory is correct, we must be able to distinguish overt and null resumptive pronouns on the one hand from null and overt traces on the other hand.

4.2.1 On How to Distinguish Traces from (null) Resumptive Pronouns

In the introduction to this section I laid out why it is not always easy to be sure what kind of element occupies the complement position of a preposition. What criteria are there to distinguish traces from null-resumptive pronouns? What criteria can be used to distinguish overt resumptive pronouns from spelled out traces if the latter exist?

If every *wh*-construction involves either a trace or a resumptive pronoun and if this exhausts the space of possibilities, then two tests that can diagnose the absence of a trace and the absence of a resumptive pronoun would be able to pin down the category of the element in question – at least if the tests are applicable.

The first possible diagnostic that I will discuss is a test for the absence of a trace. The standard assumption is that in English the empty category forming the complement of a stranded preposition is a trace in examples like (230). The gaps in (230) can be assumed to be filled by a trace, because they pass the diagnostics for *wh*-movement

⁹² Muysken 1980 shows that this spelled out trace does not ameliorate standard island conditions. Island amelioration requires the presence of a resumptive (or intrusive) pronoun with very distinct properties from those of the putative spelled out trace.

pioneered in Chomsky 1977. Importantly, the filler gap relation is potentially unbounded as (231a) and (231b) show but subject to standard locality constraints such as the *wh*-island constraint (231c) and the complex NP constraint (CNPC) (231d).

- (230) a. ✓Which candidate have you voted for __?
 b. ✓Who did you take pictures of __?
 c. ✓What bed did you sleep in __?
 d. ✓Who did you take advantage of __?
- (231) a. ✓Which candidate do you think Peter has voted for __?
 b. ✓Who do you think Mary would like Frank to say that he took pictures of __?
 c. *Which candidate do you wonder who voted for __?
 d. *Which candidate do you believe the claim that W. voted for __?

Resumptive Pronouns are often assumed to repair island violations. A typical example where this is indeed the case is the Irish example (232) from (McCloskey 1990:209). In (232a) there is a resumptive pronoun within the complex NP and the example is acceptable, but in (232b) there is a trace and the example is unacceptable.

- (232) a. ✓na dánta sin nach bhfuil fhios againn cén áit
 the poems DEM COMP_{pro}+ NEG is knowledge at-us what place
 ar cumadh iad
 COMP were-composed them
 ‘those poems that we do not know where they were composed’
- b. *na dánta sin nach bhfuil fhios againn cén áit
 the poems DEM COMP_{pro}+NEG is knowledge at-us what place
 ar cumadh
 COMP were-composed

As a diagnostic for the presence of a resumptive pronoun, locality obviation effects are less than perfect, since there are many cases where resumptive pronouns do not ameliorate island violations (for extensive discussion see Boeckx 2001). As far as I can see there is only an implicational relationship of the following type. If a particular construction in the *wh*-movement family in a given language is allowed to flout an island constraint that is otherwise operative in the language, then this construction involves a

resumptive pronoun. The absence of locality obviation therefore does not diagnose the presence of a trace, but the presence of locality obviation does diagnose the absence of one. This gives us the first half of the minimally required diagnostic tools.

The next question is what diagnostics might be used to test for the absence of a resumptive pronoun.⁹³

I suggest the following generalization as a diagnostic for the absence of a resumptive pronoun. Resumptive pronouns appear to be incompatible with comparatives of inequality. Traces are compatible with all standard *wh*-constructions (Chomsky 1977), but resumptive pronouns are systematically impossible in comparative clauses.⁹⁴ The following set of data from Hebrew data (Yael Sharvit p.c.) exemplify this. Example (233) shows that a D-linked *wh*-phrase in an interrogative can be resumed by a pronoun.

Example (234) shows the same for a relative clause. Except for the comparative as in

⁹³ Boeckx 2001:58, following an observation about Hebrew made by Edith Doron (Doron 1982, which I haven’t seen) claims that a cross-linguistically extremely robust generalization is that resumptive pronouns do not appear with non-D-linked *wh*-phrases (Pesetsky 1987) in interrogatives. If true this generalization could be used as the desired diagnostic for the absence of a resumptive pronoun. Obviously, the test would only apply if the language in question allows resumptive pronouns in interrogatives in the first place. Unfortunately, Boeckx does not give clear criteria for applying the test in non-trivial cases like the Irish examples (1) and (2) from McCloskey (1990:228, 234).

- (1) ✓Goidé a ndearna tu leis è?
 what COMP_{pro} did you with-it it
 ‘What did you do it with?’
- (2) ✓Cé a raibh tú ag caint leis ?
 Who COMP_{pro} were you talk-PROG with-him
 ‘Who were you talking to?’

(235) and (236), all of the standard *wh*-movement constructions allow the occurrence of resumptive pronouns in one language or another and sometimes all are possible in the same language. McCloskey 1979, 1990 runs through the whole list for Irish; Tuller 1986 reports restrictions for Hausa; etc.

- (233) ✓eyze sfarim dani kara (otam)
 which books Dani read them
 'Which books did Dani read (them)?'
- (234) ✓ha-sfarim she Dani kara (otam)
 the-books that Dani read them
 'the books that Dani has read'
- (235) Dani kara yoter sfarim me-asher Yosi kara ✓(*otam)
 Dani read more books than-that Yosi read them
 'Dani has read more books than Yosi has read.'
- (236) Dani diber al yoter sfarim me-asher Yosi kara ✓(*otam)
 Dani talked about more books than-that Yosi read them
 'Dani has talked about more books than Yosi has read'

There are contexts where resumptive pronouns are forced to occur. In particular, prepositions require the presence of a resumptive pronoun in their complement (237) and (238).⁹⁵ Since prepositions are incompatible with the absence of a resumptive pronoun and since resumptive pronouns are incompatible with the comparative, there is no grammatical variant for the clausal comparatives in (239) and (240). The intended meaning of (239) must be expressed as in (241).

⁹⁴ In many languages the distribution of resumptive pronouns is in fact much narrower than the distribution of traces. True resumptive pronouns often occur only in relative clauses (see Boeckx 2001 for much pertinent discussion).

⁹⁵ I follow Shlonsky 1997 and assume that there is a null pronoun in the complement of the preposition, which, in that case, is inflected. The facts discussed in this section are also compatible with the analysis where the inflection on the preposition is treated as a clitic. See section 4.4 for the reason to assume Shlonsky's analysis.

- (237) eyze sfarim dani diber {✓al-eyhem |*al}
 which books Dani talked on-them on
 'Which books did Dani talk about?'
- (238) ha-sfarim she Dani diber {✓al-eyhem |*al}
 the-books that Dani talked on-them on
 'the books that Dani has talked about?'
- (239) Dani diber al yoter sfarim me-asher Yosi diber {*aleyhem |*al}.
 Dani talked about more books than-that Yosi talked on-them on
 Intended: 'Dani has talked about more books than Yosi has talked about.'
- (240) Dani kara yoter sfarim me-asher Yosi diber {*aleyhem |*al}
 Dani read more books than-that Yosi talked on-them on
 Intended: 'Dani has read more books than Yosi has talked about.'
- (241) ✓Dan diber al yoter sfarim me-asher Yosi
 Dan talked about more books than-that Yosi
 'Dani talked about more books than Yosi.'

In English this restriction is not operative as the translation indicates. The same restriction against resumptive pronouns in comparatives seems to hold more generally (Ur Shlonsky (p.c.) for Arabic and Arsalan Kahnemuyipour (p.c.) for Persian).⁹⁶

Irish is a bit more complicated. Although McCloskey (1979:135, 156) had claimed that resumptive pronouns are incompatible with the comparative McCloskey (1990) claims that resumptive pronouns are possible in comparative clauses. According to McCloskey (p.c.) the relevant examples are very rare. They are of two kinds. The first kind is exemplified in (242).

⁹⁶ Zribi-Hertz 1984 has argued that when prepositions in French occur without a complement, so called orphan prepositions, the complement of the preposition is occupied by a null pronominal *pro*. *pro* can act as a resumptive pronoun. It is tempting to try to show that the generalization made here holds for French. A number of factors interfere with a straightforward construction of the relevant examples. (I would like to thank Mélanie Jouitteau and Anne Zribi-Hertz for discussion of this issue.) To the extent that the relevant examples can be constructed in certain dialects of French, the expectation appears to be borne out.

- (242) ✓níos mó mine ná a raibh gnaithe acu fiin leis
 more meal than COMP_{pro} was need at-themselves with-it
 'more corn-meal than they needed themselves'

This type of example does not seem to be a serious problem for the generalization that resumptive pronouns are incompatible with comparatives. The reason for this is that the putative comparative clause might be analyzed as a headless relative clause (for the suggestion that constructions with headless relative clauses can simulate the effect of a comparative see Beck, Oda and Sugisaki 2002; Besten 1978, 1981. In an email (12/22/02), Jim McCloskey comments on this suggestion regarding (242) as follows: "Here, an analysis in terms of the headless relative is rather plausible, since the string *a raibh gnaithe acu fiin leis* is independently well-formed as a headless relative with the meaning *all that they themselves needed*. So [242], I think, is plausibly paraphrased as *more corn-meal than all that they needed*." I will then assume that (242) is not a counterexample to the generalization that resumptives do not occur in comparatives.

The second type of comparative with a resumptive pronoun is shown in examples like (243) (from McCloskey 1990:239 and (244)-(246) (Jim McCloskey, p.c.). These examples are certainly more problematic for my generalization. Notice that all of the examples are equatives. The examples that I used to test the generalization in other languages all involve comparison of inequality. I suspect, that equatives have a somewhat different syntax than comparatives of inequality.⁹⁷

⁹⁷ Price (1990) mentions that the range of constructions for comparatives of equality is greater than the range of constructions for comparatives of inequality in Spanish.

- (243) ✓Do fuair sé leaba chó math agusar lui sé riamh uirthi —
 get-PAST he bed as good as COMP_{pro} lie-PAST he ever on_{3sg.fem}
 'He got a bed as good as he ever lay on (it).'
 (244) ✓fairrge comh fiadhain agus comh garbh agus a gcualaidh mi a tuairisc
 sea as wild and as rough as C heard I its report
 'a sea as wild and as rough as I have ever heard tell of'
 (245) ✓phiosa breacoirce chomh maith agus ar leag fear nó bean súilariamh air
 piece fine corn as good as C laid man or womaneye ever on-it
 'as fine a patch of corn as man or woman ever laid eyes on'
 (246) ✓capall comh scsipeamhail a's shuidh tú ar a mhuin ariamh
 horse as lively as sat you on its back ever
 'a horse as lively as you ever rode'

The generalization seems to be quite strong otherwise. We can now ask why it should hold. I would like to suggest, albeit tentatively, that argument traces in comparative clauses are always interpreted functionally. Consider example (247). The sentence means, informally, something like (248).

- (247) John has taken out more books from the library than he will read *t*.
 (248) The maximal number *n* such John has taken *n*-many books from the library is greater than the maximal number *m* such that John will read *m*-many books.

The object position is filled by an expression that neither refers itself, nor is it itself bound. Rather the argument position of the quantifier *many* is bound. The object is interpreted as a function from numbers to quantifiers or cardinal expression. It has been observed that resumptive pronouns allow functional readings only if the function denoted by the resumptive pronoun is a 'natural function' (see Sharvit 1999). To account for the fact that resumptive pronouns do not occur in comparatives, I assume, as seems natural, that the function from numbers to cardinal expressions is not a particularly natural function. In fact, Sharvit claims that resumptive pronouns in general can have only such denotations as the pronoun can have in its non-resumptive, non-A'-bound use. Surely the

function above is not a possible denotation for an independent pronoun. This would explain why resumptive pronouns are incompatible with comparatives.⁹⁸

I conclude the discussion here and will assume the generalization that resumptive pronouns are incompatible with comparatives of inequality. We are now in principle in a good position to examine the claim that in Papiamentu and some other languages traces of *wh*-movement in the complement position of a preposition can be morphologically realized. If these items in the complement position of a preposition ameliorate island violations that are otherwise operative in the language, then we have an argument that these items are not traces but resumptive pronouns. On the other hand, if these items are compatible with comparatives, we have an argument that they are not resumptive pronouns.

4.2.2 *Frisian*

We can immediately apply this reasoning to Frisian. Hoekstra (1995:107-109) discusses a contrast between Dutch and Frisian. Consider Dutch first. It had been observed (see van Riemsdijk (1978b); den Besten 1978; 1981) that comparative movement is inapplicable to the complement of a preposition in Dutch.

- (249) *Jan heeft meer geld verdiend dan zijn vrouw op gerekend had.
 Jan has more money earned than his wife on counted had
 intended: 'Jan made more money than his wife had expected.'

⁹⁸ This leaves open the door that in the Irish equatives discussed we might be able to find a natural function. They have the flavor of generic quantification, which makes the search for a general property, a 'natural function' more plausible.

Hoekstra accounts for the ungrammaticality of (249), following den Besten (1978), by assuming that the formation of such examples would involve the movement of an amount phrase, a full DP, out of PP. However, extraction of full DPs from the complement position of P° is uncontroversially disallowed in Dutch. The account thus unifies the status of (249) with that of (250), where the overt extraction of an amount phrase from the PP is disallowed.

- (250) *Hoevel geld had ze op gerekend?
 how-much money has she on counted

Interestingly, the corresponding Frisian sentences are grammatical according to Hoekstra. Thus in Frisian we get (251) and (252).

- (251) ✓Jan hat mear jild fertsjinne as dat syn frou op rekkene hie.
 Jan has more money earned than that his wife on counted had
 'Jan made more money than his wife had expected.'
 (252) ✓Hoefolle jild hie se op rekkene?
 how-much money has she on counted
 'How much money did she count on?'

Hoekstra suggests that neither Frisian nor Dutch allow preposition stranding in the sense that neither language allows movement of a full DP complement of P° to a position outside of PP. The contrast between Frisian and Dutch just seen arises, Hoekstra claims, because Frisian in contrast to Dutch allows a null resumptive pronoun as the complement of P°. If the generalization concerning resumptive pronouns suggested in this subsection is correct, then Hoekstra's conclusion about Frisian cannot be right. The most plausible analysis for Frisian is then the one rejected by Hoekstra, namely, that Frisian unlike Dutch allows preposition stranding. Hoekstra rejects this analysis despite the existence in Frisian of examples like (252) and (253) (Hoekstra 1995:97, 105), which

look like bona fide cases of P-stranding. Again the Dutch direct translations of the examples in (254) are ungrammatical.

- (253) a. ✓De bern wurdt net nei harke.
 the children is not to listened
 'The children, people didn't listen to.'
 b. ✓Wa hast mei praat?
 Who have-you with talked
 'Who did you talk to?'
- (254) a. *De kinderen word't niet naar geluisterd.
 the children was not after listened
 b. *Wie heb je mee gepraat?
 who have you with talked

There are two main reasons why Hoekstra does not assume a stranding analysis for Frisian. First, circumpositions do not allow stranding as shown in (255).

- (255) a. ✓Hja ha nei de útstalling ta west.
 they have to the exhibition to been
 'They have been to the exhibition.'
 b. *Hokker útstalling ha se nei ta west?
 which exhibition have they to to been

Second, not all prepositions participate in the stranding operation in Frisian. In particular, a certain class of prepositions like *oarekant* – 'on the other side of', *fanwegen* – 'because of' do not participate in P-stranding. Even in English there are a few prepositions (and/or postpositions) like *notwithstanding*, *ago*, and *during* that do not allow stranding according to Culicover (1999:82). This is certainly not a compelling reason to reject the idea that English has a syntactic process of preposition stranding.

Hoekstra discusses several ways in which Frisian differs from Dutch, all of which are easily compatible with an analysis where Frisian does but Dutch does not have preposition stranding. Hoekstra does not discuss locality data that might show that there is a resumptive pronoun present in the complement position of the relevant prepositions.

Instead he offers the following facts to distinguish English, where the empty complement of a preposition is a trace, from Frisian, where he assumes it is a resumptive pronoun. In example (256a) sentence stress cannot fall on the final preposition in Frisian, although the sentence is acceptable with stress elsewhere, but in the English equivalent (256b) it can fall on the preposition. On the other hand Frisian and English behave the same way with respect to the possibility of placing sentential stress on the preposition in (257a-b).⁹⁹

- (256) a. {*Wa |*Hokker tiidrek} giet dat boek OER?
 who which period goes that book over
 b. {✓Who |✓Which period} is that book ABOUT?
- (257) a. ✓Wêr ljepte er OER?
 what jumpedhe over
 b. ✓What did he jump OVER?

Hoekstra claims that the facts in (256) illustrate a genuine difference between English and Frisian, which he casts in terms of trace vs. resumptive pronoun, and that the difference between the prepositions in (256) and (257) is that the former are functional whereas the latter are lexical. He offers the following characterization of functional prepositions. "[T]hey are semantically empty, fixed prepositions introducing a prepositional object" (Hoekstra 1995:109). He then offers the following generalization. "The ungrammaticality of [256a] seems to follow from the fact that a PP consisting of a functional adposition + a (non interrogative) pronoun can never be in focus (cf. Hoekstra 1991:72)" (Hoekstra 1995:110). Frisian then does not allow stress on the preposition because it is made up of a functional preposition and a non-interrogative pronoun.

⁹⁹ The fact that stress can easily fall on the preposition in (257a) makes a purely phonological account of (256a) unlikely.

English does allow stress on the preposition, because the PP is made up of a functional preposition and a trace rather than a non-interrogative pronoun.

I find Hoekstra's characterization of *oer* and *about* in (256) as functional prepositions dubious. Hoekstra is presumably right as far as Frisian goes. *Go about* is presumably a phrasal verb in Frisian so that there is no choice of P^o; hence, *oer* must occur with *giet*. This is clearly not true for English though. A book can be about, by, or for someone and it can be about or from some period. English *about* in the example is not a fixed preposition and therefore is not, by Hoekstra's criterion, a functional preposition.

To amplify this point, consider the examples in (258). *About* in (258a) and in (258b) take non-interrogative pronouns as complements. Yet they can be stressed; hence they are not functional prepositions by Hoekstra's criteria.¹⁰⁰

- (258) a. ✓The book is ABOUT him, not BY/FOR him.
b. ✓You say the book is FROM the second world war, but it's really ABOUT it.

On Hoekstra's own assumptions then, *about* in the English example in (256b) is not a functional preposition and the pair is irrelevant. Once we control for these factors and pick a preposition that is truly uniquely determined by the predicate in the way that *oer* is determined by the verb in (256a), the English stress facts change. This is illustrated in (259), where the verb *deal* is used, which on the intended reading (forced by the subject *book*) selects for a PP headed by *with*. (259a) shows that *with* is not stressable in this context and it therefore passes Hoekstra's test for being 'functional'.

¹⁰⁰ Unfortunately I do not at present have comparable Frisian data.

- (259) a. *..., but it really deals WITH it/him.
b. { *Who | *What period } does the book deal WITH?

Example (259b), which should be compared to (256a), then shows that there is no real difference between Frisian and English that would have to be accounted for in terms of the resumptive pronoun vs. trace distinction. The alleged difference between Frisian and English does not seem to exist. If it doesn't, the argument collapses.

We have seen that Hoekstra's reasons for assuming that Frisian PPs with an empty complement contain a null resumptive pronoun are not compelling. The fact that PPs with empty complements are allowed under comparative movement, example (251), when looked at against the backdrop of the crosslinguistic generalization developed in this section, is a strong argument against the assumption that prepositions with empty complements contain a null resumptive in Frisian and for the assumption that there is actual stranding. I thus believe that Hoekstra's facts show the exact opposite of what he thinks they indicate.¹⁰¹

¹⁰¹ It would be nice to be able to discuss the claim made in Alexandre and Hagemeijer 2002; Dijkhoff 1983; Hagemeijer 2000; Muysken 1977, 1980 for various Creoles, including most prominently Papiamentu, that these languages contain an audible, spelled out form of a trace with properties quite distinct from a resumptive pronoun. At the moment I do not have the relevant data at my disposal and have to leave this issue for further investigation, however, Dijkhoff 1983 claims that the spelled out trace is compatible with the comparative in Papiamentu which lends credence to the claim that the relevant item is indeed a spelled out trace. If this is true we could add Papiamentu to the short list of languages that allow P-stranding.

4.2.3 Conclusion

In this section I have discussed the possibility of teasing apart the syntactic function of various overt and covert elements that appear as the complement of prepositions. I have followed the standard claim that the amelioration of island effects otherwise operative within a language provides support for the hypothesis that a particular element is a resumptive pronoun – but crucially not the other way around; the absence of island amelioration tells us nothing. I have also claimed that resumptive pronouns never occur in comparative movement constructions. I concluded on the basis of the generalization about resumptives in comparatives that Frisian allows P-stranding. This conclusion is the exact opposite of that reached by Hoekstra (1995) on the basis of the same data.

4.3 No P-stranding in German and Dutch!

As discussed in van Riemsdijk's (1978a) German and Dutch do not have bona fide examples of extraction of full DPs from the complement position of P°. Example (260) illustrates the absence of *wh*-movement out of PP and (262) the absence of A-movement out of PP, i.e. the absence of pseudopassives in German. Pied-piping of the preposition under *wh*-movement leads to grammatical results, i.e. (260b) corresponds to the grammatical (261).¹⁰²

¹⁰² Most of the arguments presented in this section carry over to Dutch and some of the arguments I make a calques on arguments that have been made for Dutch. This will be noted. The overall conclusion I draw for German, namely that it does not allow P-stranding at all, carries over to Dutch as well.

- (260) a. *Was hast du mit gerechnet?
 what_{dat} have you with counted
 'What did you expect?'
 b. *Welchen Kandidaten hast du für gestimmt?
 which candidate have you for voted
 'Which candidate did you vote for?'
 c. *Welchem Bett hast du {in |drin} geschlafen?
 which bed have you in DR-in slept
 'Which bed did you sleep in?'
 (261) ✓Für welchen Kandidaten hast du gestimmt?
 for which candidate have you voted
 'Which candidate did you vote for?'
 (262) a. *Frank wurde vom Präsidenten mit geredet.
 Frank was by-the president with talked
 'Frank was talked to by the president.'
 b. *Der Kandidat von den Grünen wurde am meisten für gestimmt.
 the candidate from the Greens was at-the most for voted.
 'The candidate of the Green Party was voted for the most.'
 c. *Dieses Bett ist {in |drin} geschlafen worden.
 this bed is in DR-in slept been
 'This bed has been slept in.'

Data of this sort would straightforwardly lead to a categorization of German as a non-P-stranding language. Indeed German behaves like other non-P-stranding languages in almost every respect. It doesn't allow subcomparison on the complement of P° (263) – but does allow subcomparison of number into PP (264). The same is true in Dutch, where comparatives are disallowed on the complement of P° (see the discussion of Hoekstra 1995 above and Besten 1978, 1981), but subcomparison into PPs is allowed (Corver 1990, who also reports that Zwarts 1978 disagrees).

- (263) a. *Ich habe gegen mehr Kandidaten gestimmt als du für gestimmt hast.
 I have against more candidates voted than you for voted have
 'I voted against more candidates than you voted for.'
 b. *Ich bin mit mehr Typen ausgegangen als du mit ausgegangen bist.
 I am with more guys gone-out than you with gone-out are
 'I have gone out with more guys than you have gone out with.'

- c. *Ich habe mehr Büchergelesen als du {dar-; dr-; Ø-} über gehörlhast.
 I have more books read than you dar-; dr-; about heard have
 'I have read more books than you have heard about.'
- (264) ✓Ich habe nach mehr Büchern über die französische Revolution gefragt,
 I have after more books over the French revolution asked
 als du nach Büchern über die englische (Revolution) gefragt hast.
 than you after books about the English revolution asked has
 'I have asked about more books on the French revolution than you asked
 about books on the English revolution.'

Finally, neither German nor Dutch allow clitic pronouns especially with

inanimate reference as the complement of prepositions (Müller 2000 for detailed

discussion of the German facts and Riemsdijk 1978a for the corresponding claim for

Dutch).¹⁰³ As will be discussed in section 4.4, languages that allow preposition stranding

usually allow clitic pronouns as the complement of P°.

- (265) a. Ich habe mich auf {*'n; ✓ihn} verlassen.
 I have myself on him_{cl} him relied
 'I have relied on him.'
- b. Ich habe mich auf {✓'s Fahrrad; *'s} gesetzt.
 I have me on the bike; it_{cl} sat
 'I sat down on {the bike; it}.'

Dutch and German then look like typical examples of non-P-stranding languages.

The trouble is that there are cases where certain prepositions can occur without an overt

complement and that give the appearance of preposition stranding as in (266) and (267).

- (266) a. ✓Wo hast du mit gerechnet?
 where have you with counted
 'What did you count on?'

¹⁰³ The generalization is true with the exception of *ohne* – 'without', which has a number of other unusual properties as well (Riemsdijk 1975 makes the same point for Zürichdeutsch), for example that it can appear with a null complement in situations where this possibility is barred for other prepositions.

- b. ✓Wo hast du für gestimmt?
 where have you for voted
 'What did you vote for?'
- c. Wo hast du {*in | ✓drin} geschlafen?¹⁰⁴
 where have you in DR-in slept
 'What did you sleep in?'
- (267) a. ✓Da habe ich nicht mit gerechnet.
 there have I not with counted
 'I didn't expect that.'
- b. ✓Peter hat da gestern auch für gestimmt?
 Peter has there yesterday also for voted
 'Peter also voted for that yesterday.'
- c. Da hat niemand {*in | ✓drin} geschlafen.
 there has nobody in DR-in slept
 'Nobody has slept in it.'

These examples look like preposition stranding because there seems to be a gap

within PP and there is a displaced overt element that can be associated with this gap. In

the examples given so far, the displaced elements are the locatives *wo* – 'where' and *da* –

'there'. van Riemsdijk (1978a) calls these element R-pronouns.¹⁰⁵ He assumes that the R-

pronouns despite their independent life as adverbial modifiers are proforms for the DP

complement of prepositions and are thus pronouns. There are two independent claims

here: (i) R-pronouns are pronouns, i.e. proforms for DPs; and (ii) R-pronouns originate in

the complement position of P°. In the literature on German and Dutch, claim (i) is not

commonly held (but see Müller 2000; Riemsdijk 1978a). Claim (ii) on the other hand is

¹⁰⁴ The obligatory appearance of the form *drin* in this example is subject to a fair amount of dialectal variation as is the acceptability of these examples of preposition stranding in general (see Fleischer 2002 for discussion of this variation). The data discussed here are based on my own dialect. I turn to the *in-drin* alternation in subsection 4.3.4.

¹⁰⁵ They are called R-pronouns, because in Dutch they all contain an /r/-phoneme absent in other pronominal forms.

fairly commonly made (Bayer 1990; Bennis and Hoekstra 1984; Besten and Webelhuth 1990; Corver 1990; Fanselow 1983, 1991; Grewendorf 1990; Koopman 1997; Müller 2000; Riemsdijk 1978a, 1990; Riemsdijk and Williams 1985 but see Oppenrieder 1991; Trissler 1993 for counterarguments). Eventually I will follow Oppenrieder 1991; Trissler 1993 and reject both (i) and (ii).

I will first give some reasons for rejecting (i) (see Bennis and Hoekstra 1984; Trissler 1993). The first argument is based on an observation from Riemsdijk 1978b. The second argument is directly from Bennis and Hoekstra 1984; Trissler 1993. I then briefly illustrate the reasons that lead Trissler to reject the position according to which R-pronouns are not pronouns but do originate as complements of P°.

4.3.1 *R-R-pronouns-pronouns?*¹⁰⁶

van Riemsdijk 1978b discusses an effect that arises with R-pronouns in Dutch and that we can understand as a Relativized Minimality effect. The same effect is found in German and I'll illustrate it with German examples which are in relevant respects identical to van Riemsdijk's original Dutch examples. The analysis as a Relativized Minimality effect implies that R-words are not pronouns.

As noted, all R-words have homophonous locative adverbials. This is a coincidence if we treat R-words as proforms for DPs. Interestingly, R-pronouns interact for locality with their homophonous locative adverbials rather than interacting with pronouns. Consider example (268). The example only has the reading given in (268a) but

¹⁰⁶ Read: Are R-pronouns pronouns?

not the reading in (268b). There are two potential R-pronouns in the sentence. Only the closer one, *da*, can be construed with the gap in the PP. When the locative *da* is absent as in (269), the question is about the gap in the PP.

- (268) Wo hat er da gestern drüber nachgedacht?
 where has he there yesterday dr-about thought
 a. ✓ 'Where did he think about that yesterday?'
 b. * 'What did he think about there yesterday?'
 (269) ✓Wo hat er gestern drüber nachgedacht?
 where has he yesterday dr-about thought
 'What did he think about yesterday?'

Example (269) also shows that R-pronouns do not interact in this way with regular pronouns. There is a subject pronoun in (269) intervening between PP and *wo* in the same way *da* intervened between *wo* and the PP in (268). Yet, the pronoun does not block the association of *wo* with the PP. We can make sense of this observation (van Riemsdijk's) by appealing to Relativized Minimality (Rizzi 1990). If *wo* and *da* belong to the same class of items, then they should not be able to cross. Of course, *wo* is a *wh*-element and might thus be expected to cross the non-*wh*-word *da*. As discussed in van Riemsdijk 1978a, R-words can move from the PP where they originate into a position in the upper middle field of the clause, immediately following weak pronouns (270).

- (270) ✓... weil er ihr da gestern [t_{da} drüber]berichtet hat
 ...because he her there yesterday about reported has
 '...because he reported to her about that yesterday.'

Suppose that *wh*-movement of *wo* is a two step process. First *wo* undergoes movement to the R-position, and from there it undergoes *wh*-movement. Movement to the R-position is sensitive to intervention by R-words, but the further movement to [Spec, CP] is not. This account bears obvious similarities to the idea that object *wh*-movement

passes through [Spec, Arg] (Bošković 1997a; 1997b; Ura 1993). This blocks construal of *wo* with PP in (268). On the other hand, two elements that belong to different classes do not interfere with each others movements.

On the basis of this discussion, the intervention of the pronoun in (268) and (269) is irrelevant. Notice however that the pronoun also does not intervene with R-movement in an example like (271), where R-movement lands above the site of the pronoun *ihr*.

- (271) ✓...weil er da gestern IHR drüber berichtet hat.
 ...because he there yesterday her about reported has
 '...because he reported to HER about it yesterday.'

The lack of interaction between pronouns and the R-word in (269) then indicates that they belong to different classes. The conclusion to draw from such facts is that R-pronouns are not pronouns at all but rather locative proforms.

We have just seen data which suggest that pronouns and R-words belong to different classes. We can also show that traces of R-words and traces of DPs belong to different classes. The argument and the data are taken from Trissler (1993:260-261), who in turn adapts an argument about Dutch from Bennis and Hoekstra 1984.¹⁰⁷ The argument has to do with parasitic gap licensing. In the examples below '*t*' indicates the rough position of the real trace and '*e*' the rough location of the parasitic gap under the analysis of Bennis and Hoekstra. The trace and the site of the parasitic gap precede the prepositions (postposition) because Bennis and Hoekstra give a postpositional analysis to

¹⁰⁷ Trissler speaks a dialect, not mine, that allows relative clause formation in general with the relative particle *wo*. The construction is similar to *that*-relative clauses in English. Despite the fact that I don't generally allow relative clauses with the relative particle *wo*, the contrasts discovered by Trissler still hold for me.

R-words. For ease of reference I will distinguish between DP-gaps, gaps that can be anteceded by regular DP relative and interrogative pronouns, and R-gaps, gaps that require an R-word as their antecedent.

- (272) a. ✓ein wichtiges Buch, das statt nur *e* zu zitieren
 an important book which instead-of only to quote
 jeder *t* kaufen sollte
 everybody buy should
 'an important book that everybody should buy instead of just citing it'
 b. ??ein wichtiges Buch, wo statt nur *e* zu zitieren
 an important book where instead-of only to quote
 jeder *t* kaufen sollte
 everybody buy should
 'an important book that everybody should buy instead of just citing it'
 c. ??ein Vorhaben, wo er statt dauernd [*e* von] zu reden
 a project which he instead-of always about to talk
 mal besser [*t* mit] anfangensollte
 PRT better with start should
 'a project that he should start with instead of constantly talking about it'
 d. *ein wichtiges Buch, wo statt dauernd [*e* von] zu reden
 an important book where instead-of always about to talk
 jeder *t* kaufen sollte
 everybody buy should
 'an important book that everybody should buy instead of just talking about it'
 e. **ein wichtiges Buch, das statt dauernd [*e* von] zu reden
 an important book that instead-of always about to talk
 jeder *t* kaufen sollte
 everybody buy should
 'an important book that everybody should by instead of just talking about it'
- (273) ✓/??This is the guy that I suggested [to *e*] that my brother could offer *t* a job.

Example (272a) shows that parasitic gaps can easily be licensed by the relative pronoun *das*. Example (272b) shows that *wo* can antecede DP-gaps and that it can, in that case, marginally license a parasitic DP-gap. Example (272d) shows that sentences where *wo* binds an R-gap and a parasitic R-gap have roughly the same status as examples where *wo* binds a DP-gap and a parasitic DP-gap (272b). Examples (272e) and (272c) are

completely ungrammatical. In them *wo* is associated with a DP-gap and a parasitic R-gap and an R-gap and a parasitic DP-gap respectively. DP gaps and R-gaps do not mix, it seems. English exhibits markedly different behavior as (273), which is at worst slightly degraded, shows (Bennis and Hoekstra 1984:61).

Bennis and Hoekstra 1984 and Trissler (p. 261) conclude from these data, correctly I believe, that “traces of R-pronouns and traces of DP/NP-complements belong to different categories” (Trissler, 1993:261 [translation K.A.]). This would follow if R-pronouns are not DP/NPs. The silent operator that moves to [Spec, CP] in examples (272b) is a DP, the silent operator that moves to [Spec, CP] in (272c) is not. In English DPs can be extracted from PPs. The DP-operator is associated with a DP-gap and a parasitic DP-gap. This explains why (273) is acceptable.

It seems reasonable to conclude that the category left behind by R-words is not the same as the category left behind by DPs, which follows if R-words and DPs belong to different categories, i.e. if R-words are not pronouns.

We have seen two sets of data that suggest contra Müller 2000; Riemsdijk 1978a that R-words are not DP proforms.¹⁰⁸ The facts fall into place if we assume that R-words

¹⁰⁸ Müller’s (2000) optimality theoretic account of the distribution of R-pronouns is based on the fact that weak pronouns in German have a requirement to appear in the Wackernagel position towards the beginning of the clause. Since German disallows P-stranding, weak pronouns with a Wackernagel requirement cannot appear as the complement of prepositions. This reasoning is plausible and has a long tradition (Behagel 1932 §1481; Blatz 1896:256). It explains why weak pronouns cannot occur as the complement of P° in German. However, Müller assumes that R-words are the unfaithful realization of weak pronouns that becomes optimal if a high ranking Wackernagel requirement is violated.

are adverbs like their locative homophones – indeed traditional grammar in German calls these items pronominal adverbs (Pronominaladverbien). Notice that both arguments reported here were originally based on Dutch. The conclusion reached for German therefore holds for Dutch as well.

4.3.2 Are R-words P°-complements

I introduced two assumptions from van Riemsdijk (1978a) above: (i) R-pronouns are pronouns, i.e. proforms for DPs; and (ii) R-pronouns originate in the complement position of P°. We have now seen reasons to reject (i). This leaves the option of assuming (ii) while rejecting (i). Rejecting (i) amounts to claiming that there are significant inherent differences between pronouns (and DPs more generally) and R-words. This would allow us to pin the difference in behaviours on the inherent differences between DPs and R-words. In particular, the inherent differences might allow deriving or at least describing the fact that DPs never strand P° while R-words apparently do.

The reasoning carries over to other instances of weak pronouns that cannot reach the Wackernagel position. Consider example (i). The adjective *gewöhnt* – ‘used to’ is a rare case of an adjective in German that both assigns accusative case and may appear adnominally. It may in fact be the only such adjective in the present language. With a full DP as the adjectival complement, the example is acceptable. However, a weak pronoun cannot replace this DP. This follows, plausibly from the Wackernagel requirement on weak pronouns and the assumption that the complex NP forms an island. However, contrary to Müller’s prediction the R-word *da* cannot replace the weak pronoun. This is obviously problematic for Müller.

- (i) Ich kenne [_{NP-island} einen [_{AP} {✓schweres Essen | *es | *da} gewöhnten] Mann]
 I know a heavy food it_{prn.wk} R_{dem} used-to man
 ‘I know a man used to heavy food.’

My general claim laid out in the introduction to this chapter is that the absence of P-stranding in general and in German examples like (260) and (262) above in particular follows from the Stranding Generalization. The Stranding Generalization does not allow for exceptions. The complement of a particular head is either frozen or not depending on whether the head is a phase head or not. The properties of the complement are irrelevant for the operation of the Stranding Generalization. It follows straightforwardly from this logic that R-words cannot begin their lives as the complements of P° no matter how different they are from DPs. In other words I crucially have to reject (ii) above.¹⁰⁹

Clearly then I cannot adopt the position that R-words originate as the complement of P° and that the distinction between DPs, which do not extract, and R-words has to do with the inherent features of R-words. This is essentially the position taken by the above authors who adopt position (ii) while rejecting (i).

Oppenrieder 1991; Trissler 1993 argue that no significant generalizations are captured by making assumption (ii) while rejecting (i). Trissler in particular shows that under all existing accounts that adopt (ii), the most straightforward extensions of the proposal must be blocked, which essentially amounts to stipulating not only the inherent properties of R-words but also their syntax. Since the syntax ends up stipulated, the inherent properties of R-words end up not playing a real role even in these accounts. They are thus at best on a par with the present approach.

¹⁰⁹ The observant reader will have noticed that the discussion from the previous subsection does not play a role in these considerations. I must reject (ii) whether R-words are pronouns or not. I began by clarifying the categorial status of R-words because these assumptions influence what form (ii) can take. My prior rejection of (i) greatly simplifies

Take as an example of the kind of issue raised by Trissler the claim made in Bennis and Hoekstra 1984 that R-words appear to the left of prepositions because they are case resistant and that prepositions assign case to the right. This amounts to saying that prepositions can take complements in either direction and the actual place where the complement appears is in principle free. In the ideal case it would be constrained only by the inherent properties of the complement itself. The relevant property in this case is the need for case or the inherent resistance to it. This derives the pattern in (274).

- (274) a. ✓mit dem Mann
 with the man
 b. ✓damit
 there-with
 c. *dem Mann mit
 the man with
 d. *mit da
 with there

A first problem (pointed out to me by Željko Bošković p.c.) has to do with the question whether prepositions obligatorily assign Case or not. If P° obligatorily assigns Case, then we lose the explanation of (274b), where by Bennis and Hoekstra's assumption P° does not assign Case. However, if P° does not necessarily assign Case, then it becomes unclear why (274d) is ungrammatical. This is clearly a problem for the Case resistance account of the distribution of R-words.

Trissler asks how the account might extend to other cases. Prepositions can take PPs as their complements. Assuming that prepositional phrases do not require case marking, we expect PPs to be able to occur on either side of prepositions in German

the discussion of (ii).

(275a) and (275b) – if we assumed that they are Case resistant, they should actually only occur to the left of prepositions, i.e. only (275b) should be acceptable. Since for Bennis and Hoekstra extractability of the complement of P° is linked to its appearing in a postpositional structure, we also expect PPs to be extractable from PPs, as exemplified in (275c) and (275d). Neither expectation is borne out.

- (275) a. ✓von vor der Wahl
from before the elections
b. *vor der Wahl von
before the elections from
c. ✓Von vor der Wahl ist dieses Plakat.
from before the election is this poster.
'This poster is from before the elections.'
d. *Vor der Wahl ist dieses Plakat von.
before the elections is this poster from

The linkage of Case properties with directionality and extractability then doesn't appear to work. Consider now the question whether *da* can really be said to be case resistant. The preposition *von* – 'from' in its directional use can assign Case to its complement (276a). The DP *dem Bahnhof* designates the origin of the movement. In place of the DP we can use *da* as a locative (276b). In this case the order must be as given in (276b) and (276d). The reverse order (276c) is grammatical but does not have the directional interpretation. In its locative use *da* cannot be extracted (276e).

- (276) a. ✓{von dem |vom} Bahnhof
from the from.the station
b. ✓von da
from there
c. davon
there.from
d. ✓Das Auto kam von da.
the car came from there
'The car came from there.'

- e. *{Da |Wo} kam das Auto von.
there where came the car from

There are two possible conclusions. *Da* in (276) is a homophone of the R-word *da* with substantially different properties or the Case resistance/leftward complement story is wrong. Taking the first option brings us back to the evidence from section 4.3.1 where I showed that the R-word *da* and the locative *da* are substantially the same category. This leaves us with the second option. We give up the leftward generation story (because of (276b) and (276d)); we give up the Case resistance story (because of (276b), (276d) and (275); and we give up the notion that *da* can ever be extracted from the complement of P° (because of (276e)). For the paradigms just discussed, no significant generalizations are captured by assuming that R-words are leftward complements of P° . Trissler shows that similar problems arise not only under Bennis and Hoekstra's version of the thesis that R-words are born as complements of P° but under every existing implementation of that assumption. Given that we gain little or nothing by assuming that R-words are generated as the complements of prepositions, I will reject this assumption as well as the assumption that they are pronouns.

4.3.3 So what are R-words? And where?

I have so far ignored a fact in the discussion that becomes relevant now. Some prepositions undergo (d)r-epenthesis when they combine with an R-word. In particular

the prepositions that have an initial vowel participate in this alternation (see Fleischer 2002 for a complete list).¹¹⁰

- | | | | |
|----------|--------|----------------------------|---------------------------------------|
| (277) a. | auf | – darauf, worauf | on, there -on, where-on |
| | über | – darüber, worüber | over, there-over, where-over |
| | unter | – da(d)runter, wo(d)runter | under, there-under, where-under |
| | in | – da(d)rin, wo(d)rin | in, there-in, where-in |
| b. | von | – davon, wovon | from, there-from, where-from |
| | vor | – davor, wovor | before, there-before, where-before |
| | gegen | – dagegen, wogegen | against, there-against, where-against |
| | hinter | – dahinter, wohinter | behind, there-behind, where-behind |
| | für | – dafür, wofür | for, there-for, where-for |

The alternation is obligatorily retained when the R-word and the preposition are split. Recall example (266c), which illustrates this.¹¹¹

- (266) c. Wo hast du {*in |✓drin} geschlafen?
 where have you slept
 ‘What did you sleep in?’

Items like *drauf*, *drüber*, *drunter*, *drin* etc. do not take complements. It is plausible that they are different lexical items from the prepositions *auf*, *über*, *unter*, *in*, etc. The members of the latter class systematically take complements, the members of the former do not. We could in fact assume that (d)r- is the complement of P° incorporated into it. If it is sound to postulate different lexical items for *drauf* and *auf*, then it seems equally sound to postulate that there are two homophonous lexical items for each of the remaining prepositions like *von*, *vor*, *gegen*, *hinter*, and *für* that do not participate in (d)r-

¹¹⁰ The same kind of morphological reflex of the presence/absence of R-words is also found in Dutch though there the distinction is marked only on two prepositions: *met DP* vs. *waarmee* – ‘with’ and *tot DP* vs. *waartoe* – ‘to’.

¹¹¹ These are all presumably very low level constraints. Fleischer 2002 in fact records a lot of dialectal variation.

epenthesis.¹¹² Given their distribution, we have to treat these new lexical items as zero-place prepositions. This has been suggested for particles (see already Emonds 1972) with which the items under discussion are homophonous. This is certainly not an accident.

So far my suggestion says nothing about the fact that the postulated zero-place prepositions obligatorily co-occur with R-words and that the relation between the R-word and the preposition must be local. The relation between R-words and the PPs they associate with is constrained by standard locality conditions. Both facts were elegantly captured under the approach where the R-word originate in the complement position of P°.

In the nominal domain both of these properties are not surprising to find even outside of movement constructions. We call items that obligatorily take a local antecedent anaphors. We could simply assume then that the postulated zero-place prepositions are prepositional anaphors. Alternatively we could assume (Oppenrieder 1991) that R-words are generated in a PP-adjoined position or in [Spec, PP] (Trissler 1993).

To decide between these options consider the following facts.

We need to allow base-generation of R-pronouns in [Spec, PP] or a PP-adjoined position independently since they occur in that position even in PPs where the complement is lexically filled. The Dutch example (278) from Corver (1990:72 fn. 38 ex. (i)) and the German example (279) may serve as an illustration.

¹¹² This move of describing the relevant facts in the lexicon has the advantage that it can easily capture the fact that only certain prepositions combine with R-pronouns at all. *Wegen* – ‘because of’, *trotz* – ‘despite’, *ohne* – ‘without’, *seit* – ‘since’, *mittels* – ‘by means of’ and in fact most prepositions that are not of the oldest stock of German prepositions (see Müller 2000) cannot combine with R-words.

- (278) ✓Ergens (diep) (bij Marie) (achter) in de keel zit een zweer
 somewhere deep with Marie back in the throat is an abscess
 'There is an abscess somewhere far back in Mary's throat.'
- (279) ✓Wo im Hals tut es dir denn weh?
 where in.the throat does you it then hurt?
 'Where is the pain located in your throat?'

Since we need to allow the possibility of generating R-words in [Spec, PP] or PP-adjointed position anyway, the most plausible assumption would be that they are always generated in that position.

Corroborating evidence comes from extraction out of PP (see Corver 1990 chapter 9 for extensive discussion). A generalization that is often claimed to restrict the occurrence of R-words in German is that P° must be left-adjacent to the verb before and after movement of the R-word. As shown in Fleischer 2002; Oppenrieder 1991; Trissler 1993, there are counterexamples to the generalization. The examples in (280) are taken from Trissler 1993:288. Examples from actual use like those in (281) are discussed in Fleischer (2002:135, the East Franconian example (281b) is Fleischer's (p. 136) example (12c')). Fleischer notes that Abraham (1995:337) judges these to be ungrammatical.¹¹³

- (280) a. ✓Da hat er ihm mit auf den Kopf geschlagen.
 there has he him with on the head hit
 'He hit him on the head with that.'
- b. ✓Da ist er doch hoffentlich mit zum Arzt gegangen.
 there is he PRT hopefully with to.the doctor went
 'Hopefully he went to see a doctor about this.'
- (281) a. ✓Da will ich nichts mehr hören von.
 there want I nothing more hear from
 'I don't want to hear anything about this any more.'

¹¹³ Corver (1990:42) provides Dutch examples where PP is a complement of AP and can be stranded by R-movement if it is either left or right adjacent to the adjective, but not if it is extraposed out of AP. This suggests that speakers who allow the patterns in the examples in the text can generate them without extraposition.

- b. dā hāt er niks gsächt driwer
 there has he nothing said DR-about
 'He didn't say anything about it.'

Examples like these exist, but they are rare. In the usual case R-words cannot split from P° if P° does not occur adjacent to the selecting head. In fact, as discussed in Corver 1990 chapter 9, extraction from PP degrades if the PP is an adjunct or if it is extraposed. This corroborates further that the relation between R-words and zero-place PPs is mediated by movement since we know that extraction out of adjuncts and out of moved constituents is in general degraded. Similar conditions are needed for English, too, where generally P-stranding degrades if a PP is not argumental or if it is extraposed (see section 4.5.3). The same conditions also govern the extractability of left branch modifiers of PPs (Corver 1990 chapter 9).

I conclude that R-words are base generated as the specifiers (or adjuncts) of distinguished class of zero-place prepositions. This position virtually forces us to adopt the position that (d)r- is the complement of P°, since in the absence of a complement there is no distinction between the complement and the specifier.^{114, 115}

¹¹⁴ As discussed in chapter 2, we must assume that the intervening features on P° do not percolate. This being so, the P'-sister of the R-word in its base position does not block extraction out of PP.

¹¹⁵ Fleischer 2002 notes examples from various dialects where these zero place prepositions have pronominal uses rather than anaphoric uses in the sense that they do not even require a local R-pronoun (Fleischer 2002:128 ex. 127c citing the Badisches Wörterbuch (1925)).

(i) Bhalt deiⁿ Messer, ich kann nigs mit schneide!
 keep your knife I can nothing with cut
 'Keep your knife, I cannot cut anything with it.'

4.3.4 Extraction out of PP in German

Assuming that German does not allow P-stranding, it is interesting to note that just like the Slavic non-P-stranding languages and like French, German allows extraction out of PP (see already the subcomparative examples (263)). The clearest examples turn out to come from R-words. We just reached agreement with the rest of the literature on German that R-pronouns are base generated as a constituent (though not the complement, which, when visible, is occupied by *(d)r-* with the preposition they associate with. Now consider examples (282) and (283). The *für*-PP and the *gegen*-PP are clearly part of NPs headed by *Argumente*. It is impossible for example to drop *über/an Argumente* without also dropping the *gegen/für* PP. It is impossible to pronominalize *Argumente* without dropping the *gegen/für* PP etc. Therefore the examples involve extraction of the R-pronoun from the embedded PP across the embedding P° and to the clause initial position. It is hard to imagine clearer evidence for extraction out of PP. Notice that P-stranding here is also completely impossible.¹¹⁶

- (282) ?Da habe ich lange [_{PP} über Argumente [_{PP} gegen t_{da}]] nachgedacht.
 there have I long over arguments against thought
 'I thought about arguments against that for a long time.'
- (283) ✓Da kann ich mich [_{PP} an kein Argument [_{PP} für t_{da}]] erinnern.
 there can I self on no argument for remember
 'I cannot recall an argument for that.'

¹¹⁶ Corver discusses structurally similar examples in Dutch (Corver 1990:298). Riemsdijk 1978a:193 judges Dutch examples very similar to the ones given here ungrammatical.

Another relevant example of extraction from PP is given in (284) for which I assume the analysis in (285). The claim is that the PP [*über welches Thema*] originates as part of the DP [*einem Buch über welches Thema*] and is extracted from within the containing PP headed by *nach*.

- (284) ✓Über welches Thema hast du mich noch mal nach einem Buch gefragt?
 About which topic have you me again after a book asked
 'Which topic did you ask me about a book on again?'
- (285) [_{PP} über welches Thema] ... [_{PP} nach [_{DP} einem Buch t_{über welches Thema}]]

We need to rule out several possible alternative analyses. First of all we need to make sure that the extracted PP can indeed be contained within the PP headed by *nach*. That this is the case is shown in example (286a), where the entire *nach* PP is pied piped under question formation.¹¹⁷ (286b) and (286c) show that the *über*-PP cannot act as an independent argument to the verb (in other words, the factors discussed in Bach and Horn

¹¹⁷ It is impossible that *über den Krieg* in (286a) is not contained within PP but is, maybe, part of a VP that is remnant moved to the front. The remnant movement derivation is ruled out, because remnant VP movement does not occur under question formation (Koopman 1997). See examples (i) and (ii), the latter of which is built on an example from Müller 1998 discussed above in chapter 3. The example involves remnant VP fronting where the verb has moved out of its VP.

- (i) a. ✓Ein Buch vorgelesen hast du ihr.
 a book read have you to-her
 'I read a book to her.'
- b. *{welches Buch |*was} vorgelesen hast du ihr.
 which book what read have you to-her
- (ii) a. ✓Kindern Bonbons gibt man besser nicht.
 childrensweets gives one better not
 'It's better not to give sweets to children.'
- b. *Wem Bonbons gibt man besser nicht?
 who sweets gives one better not
- c. *Kindern was gibt man besser nicht?
 children what gives one better not

1976; Cattell 1976; Chomsky 1977 are controlled for). Example (286d) shows that *nach* does not show any exceptional possibility of being stranded. It follows that the *über*-PP in (285) starts out within the *nach*-PP and is extracted out of it.

- (286) a. ✓Nach welchem Buch über den Krieg hast du mich noch mal gefragt?
 After which book about the war have you me again asked
 'Which book did you ask me about again?'
 b. ?*Über welches Thema hast du mich noch mal gefragt?
 about which topic have you me again asked
 c. ?*Du hast mich danach über den Krieg gefragt.
 you have me there-after about the war asked
 d. *Welchem Buch hast du mich noch mal nach gefragt?
 Which book have you me again after asked

A similar example is given in (287), with the relevant controls in (288). Notice that example (287) is not fully grammatical, but there is a very sharp contrast between it and the stranding example (288d).

- (287) ?Von welchem Maler würdest du dich über ein Portrait freuen?
 From which painter would you yourself about a portrait be-happy
 'Which painter would you be happy about a portrait by?'
 (288) a. ✓Über welches Portrait von Van Gogh würdest du dich freuen?
 About which portrait from Van Gogh would you yourself be-happy
 'Which portrait would you be happy about?'
 b. *Von welchem Maler würdest du dich freuen?
 From which painter would you yourself be-happy
 c. *Du würdest dich darüber von van Gogh freuen.
 you would yourself there-about from van Gogh be-happy
 d. *Welches Portrait würdest du dich über freuen?
 Which portrait would you yourself about be-happy

The next class of cases of extraction from PP comes from *was-für*-split. A simple case of *was-für*-split is given in (289) where (289a) is the unsplit version and (289b) the

split version. *Was-für*-split in Germanic is discussed in Beck 1996; Bennis, Corver and Dikken 1997; Besten 1985; Corver 1991; Diesing 1992; Kratzer 1996; Leu 2003 for a brief survey; Pafel 1996; Vikner 1995.

- (289) a. ✓Was für Bücher hast du gelesen?
 Whatfor books have you read
 b. ✓Was hast du für Bücher gelesen?
 Whathave you for books read
 'What kind of books did you read?'

There are two plausible analyses for *was-für*-split, the direct extraction analysis (289b') and the remnant movement analysis (289b''). The situation is entirely parallel to the Slavic case discussed at some length above in section 4.1.2.

- (289) b.' Direct Extraction: Was ... [DP t_{was} [für Bücher]]
 (289) b.'' Remnant Movement: [Was t_{für Bücher}] ... [für Bücher] t_{was für Bücher}

Now consider what happens if a *was-für*-DP is embedded within a PP. If we assume the direct extraction analysis of *was-für*-split, we expect the (a) examples in (288) through (291) to be ungrammatical. The clause initial position must be occupied by a *wh*-phrase or a *wh*-phrase with pied-piped material. If *was* underwent direct extraction in *was-für* split constructions we would not expect it to form a constituent with the embedding preposition. Hence, the (a) examples cannot be generated. The (b) examples may or may not be predicted to be grammatical, depending on whether extraction from PP is to be allowed. The remnant movement analysis gives rise to exactly the opposite

-
- d. *Wem was gibt man besser nicht?
 who what gives one better not

expectations. The actual judgments follow straightforwardly from the remnant movement analysis but are very problematic under the direct extraction analysis.¹¹⁸

- (290) a. ✓Über was der alles für Sachenbescheid weiß!
About what that all for things is knowledgeable

¹¹⁸ There is a construction in Russian that might provide further evidence for extraction out of PP. The construction is known as the *čto-za*-construction and it is broadly similar to *was-für*-construction in Germanic. In Russian the *čto-za* construction is limited to nominative and accusative environments. Directional prepositions are also incompatible with the *čto-za*-construction in Russian.

There are some cases where *na* and *v* can be selected by a verb and assign accusative case without the directional interpretation. Two such cases are given in (i) and (ii). According to Natalja Rakhlin these examples are compatible with *čto-za*-split. The preposition must move together with the *wh*-word *čto* – ‘what’ (ib) and (iib). Direct extraction of *čto* from the PP is sharply ungrammatical as (ic) and (iic) show. This is fairly direct evidence that *čto-za* split is derived by remnant movement. For reasons that I do not understand the unsplit versions of (ib) and (iib) given in (id) and (iic) are slightly degraded. Other speakers I have consulted disagree with the judgments and find (ib) and (iib) ungrammatical – though (iic) “may be probably said in colloquial unprepared speech” (Yakov Testele, in an email 2003-06-01).

- (i) a. ✓On nadeetsja na podarok.
he hope.refl on gift.acc
‘He is hoping for a gift.’
b. ✓Na čto on nadeetsja za podarok?
on what he hope.refl for gift.acc
‘What kind of gift is he hoping for?’
c. **Čto on nadeetsja na za podarok?
what he hopes on for gift
d. *Na čto za podarok on nadeetsja?
on what for gift.acc he hope.refl
- (ii) a. ✓On odet v brjuki.
he dressed in pants.acc
‘He is wearing pants.’
b. ✓Vo čto on odet za brjuki?
in what he dressed for pants.acc
‘What kind of pants is he wearing?’
c. **Čto on odet v(o) za brjuki?
what he dressed in for pants.acc
d. *Vo čto za brjuki on odet?
in what for pants.acc he dressed

- b. *Was der alles über für Sachenbescheid weiß!
What that all about for things is knowledgeable
c. ✓Über was für Sachen der alles bescheid weiß!
About what for things that all is knowledgeable
‘All the things he knows about!’
- (291) a. ✓Über was sich manche Leute nur für einen Quatsch aufregen!
About what self some people only for a nonsense get-upset
b. *Was sich manche Leute über für einen Quatsch aufregen!
What self some people about for a nonsense get-upset
c. ✓Über was für einen Quatsch sich manche Leute aufregen!
About what for a nonsense self some people get-upset
‘What nonsense some people can get upset about!’
- (292) a. A: I didn’t expect THIS kind of people to show up at my party. –
B: ✓Mit was hattest du denn für Leuten gerechnet?
With what had you PRT for people reckoned
b. *Was hattest du denn mit für Leuten gerechnet?
what had you PRT with for people reckoned
c. ✓Mit was für Leuten hattest du denn gerechnet?
with what for people had you PRT reckoned
‘What kind of people had you expected?’
- (293) a. ✓An was hatten sie denn so für Preise gedacht?
on what had you then so for prizes thought
‘What kinds of prizes were you thinking about?’
b. *Was hatten sie denn so an für Preise gedacht?
what had you then so on for prizes thought
c. ✓An was für Preise hatten sie denn so gedacht?
on what for prizes had you then so thought
‘What kinds of prizes were you thinking about?’

These examples, though perfectly acceptable with the right intonation, are not neutral. One class of examples that favours *was-für*-split into PPs are exclamations as in (290) and (291).¹¹⁹ In these cases there is stress on the nouns *Sachen* and *Quatsch* respectively. The questions behave strikingly differently. Without sufficient context, the unsplit

¹¹⁹ Case on the DP following *für* is assigned by the moved preposition. The only example that shows explicit case morphology on the DP is (291). Some speakers find example (291) degraded while accepting the others. This probably has to do with the licensing of overt case morphology.

questions (292c) and (293c) are strongly preferred.¹²⁰ In fact (292a) and (293a) are unacceptable if they are given the intonation of standard examples of *was-für*-split as in (289b), where stress falls on *Bücher*. In the questions, the nouns *Leute* and *Preise* cannot be stressed. Rather stress must fall on *denn* in (292a) and on *gedacht* in (293a).¹²¹

Yet another case of extraction from PP might be relative clause extraposition as in (294) if it is assumed to involve movement.¹²²

- (294) ✓Ich habe gestern von Sachen geträumt, die mir gut gefallen.
 I have yesterday from things dreamt that me good please
 'I dreamt about things yesterday that please me a lot.'

¹²⁰ Sigrid Beck (p.c.) points out that such examples are perfectly acceptable as indirect questions.

¹²¹ If German split partitives are derived via movement, there is clear evidence for extraction out of PP. Compare (i) with the split variant (ii) and (iii) with the split variant (iv). However, the behavior of split partitives is quite distinct from the behavior of other extractions from DP in German, which casts some doubt on an extraction analysis.

- (i) ✓Viele von denen gefallen mir.
 many of them please me
 'Many of them please me.'
- (ii) ✓Von denen gefallen mir viele t_{von denen}.
 of them please me many.
 'Many of them please me.'
- (iii) ✓An die meisten von denen kann ich mich noch erinnern.
 on the most of them can I refl. still remember
 'I can still remember most of them.'
- (iv) ✓Von denen kann ich mich noch [an die meisten t_{von denen}] erinnern
 of them can I refl. still on the most remember
 'I can still remember most of them.'

¹²² Interestingly, the same is true in English as the gloss indicates. This is somewhat surprising. For unknown reasons, P-stranding is disallowed under another type of rightward movement, Heavy NP Shift (Ross 1967). (Relevant data was discussed in section 3.1.3.) Here we have a case where English rightward movement allows subextraction out of PP but P-stranding is disallowed, i.e. we have a case of the complement non-complement distinction for prepositions even in English.

These are bona fide examples of extraction out of PP. If we treat German as a non-P-stranding language as I argued above, then we can see that PPs are not islands for extraction in German, rather PPs obey the Stranding Generalization (193).

4.4 Pronominal Clitics as the complement of P°¹²³

In this section I discuss the generalization that in most languages pronominal clitics do not readily appear as the complements of P° (Abels 2003b, in press) and show how it follows from the assumptions made in this thesis once the crucial additional assumption is made the clitics must undergo syntactic movement processes that other DPs are exempt from.¹²⁴ The data presented in this section then provides a new argument for the assumption that clitics obligatorily move. Descriptively, the ban is captured by filter (295), illustrated by the examples in (296) from various languages.¹²⁵ In all cases when a full pronoun is used instead of the clitic pronoun, the example becomes acceptable.

¹²³ For discussion of the content of this section I thank the audiences at the Acme Balkanica Workshop (Montreal, April 2001), at FDSL 4 (Potsdam, November 2001), of Arhonto Terzi's class on prepositions at the 8th EGG-School (Niš, August 2001), at FASL 11 (Amherst, May 2002), and my class on P-Stranding at the 9th EGG-School (Novi Sad, July/August 2002). Thanks also to Boban Arsenijević, Željko Bošković, Wayles Browne, Anders Holmberg, Howard Lasnik, Mariana Lambova, Luisa Martí, Ad Neeleman, Christer Platzack, Anna-Lena Wiklund, and finally the two FASL 11 reviewers.

¹²⁴ The observation that clitics cannot appear as the complements of prepositions is not new at all. The observation is often made in passing without attempting to give an account.

¹²⁵ In Greek there are some prepositions that do take clitics as their complements and, in fact, almost exclude full DPs as complements c.f. Terzi 2001 for an approach compatible with the assumptions made here. The same is true of the Bulgarian

- (295) * [P clitic]
 (296) a. Serbo Croatian
 Prema { **joj*/ ✓*njoj* } trče.
 Toward her_{cl} her_{pm} run
 ‘They run towards her.’
 b. Spanish
 Sobre { **la*; ✓*ella* } habló Pedro.
 about her_{cl} her talked Pedro
 ‘Pedro talked about her.’
 c. Greek
 Gia { **ton*; ✓*afton* } milise o Petros.
 about him_{cl} him talked the-P
 ‘Petros talked about him.’
 d. Hausa (Tuller 1986:280, 356 ex. 260)
 Sun zoo da { **shi*; ✓*shii* }.
 3p come with him_{cl} him_{pm}
 ‘They came with him.’
 e. Ich habe mich auf { **n*; ✓*ihn* } verlassen.
 I have myself on him_{cl} him relied
 ‘I have relied on him.’

The first part of this section investigates the nature of the ban in (295). The thesis is that the filter follows from two independent properties of the grammar: (i) the ban against preposition stranding and (ii) the requirement that clitics must move even where other DPs need not move. On the basis of these assumptions the typological correlation (297) is predicted, which turns out to be true.

- (297) A language allows clitic pronouns as the complements of P iff that language allows P-stranding and has clitic pronouns.

I will now show that filter (295) follows from the ban against P-stranding.

As a first step, we ask whether (295) holds universally or whether there is variation. The answer is that (295) must fall under some parameter, since there are

preposition *pomeždu* – ‘between’ Hauge 1999.

languages that violate it freely: English (298), Swedish (299), Icelandic (301), Norwegian (302), and Gbadi (303).

Reduced pronouns in English are clitics. Bošković (1997b:59; 2001a) argues for their clitic status on the basis of *wager*-class verbs and quantifier float respectively.

Lasnik 2002 argues for their clitic status on the basis of binding facts in ECM-constructions. I will thus assume that (298) is a true counterexample to (295).

- (298) ✓ We talked about ‘im for quite some time.

Swedish *na* is a clitic in the relevant dialect (see Hellan and Platzack 1999). This can be seen in (300), which illustrates that *na*, unlike full DPs, strong and weak pronouns, obligatorily object-shifts. Thus, *na* undergoes movement from which other DPs are exempt, a typical clitic behavior (examples from Anna-Lena Wiklund, p.c.).

- (299) ✓ Han trodde på na.
 He believed on her
 He believed in her.
 (300) Han möter *(na) alltid *(na) snabbt *(na)
 he meets her always her quickly her
 på lördagar före lunch.
 on Saturdays before lunch
 He always meets her on Saturday before lunch.

For the claim that Icelandic ‘*ana* and Norwegian ‘*a* are clitics, see Hellan and Platzack 1999.

- (301) ✓ Ég hugsaði um ‘ana.
 I thought about her_{cl}
 I thought about her. Hellan and Platzack 1999 p. 128 ex. 15a
 (302) ✓ Den lå under ‘a.
 it lay under her_{cl}
 It lay under her. Hellan and Platzack 1999 p. 128 ex. 15b

Finally in Gbadi (303) the pronominal clitic *É* has moved out of its

P(ostpositional)P across the question marker *ḡ* and has merged with the perfective auxiliary (Koopman 1984 p. 54 ex. 27d).

- (303) ✓ wa y- É – ḡ kIÚ jIIE
 they PERF-A- it_{cl}- Q [PP t_{cl} on] put-Q
 Have they put food on it?

English, Swedish, Icelandic, and Norwegian violate (295) on the surface while Gbadi does so at an earlier stage of the derivation, before the clitic moves to the auxiliary. These languages have one important and typologically rare property in common: they allow P-stranding - for the less well-known Gbadi see Koopman (1984). In fact, they represent all of the best known P-stranding languages.¹²⁶

The languages in (296), which were used to exemplify filter (295), all disallow P-stranding. On the basis of this, I suggest that filter (295) holds in all and only the languages that disallow P-stranding. This is generalization (297), repeated here for reference.

- (297) A language allows clitic pronouns as the complements of P iff that language allows P-stranding (and has clitic pronouns, of course).¹²⁷

¹²⁶ Based on the discussion in fn. 101 of this chapter, we might want to add Papiamentu as well. The clitic pronoun *é* can be the complement of a preposition.

- (i) ✓ *e homber ta di Korsow; mi ta hunga kune*
 the man be from Curaçao I ASP play with him_{cl}
 ‘The man is from Curaçao; I play with him.’

¹²⁷ Danish, Faroese, and Vata allow P-stranding but don’t have clitic pronouns. They are therefore unproblematic. Prince Edward Island French allows P-stranding King and Roberge 1990, but it does not readily allow clitics as the complements of P (Yves

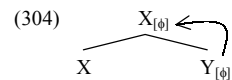
Generalization (297) was first observed in Abels in press. The fact that there is variation with respect to (295) and the correlation with P-stranding suggests that it is not clitic-hood per-se that prevents clitics from occurring as the complements of P in the general case. Generalization (297) suggests that the relevant difference between languages should be sought in the syntax of PPs. There is, in fact, an obvious connection between the syntax of pronominal clitics and P-stranding: clitics are commonly assumed to obligatorily undergo movements that strong pronouns and full DPs are exempt from.¹²⁸ Suppose that this is true. Suppose furthermore that clitics are of the right category to be complements of prepositions. Thus, if prepositions take DPs as their complements, clitics are DPs (Cardinaletti and Starke 1999) or alternatively ambiguously both D° and DP (see Chomsky 1995a; Bošković 2001b). Clitics, being the immediate complements of P°, have nowhere to move within the c-command domain of P°; thus any movement of the clitic for whatever reason must necessarily target a position above P°. Such movement will either (i) pied-pipe PP or (ii) strand the preposition (i.e. leave a trace in the complement position). There is no other option.

Roberge, p.c.). Prince Edward Island French might thus turn out to be problematic for the implication from P-stranding to allowing clitics as the complements of P.

The other half of the generalization is threatened by Hebrew and Arabic. They do not allow P-stranding but seem to allow clitics as the complements of P. These languages are not counterexamples given the fact that Länzliger and Shlonsky 1997; Shlonsky 1997 have argued that these apparent clitics are really agreement heads. The same is true for Celtic languages (see McCloskey 1986a; McCloskey and Hale 1984; Rouveret 1991; Willis 2000).

¹²⁸ For some ideas why this might be so cf. Cardinaletti and Starke 1999, Franks 1998, Bošković 2001b, Stjepanović 1999.

Consider (i) first. Although little is known about the nature of pied-piping, it is not unreasonable to assume that (i) is universally ruled out. Following Webelhuth (1992), I assume that pied-piping presupposes percolation of a feature ϕ relevant for movement from the complement to the phrase, indicated by the curved arrow in (304). Percolation of ϕ is possible under the necessary but not sufficient condition that the features of X do not clash with ϕ .¹²⁹ For *wh*-movement, percolation up to PP (and hence pied-piping of PP) is possible, because the percolated feature is a *wh*-feature. In line with the suggestion made in chapter 2, I will assume that there is an underspecified *wh*-feature present on P, but that it does not percolate. Therefore the mother of P° and its complement is not intrinsically specified for a *wh*-feature. This allows percolation of the *wh*-feature from the complement of P° to the mother node without clash. Ultimately this allows pied-piping of PP under *wh*-movement. The movement of a pronominal clitic is, presumably, linked to the categorial (D) features of the clitic, but the PP already bears categorial features, making percolation impossible. Option (i) is then ruled out universally on plausible grounds.



This leaves us with option (ii), P-stranding. It is allowed in P-stranding languages but not otherwise. Generalization (297) follows naturally. It is in fact surprising that (297) has gone unnoticed so far.

¹²⁹ The same conditions presumably apply to percolation from a specifier to the

One reason why (297) has gone unnoticed might be the fact that pronominal clitics have idiosyncratic properties in many languages. Thus, SC clitics are infamous for their second position requirement.¹³⁰ In fact, we haven't ruled out the possibility that filter (295) holds in SC because of the second position requirement.¹³¹

Suppose we were to unify filter (295) in SC with the second position requirement. The question would remain why (295) holds in so many other languages. Following the SC example, we might blame special properties of the clitics in each of these languages. Given the diversity of clitic systems, the explanation would be extremely non-uniform cross-linguistically, i.e. filter (295) would turn out to not to reflect a real common property of the languages that obey it; the fact that (297) holds would be purely coincidental—a dubious result.

Consider again the Serbo Croatian example given above and repeated here as (305).

- (305) a. *Prema *joj* trče.
 toward her_{cl} run
 b. ✓Prema *njoj* trče.
 toward her_{prn} run
 'They run towards her.'

phrase, too.

¹³⁰ For descriptions and references see Franks and King 2000 and Bošković 2001b.

¹³¹ I run through the case of Serbo Croatian with some care here. It is not difficult to construct similar controls in the other languages mentioned in (296) as well. Given the differences between the clitic systems this would lead to a lengthy and tedious discussion, though.

It turns out that existing theories of the second position effect in Serbo Croatian do not predict that filter (295) should hold without recourse to special properties of prepositions. I briefly discuss two approaches to the second position requirement to show this: (i) second position clitics are located at surface structure in a designated syntactic (Wackernagel) position, (ii) the second position requirement is a prosodic filter (for more details cf. Abels in press).

Suppose the first theory of the second position effect: clitics are in the Wackernagel position at SpellOut. Suppose for concreteness that the Wackernagel position is C° , though nothing hinges on this choice.¹³² On these assumptions, (305a) is expected to be ungrammatical given structure (306). In (306) the clitic is not in C° at spell out.

(306) [_{CP} [_{PP} prema *joj*] [C° [trče *t_{PP}*]]]

However, there is an alternative structure, (307), in which the clitic is in fact located in C° . In (307) the clitic has first moved out of the PP, which is later remnant-moved to the beginning of the clause. PPs can in principle host clitics as shown in (308) so that there is nothing wrong with the assumption that PP might be hosting *joj* in (307).

(307) [_{CP} [_{PP} prema *t_{joj}*] [C° [*joj*+ C°] [trče *t_{PP}*]]]

(308) ✓Prema Mileni *su* trčali.
Toward Milena_{dat} are-cl ran
'Toward Milena, they ran.'

¹³² Theories along these lines are Franks 1998, King 1996 Progovac 1996, Schütze 1994, Tomić 1996, Wilder and Čavar 1994 among many others. For critique see e.g. Bošković 1995; 2001b, Franks 1998, Radanović-Kocić 1988, Stjepanović 1998.

Since movement to C° characterizes clitics by assumption, it is unclear how to rule out (307) without appeal to the ban against P-stranding. Filter (295) thus follows from the ban against P-stranding, not directly from the second position effect.

On the second approach, the second position effect arises as the result of a prosodic filter that rules out sentences with clitics that are not in a prosodically defined second position.¹³³

Consider first example (309). In the example we have an element, *ispod* – 'below', which can act both as a preposition and as an adverb. There is no reason to believe that prosodically the preposition *ispod* behaves differently from the adverb *ispod*. In its adverbial use, *ispod* can host clitics as example (310) shows.¹³⁴

(309) ✓Gledao sam ispod
looked was underneath
'I was looking around underneath.'

(310) ✓Ispod mi je dostojanstva, da idem
below me_{dat.cl} is_{cl} dignity that go_{1st.sg}
'It is below my dignity to go.'

Given examples like these, there appears to be no prosodic reason why clitics should not be hosted by (at least polysyllabic) prepositions. The prosodic account of clitic placement then has nothing to say about examples like (305a). In fact, even elements that are prosodically much lighter than *ispod* may host clitics in Serbo Croatian. Consider example (311). Without the complementizer *da* the sentence is ungrammatical, since the

¹³³ For different versions of a prosodic approach to the second position effect, cf. Franks 1998; Bošković 2001b Embick and Noyer 2001 and references cited there.

¹³⁴ This type of example was first pointed out to me by Wayles Browne (p.c.).

clitics appear in the illicit first position. However, the very light monosyllabic *da* is able to host clitics. It is unlikely how a distinction could be made in purely prosodic terms between the complementizer *da*, which can host clitics, the adverbial *ispod*, which can host clitics, and the preposition *prema*, which cannot host clitics.

- (311) ✓ Ona tvrđi da smo mu je mi predstavili juče.
 She claims that are_{cl} him_{cl, dat} her_{cl, acc} we introduced yesterday
 'She claims that we introduced her to him yesterday.' (Bošković 2001b:8)

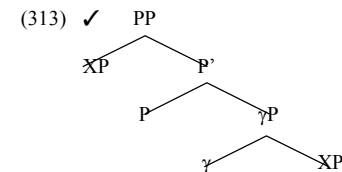
Notice further more, that the example (305a) does not improve by introducing a host for the clitic in sentence initial position. For example, the complementizer *da* – 'that' can host second position clitics (311); however, embedding *prema* under *da* (312) does not improve if in an embedded context the preposition is preceded by the complementizer *da* as in (312) (see Bošković 2001b:161 for related discussion). A purely prosodic account of second position clitics does not seem to be able to deal with these facts. The conclusion must be that other properties of the preposition must be relevant, namely the fact that prepositions cannot be stranded in Serbo Croatian.

- (312) *... da prema joj trče.
 that toward her_{cl} run

Neither the syntactic nor the purely prosodic account of the licensing of second position clitics by itself explains why (295) should hold without recourse to the ban against P-stranding. This situation is typical. To the best of my knowledge, no existing account of the second position effect predicts that filter (295) holds without appeal either to the ban against P-stranding or a specific stipulation amounting to (295).

I have argued that filter (295) should be derived from the ban against P-stranding. Deriving (295) from the ban against P-stranding, predicts that generalization (297) should hold. (297) was shown to hold across an interesting range of languages.

Remember from the introduction to this chapter that I take it that there is a parameter according to which prepositions either are or are not phase heads. The attentive reader will undoubtedly have noticed that an alternative account of clitic movement out of PP and in fact of P-stranding in general is possible without parametrization for the phase-nature of P°. Rather, P-stranding languages could systematically separate P° from its apparent complement by an abstract head γ as shown in (313). Every extraction from PP could then target XP leaving γ P in place. This would entail that every extraction from PP would actually conform to the stranding generalization (193) since every instance of extraction out of PP would be a case of licit subextraction (193a). Illicit extraction of the complement of P (193b) would then never be allowed. This solution was alluded to in fn. (91). On this account, the Stranding Generalization (193) would hold universally for PPs.



Appealing though it may seem, this account is problematic on several counts. We have good reasons to believe that the [Spec, PP] position is an A'-position. After all we use it for A'-extraction in the cases of extraction that were discussed in sections 4.1 and 4.3.4. But if [Spec, PP] is an A'-position, then it would remain mysterious how A-

movement out of PP is possible in some P-stranding languages (pseudopassive). This problem may be overcome, though. Other positions have been claimed to have ambiguous A- and A'-properties <refs??? Spec IP, Spec DP...>.

Additionally, there is no independent evidence for the existence of γ . The following facts show that we would expect to find morphological evidence for γ if γ existed. There are rare cases, where non-P-stranding languages allow pronominal clitics as the complements of P°. Examples of this kind include the Old Church Slavonic clitic *i* (**jь*) (3rd.sg.msc.acc) and the Hausa clitic *mù* (2nd.sg.obj).¹³⁵ On the account of filter (295) given here this ought to be impossible, unless some additional structure along the lines of γ in (313) were to intervene between the preposition and its apparent complement. Indeed, grammars of Old Church Slavonic state that “[a]fter prepositions an *n* is prefixed to the stem [of the clitic]: [...] *na njь* (acc. sg. masc.) etc” Lunt (1959:53). Similar comments apply to Hausa. Tuller (1986:281) contains a single example where a Hausa preposition takes a clitic/weak pronoun as its complement. Just in this case an extra morpheme is inserted; *ga* –‘to/by’ surfaces as *garee* when followed by the clitic/weak pronominal *mù* –‘you’.¹³⁶ I assume that *n* in Old Church Slavonic and *ree* in Hausa identify γ . Given that γ seems to be morphologically identified when present in

¹³⁵ Apparently Slovak still makes use of the Old Church Slavonic pattern (Michal Starke, p.c.)

¹³⁶ Notice that the insertion of a morpheme here does not necessarily indicate that the clitics have to move out of their containing PPs in these languages. It is conceivable that the extra morphemes identify structure that serves as a PP-internal landing site for clitic movement.

these languages, it would be surprising not to find any evidence for its presence in P-stranding languages – see Sugisaki 2002 for a suggestion, though.

To further test the hypothesis whether there is abstract structure present within PP in stranding but not in non-stranding languages, Željko Bošković (p.c.) suggests the following line of reasoning: Quantifiers are floated under movement, but they are never left in θ -positions (Bošković 2001a; 2002a). If quantifiers could be floated PP internally, this would count as evidence that there is PP internal movement. The converse is not true, however. The absence of PP internal Q-float does not establish the fact that there is no PP internal movement. Bošković’s theory of Q-float may illustrate this. For Q-float to be possible, at least two movement operations need to take place, one from a θ -position to a non- θ -position, which can never float Q, and then another movement operation that moves DP and floats Q in the non- θ -position. For P-stranding, it is enough if there is one PP internal position that DP can move to.

Despite the clear logic of the situation, evidence that Q-float is in fact possible within PP is hard to come by. For the moment I leave the issue unresolved.

In this subsection we have seen that the present approach to P-stranding together with the assumption that clitics undergo movement processes that other DPs are exempt from derives the fact that in non-P-stranding languages clitic pronouns are illicit as the complement of P°.

4.5 On the nature of P-stranding

I will now discuss the ramifications of the findings of this chapter and of previous work on P-stranding for a theory of P-stranding. At the end of the section I briefly discuss how various approaches to P-stranding capture the relevant generalizations.

4.5.1 The P-stranding generalizations

There is a broad range of claims in the literature that bear on the nature of P-stranding. I have discussed two in some detail in this chapter: Generalization [1] and Generalization [2].

Generalization [1]: Languages that do not allow P-stranding do not allow clitic pronouns as the complement of P

Generalization [2]: Even in non-P-stranding languages PPs are not islands.

From the literature on P-stranding, Generalization [3] emerges.

Generalization [3]: All languages that allow P-stranding under A-movement, i.e. pseudopassivization, also allow P-stranding under A'-movement.

The languages that allow both types of stranding include English, Norwegian (Merchant 1999; Vikner 1995), Swedish (Maling and Zaenen 1985a; Merchant 1999; Takami 1992; Vikner 1995), Vata with postpositions (Koopman 1984), Gbadi with postpositions (Koopman 1984), and Prince Edward Island French (King and Roberge 1990; Roberge 1998; Roberge and Rosen 1999). Stranding only under A'-movement is allowed in Icelandic (Maling and Zaenen 1985a, b, 1990), Frisian (section 4.2 of this

chapter), Danish (Herslund 1984; Merchant 1999; Takami 1988, 1992). There is no language that allows P-stranding under A-movement but not under A'-movement.¹³⁷

Generalization [4]: A language allows P-stranding under sluicing iff it allows P-stranding under question formation.

This generalization was discovered by Merchant (1999:113-117). Merchant shows on the basis of a large number of languages (German, Dutch, Greek, Russian, Polish, Bulgarian, Serbo-Croatian, Slovene, Persian, Catalan, French, Spanish, Italian, Hebrew, Moroccan Arabic, and Basque) that languages that do not allow P-stranding under question formation, do not allow P-stranding under sluicing. I repeat Merchant's examples from Serbo-Croatian here.

- (314) a. *Kim je Ana govorila sa?
 *who is Ana spoke with?
 b. Ana je govorila sa nekim, ali ne znam *(sa) kim.
 Ana is spoken with someone, but not know.1sg with whom
 ‘Ana spoke with someone, but I don’t know who?’

Languages that do allow P-stranding under question formation also allow it under sluicing. Merchant gives examples from English, Swedish, Norwegian, Danish, and Icelandic.¹³⁸

Generalization [5]: All languages that allow preposition stranding also have verbal particles (Stowell 1982a and Sugisaki 2002; Sugisaki, Lasnik and Snyder 2001; Sugisaki and Snyder 2001 for evidence from language acquisition).

¹³⁷ There do not appear to be any pseudo-middles in any P-stranding language. I do not have an explanation for that fact.

(i) *This bed sleeps on well/sleeps well on.

¹³⁸ Sluicing with P-stranding is possible in the Creoles discussed in Alexandre and Hagemeijer 2002as Tjerk Hagemeijer (p.c.) tells me.

See Áfarli 1985; Dikken 1992; Herslund 1984; Holmberg 1986 Neeleman 1994, 2001; Svenonius 2002 among many others on Germanic, Koopman 1984 on Vata and Gbadi, and King and Roberge 1990 on Prince Edward Island French. I illustrate from Prince Edward Island French (PEIF) since PEIF departs significantly from standard French in this regard (all examples from King and Roberge 1990:366-367).

- (315) a. ✓Il a pluggé {in le computer |le computer in}.
 he has plugged in the computer the computer in.
 b. ✓Pluggé {le |le computer} in!
 plug it the computer in
 c. ✓Ils avont layé {off le monde |le monde off} à la factorie
 They have layed off the people thepeople off at the factory
 d. ✓Il y a une tapée de femmes qui travaillent out
 it there has a lot of women who work out
 'There are a lot of women who work outside the home.'

It is unclear to me why this generalization should hold.¹³⁹

¹³⁹ The generalization must be taken with a bit of care; particles are not all created equal. Svenonius 2002 argues that particle constructions in Swedish, Danish, Norwegian, and English (all of them P-stranding languages) and separable prefixes in Dutch, German, Yiddish and Afrikaans (all of them, except for maybe Afrikaans non-P-stranding languages – see den Besten 1981 for discussion of comparatives in Afrikaans which suggests that Afrikaans does not allow P-stranding) are substantially different. If there is a true correlation here, this is interesting, but more work needs to be done. The most intriguing property of Svenonius' findings is that the structural possibilities that Svenonius assumes for the non-P-stranding languages is a proper superset of the structural possibilities assumed for the P-stranding languages. The structure that is not independently available in the P-stranding languages is, moreover, structurally a subset of the available (and obligatory) structure. Now, if the ability to strand P° were to ride on the presence of the bigger structure, then the non-P-stranding languages should all be P-stranding languages since they have the larger structure along with the smaller one. It is extremely difficult to see how Svenonius' assumptions could be reconciled with my take on P-stranding or in fact any approach that links availability of structure with stranding.

4.5.2 The Current Approach

As mentioned at the beginning of this chapter, I would like to suggest that the property whether a language allows P-stranding or not resides in the inherent properties of P° in that language. If P° is a phase head in a given language, P-stranding is disallowed. This is the common situation. If P° is not a phase head in a given language, P-stranding is allowed. This is the situation found in stranding languages.

As we saw throughout this chapter this derives the Generalization [2]: Extraction of the complement of P° is strongly disallowed in non-P-stranding languages, extraction out of the complement of P° is, in principle, possible.¹⁴⁰ In stranding languages both extraction of the complement of P° and extraction out of it are possible. In the previous section I discussed how Generalization [1] follows on the present account.

- (316) Parameter 1: [+/-] P° is a phase head.

There is a second independent parameter regulating whether the Case assigning property of P° can be suppressed. For the moment it is sufficient to think of this parameter as regulating whether prepositions optionally or obligatorily assign Case. In the end this might not be sufficiently precise, however.

- (317) Parameter 2: [+/-] P°'s Case may be suppressed.

Together these parameters give rise to a four way typology of languages. P-stranding will be allowed if P° is not a phase head. P-stranding will be disallowed if P° is

¹⁴⁰ The qualification 'in principle' is quite important of course. Not all extractions out of PP are possible and in fact most aren't. But the situation is the same for all

a phase head. If P° is a phase head, pseudopassives are ruled out even if P° 's Case may be suppressed. The DP complement of P° will not be able to leave PP to be case marked and such structures will be ruled out by the Case filter or some analogue thereof.

Pseudopassives will therefore only be allowed if P° is not a phase head and P° 's Case may be suppressed. This correctly captures the implicational relation from the availability of pseudopassives to the availability of P-stranding under A'-movement. Generalization [3] therefore follows.¹⁴¹

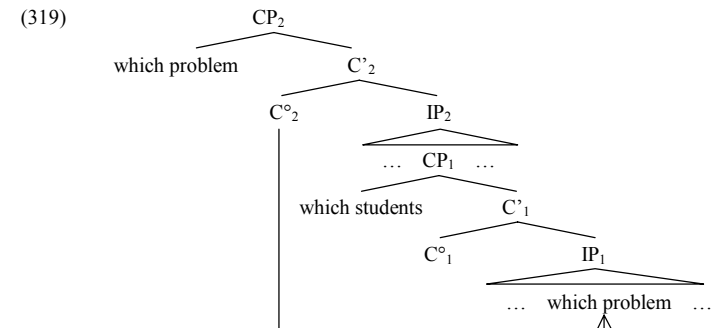
Generalization [4] is a bit more involved. It says that P-stranding does not become licit under sluicing in non-P-stranding languages. Sluicing is known to ameliorate island effects (Ross 1969; Chung, Ladusaw and McCloskey 1995; Fox and Lasnik 2003; Lasnik 2001; Merchant 1999, to appear). Consider example (318) from (Chung et al. 1995 ex. 84a). In (318a) sluicing has applied and the example is acceptable despite its complexity, example (318b) is just like (318a) except that sluicing has failed to apply and the example is degraded.

domains that allow extraction. Delimiting the conditions under which extraction is possible is not the topic of this thesis.

¹⁴¹ The system as it stands appears to overgenerate in a non-trivial way. If P° simply may or may not come with a Case assigning property and verbs may or may not assign an external theta role, then four possibilities are in principle possible: (i) P° assigns Case and V° has an external argument (e.g. ✓Peter talked about Frank.); (ii) P° does not assign Case and V° has an external argument (this case is ruled out by the Case filter); (iii) P° does not assign Case and V° does not have an external argument (e.g. Frank was talked about.); and (iv) P° assigns Case and V° does not take an external argument (e.g. *It was talked about Peter.). (iv) is predicted to be possible. The over-generation problem is only apparent, however, once we notice that the status of (iv) can plausibly be attributed to the independent ill-formedness of other impersonal passives in English (*It was danced.).

- (318) a. ✓Sandy was trying to work out which students would be able to solve a certain problem, but she didn't tell us which one.
 b. ??Sandy was trying to work out which students would be able to solve a problem, but she didn't tell us which problem she was trying to work out which students would be able to solve.

If, as suggested here, the ban against P-stranding reduces to the joint operation of two conditions, Attract Closest and Last Resort, we will have to conclude that at the very least Attract Closest may be obviated under ellipsis. To see this consider the structure in (319).

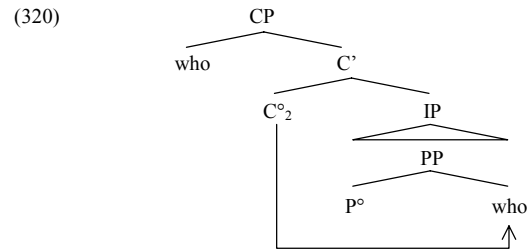


C°_2 attracts *which problem* across the interveners C°_1 and *which students*.¹⁴² This is in fact the kind of derivation argued for in Fox and Lasnik 2003, where movement under sluicing does not land in intermediate landing sites. Notice that if we assumed successive cyclic movement to take place in the sluice (with or without islands) we would

¹⁴² The situation does not really improve from the point of view of attract closest if we assume that *which problem* moves successive cyclically, because attraction of *which problem* by C°_1 would still violate attract closest since *which students* (or its trace, depending on the order of operation) intervenes. I leave it as an exercise to the reader to check that this is the predicted result given the definitions developed in chapter 2.

have to posit a structure for the sluice which is non-identical to its antecedent. The sluice would be non-identical not only because of the additional intermediate traces but also because the features on the intermediate landing sites would receive different values in the antecedent and in the sluice. Suppose that this is impossible, i.e. suppose that feature identity for the phase heads within the sluice is part of the identity condition on sluicing. Clearly this cannot be the only identity condition there is. What I am claiming is that it should be part of the ultimate formulation of the condition on recoverability of deletion or follow from that ultimate formulation.

Now consider what would happen if sluicing were applied under P-stranding in a non-P-stranding language (320) – recall the relevant example (314b) above, which is ungrammatical without pied-piping.



Again attraction of *who* violates Attract Closest. There is no distinction there. But notice that *who* is in a local feature satisfaction relation with P° whereas *which problem* in (319) was not in a feature satisfaction relation with any of the c-commanding phase heads. As a result, the phase heads in (319) could be identical after feature valuation to the corresponding heads in the antecedent, but in (320) this is not possible. Due to the

presence of the *wh*-word (i.e. a *wh*-feature bearer in the terminology of chapter 2) P° receives a feature specification different from that of the preposition in the antecedent clause, where P° was not in a feature satisfaction relation with a *wh*-feature bearer. The *wh*-feature on P° ends up with a different value in the sluice from the value in the antecedent and the structure is ruled out by the condition that the features of phase heads must be identical in the sluice and in its antecedent.¹⁴³ The reasoning so far derives why P-stranding does not become licit in a non-stranding language, despite the fact that violations of Attract Closest are apparently tolerated. The answer is that P-stranding would violate the assumed identity condition on ellipsis.

In a P-stranding language, the situation is quite different. Since P° in such a language is not a phase head, it does not come with a full set of mover-features. In particular P° does not come with a *wh*-feature.¹⁴⁴ Therefore, the identity condition will not be violated and the structures with P-stranding are predicted to be grammatical in a P-stranding language.

¹⁴³ I must leave for further research the question of how the assumed identity condition interacts with and is compatible with vehicle change in ellipsis (Fiengo and May 1994; Merchant 1999).

¹⁴⁴ It may be that in English P° is a phase head optionally. If we assume this, PP should behave like PPs in a non-stranding language when the option is chosen where P° is a phase head. In other words, if the option to treat P° as a phase head is chosen, pied-piping of P° is forced just like it is in a non-stranding language. This optionality might account for the optionality of pied-piping in English. Notice that Postal's 1972 problem of stranding P° in an intermediate landing site of movement does not arise under this view.

If indeed P° is optionally a phase head in English, then we would not say that P° does not come with a *wh*-feature, but that it does not come with one necessarily. The possibility of not having a *wh*-feature on P° is sufficient to ensure the correct result.

Let's call this account of Generalization [4] the Identity Solution.

It might be objected at this point that the identity condition on phase heads is a stipulation that, at present, doesn't do any work but to derive Generalization [4]. Had the facts come out some other way, say that non-P-stranding languages allow stranding under sluicing, we could have simply assumed that feature identity on phase heads is not relevant to ellipsis but that it is subject to vehicle change. The point is well taken. The question is though what an alternative to the Identity Solution might look like.

Recall that in chapter 2 I made two logically independent moves. I derived the ban against moving from complement to specifier position of the same head from Last Resort and I made steps towards the unification of Attract Closest and the Head Constraint. We could reject the second part and fall back on a position that has two distinct locality principles: Attract Closest and the Head Constraint. We could then capture Generalization [4] by assuming that the Head Constraint always holds but that Attract Closest can be obviated under sluicing. The account would run along the following lines. The DP complement of a P^o can never reach the escape hatch position in [Spec, PP] by Last Resort. Extraction of the complement out of PP in a non-stranding language therefore always violates both Attract Closest (because of the intervening P^o head – see (320)) and the Head Constraint. Attract Closest can be obviated, but the Head Constraint by assumption cannot. Therefore, P-stranding in a non-P-stranding language is disallowed under sluicing. This view entails that other movement operations within the sluice must be successive cyclic though possibly violating Attract Closest. If they were not successive cyclic, they, too, would fall prey to the Head Constraint.

Call this the Locality Solution.

Although the Locality Solution would account for Generalization [4], I do not think that it is any more desirable than the Identity Solution advocated above where Generalization [4] is attributed to the identity condition on phase heads. The number of assumptions needed to make the account in terms of an independently stipulated Head Constraint work is at least as great as the number of assumptions needed under my account – and the stipulations under the Locality Solution are less desirable as I will show now.

Recall the assumptions that went into the Locality Solution. We assumed that the Head Constraint and Attract Closest are different conditions. We assumed that Attract Closest may be obviated under ellipsis but that the Head Constraint may not. This was the crucial assumption to enforce pied-piping even under sluicing. This assumption would be strictly driven by the facts. It is no more natural than to assume that both can be obviated, that none can be obviated, or that only the Head Constraint but not Attract Closest can be obviated under ellipsis. In any case, the Head Constraint holds by assumption. Finally, we need to make an assumption regarding the identity condition on sluicing. Because the Head Constraint is assumed to hold (otherwise pied-piping is not enforced), the sluice must contain intermediate traces. These intermediate traces are not present in the antecedent because of the head constraint, we have to assume that intermediate traces within the sluice do not count for identity.

The identity condition used for the Identity Solution makes the theory more restrictive; it allows less cases of ellipsis. The identity condition for the Locality Solution

does the opposite; it makes the theory less restrictive. This is certainly desirable from a theoretical point of view.¹⁴⁵

To conclude this discussion, the most obvious alternative account of Generalization [4] is inferior to the one suggested here. Not only does it give up the possibility of having a unified notion of locality and a non-stipulative approach to escape hatches, it also loosens the identity condition on ellipsis in an undesirable way.

Generalization [5] is not accounted for under the present theory. Partly, this is owed to the fact that it is unclear what Generalization [5] actually amounts to (see fn. 139).

I will therefore not attempt to link this generalization in one way or another to the parameters that I assume. In any case, a further parametric choice will have to be introduced since, at least on the most common understanding of Generalization [5] some non-P-stranding languages do and some don't have verbal particles. The situation, as mentioned, is not clear however, since Svenonius 2002 points out a number of significant differences between the verb particle constructions in what seem to be the Germanic P-stranding languages on the one hand and the non-stranding languages on the other.

4.5.3 Previous Approaches

There are three main approaches to the syntax of P-stranding that I would like to discuss briefly in the remainder of this chapter. First I will discuss escape hatch based

¹⁴⁵ Fox and Lasnik 2003 argue that on empirical grounds intermediate traces within an ellipsis site do count for identity and that an antecedent without such traces is not identical to an ellipsis site with them.

approaches (Koopman 1997; Riemsdijk 1978a, 1990). My own approach has a great conceptual affinity to the escape hatch based theories, except that I would argue that they get things exactly backwards. Then I will turn, in chronological order, to approaches based on literal reanalysis (Hornstein and Weinberg 1981; Stowell 1982a). The literature on this approach is vast and mostly negative. I will give a small sampling of the main arguments. Then I turn to approaches that rely on government (this tradition was inspired by Kayne 1981; 1984 chapter 5).^{146, 147, 148} I will not return in this discussion to Generalization [1] since I believe that every theory of (non-)P-stranding can capture it.

¹⁴⁶ For comments on the approach in terms of underspecification/category neutralization proposed in Pollock 1989a; Rizzi 1990 see Newmeyer 1998.

¹⁴⁷ Law 1998; Salles 1997 suggest to account for the impossibility of P-stranding in German, Dutch and the Romance languages on the basis of a distinction between the Germanic languages that allow P-stranding on the one hand and German, Dutch, and the Romance languages on the other hand. In Romance and in German and Dutch determiners can amalgamate with prepositions (French: *de le* → *du*; *de les* → *des*; Portuguese: *em o* → *no*; *por o* → *pelo*; German: *von dem* → *vom*; *in dem* → *im*; Dutch: *in het* → *in 't*; *voor het* → *voor 't* etc.) while no such amalgamation is observed in English and other Germanic P-stranding languages. The authors suggest that the presence of [P+article] contraction blocks P-stranding and that languages that allow P-stranding are marked by the absence of [P+article] contraction. It is unclear what this generalization if true would follow from.

In any case, the generalization is counterexemplified by Prince Edward Island French (King and Roberge 1990). Prince Edward Island French appears to show the regular French pattern of [P+article] contraction while allowing P-stranding. The examples in (i) are taken from (King and Roberge's (1990 p. 356 ex (7c) and p. 362 ex. (22b))) discussion of P-stranding in Prince Edward Island French.

(i) a. ✓Robert a été parlé beaucoup de t_{Robert} au meeting.
Robert has been talked a-lot of at+the meeting
'Robert was talked about a lot at the meeting.'
b. ✓Ça fait trois locataires qu' on a du trouble avec
it makes three tenants that we have of+the trouble with
'We have had three tenants whom we had trouble with.'

van Riemsdijk's (1978a) analysis of P-stranding is, as mentioned, the closest to the analysis pursued in this chapter. However, van Riemsdijk uses the same conceptual ingredients to very different effect. He suggests that PPs are universally bounding nodes and that there is a parameter according to which PPs either have or do not have a COMP

If we are to conclude that Papiamentu allows P-stranding, then Papiamentu also counterexamples the generalization. The preposition *di* – 'of' and the definite article *é*, which is homophonous with the weak third person singular pronoun and the mysterious *É* discussed above, can be contracted to *dje* as shown in example (ii). (The example is from the web for further examples see Birmingham 1970:41, 108)

- (ii) e porta dje Jeep a habri
 the door of-the Jeep ASP open
 'The door of the Jeep opened.'

¹⁴⁸ Sugisaki 2002 suggests that the availability of P-stranding might be linked to the availability of additional structure in P-stranding languages (concretely a pP layer above PP), which is absent in non-P-stranding languages. The diagnostic he gives for the presence of this pP layer is the well known difference between French and English illustrated in (i) and (ii). Sugisaki suggests that the additional interpretation that the PP has in English is indicative of the presence of an additional layer of structure. To accommodate languages like German or Russian where prepositions behave semantically like English prepositions, Sugisaki suggests a parameter concerning the movement of the lexical P to the upper layer of structure. Movement of P, on Sugisaki's analysis, allows P-stranding. This predicts an entailment relation. If a language allows only the location interpretation in an example like (i) it could not possibly allow P-stranding. The obvious test case would be Prince Edward Island French. I do not know at present what the facts are. It is also not clear to me how to account in Sugisaki's system for Generalization [3] and Generalization [4] above. For Sugisaki, DPs never move out of the traditional PP. In P-stranding-languages the lower PP shell in a layered [_{PP} ... [_{PP} ...]] moves. In pseudopassives, there would then be PP-subjects, which, moreover, induce agreement on the verb. (See Bresnan and Kanerva 1992; Conway 1996 for arguments that PPs never occupy subject position in English.) Sugisaki himself notes that Generalization [5] remains problematic, but that seems to be a common problem of all but the reanalysis theories of P-stranding.

- (i) The bottle floated under the bridge
 ✓the bridge=location, ✓ the bridge=goal

position that can serve as an escape hatch. Updating the terminology (see Bošković 2001a; 2002a), this says that there is a parameter whether PPs have a specifier position that can serve as an escape hatch or not. Or, in Chomsky's (2000; 2001a; 2001b) most recent terminology: there is a parameter whether the phase head P° can be optionally assigned a (EP)P feature. Languages that allow projecting the escape hatch position allow preposition stranding according to van Riemsdijk, all others don't.¹⁴⁹ This is in its bare bones van Riemsdijk's theory of P-stranding under A'-movement. The obvious problem is that van Riemsdijk's theory makes PPs in non-P-stranding languages islands. Since there is no [Spec, PP] position in non-P-stranding languages, nothing should ever be able to escape from PP. In sections 4.1 and 4.3.4 of this chapter we saw that this prediction is not borne out. In other words, van Riemsdijk does not capture Generalization [2]. Notice though that van Riemsdijk's system allows subextraction out of PP in a P-stranding language.

- (321) Which building did you bungee jump off [the roof [of t]]?

Koopman's (1997) system also relies on various specifiers as escape hatches. She derives the ban against P-stranding from the ban against Left Branch Extractions (Ross 1967). Koopman assumes that in non-P-stranding languages the lexical P and its complement must move to the specifier of a higher functional head Place°. In Dutch (and

-
- (ii) La bouteille a flotté sous le pont.
 The bottle has floated under the bridge
 'The bottle floated under the bridge'
 ✓the bridge=location; *the bridge=goal

presumably German) [Spec, PlaceP] can either be occupied by the lexical PP or by an R-word (322a) – other items cannot satisfy the checking requirements of Place°. R-words can use [Spec, PlaceP] as an escape hatch, but the complement of the lexical P° cannot. It is frozen in place because extraction of DP from PP would violate the Left Branch Condition (Koopman 1997:17) – see (322a).¹⁵⁰ The Left Branch Condition does not distinguish the complement of the lexical preposition from its proper subparts; Generalization [2] is therefore not accounted for.

- (322) a. Dutch: [PlaceP {[PP P° DP] [R-word] [Place° [AgrP Agr° t_[P° DP]]]}]
 b. English: [PlaceP [PP P° t_{DP}] [Place° [AgrP DP [Agr° t_[P° DP]]]]]

P-stranding in a language like English is allowed because, Koopman assumes, the DP complement of the lexical preposition moves out of the lexical PP to [Spec, AgrP] followed by movement of the remnant PP to [Spec, PlaceP] – see (322b). This allows DP to move further without violating the Left Branch Condition. Presence or absence of movement of the DP complement of P° to [Spec, AgrP] therefore constitutes the stranding parameter for Koopman. Notice that extraction out of DP as in the English example (321) above would appear also to constitute a violation of the Left Branch Condition and is therefore incorrectly ruled out. Extraction out of the complement of P°

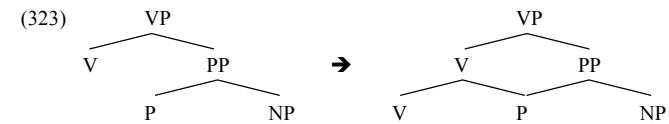
¹⁴⁹ As we saw, van Riemsdijk's position is more complicated since he assumes that in German and Dutch there is an escape hatch which can only be occupied by R-pronouns.

¹⁵⁰ Koopman suggests a split PP structure with an extensive array of functional projections. For German, I have been able to reproduce Koopman's arguments for the extended structure of PP only in contexts where PP is embedded under a verb but not in contexts where PP is embedded under a noun, an adjective or another preposition. I do

should presumably be disallowed both in stranding and in non-stranding languages according to Koopman since it would involve extraction either out of PP in [Spec, PlaceP] or from DP in [Spec, AgrP]. Koopman does not discuss her view of Left Branch Extraction in any detail, but it is difficult to see how the two cases could be distinguished in a principled way.

Van Riemsdijk (1978a) also discusses P-stranding under A-movement. He suggests that prepositions can be reanalyzed under adjacency with verbs as shown in

(323).¹⁵¹



Van Riemsdijk assumes a definition of c-command as in (324). According to this definition P does not c-command NP in the reanalyzed structure in (323). Thus the Head Constraint (40) is voided because P°, the head of $\alpha=PP$ no longer c-commands NP. This allows NP to move out of PP without going through the escape hatch position.

- (324) α c-commands β iff every node dominating α and also dominates β and α does not dominate β .
 (325) *Head Constraint*
 No rule may involve X_i (X_j) and Y in the structure $\dots X_i \dots [\alpha \dots Y \dots] \dots X_j \dots$ if **Y is c-commanded by the head of α** .
 α ranges over V''' , N''' , A''' , P'''

not know what the facts in Dutch are. The German situation suggests to me that many of Koopman's positions within the split PP are really verbal projections.

¹⁵¹ The structure has a non-accidental similarity with Head Movement Structures, see Riemsdijk (1998 for discussion).

The account has the virtue that it captures the fact that there needs to be a close relation between the verb and the preposition to allow pseudopassives and that the relationship between the verb and the prepositions needs to be closer for pseudopassives than for A'-movement. A second virtue is that the [Spec, PP] escape hatch position can be viewed unambiguously as an A'-position, since it is used only for A'-movements.¹⁵² To parameterize the system, i.e. to allow pseudopassives in some languages and disallow it in others, van Riemsdijk assumes that only some languages allow the reanalysis operation and others don't. The parameter allowing P-stranding under A'-movement ([+/- [Spec, PP]]) is thereby completely independent from the parameter allowing P-stranding under A-movement ([+/- Reanalysis]). van Riemsdijk's theory thus fails to predict Generalization [3].¹⁵³ I will not discuss possible extensions of van Riemsdijk's account to Generalization [4] and Generalization [5]. It should be clear at this point that the problems that van Riemsdijk's account faces are serious enough to warrant the search for an alternative.¹⁵⁴

¹⁵² The issue doesn't arise under the present account, since P-stranding languages never need to use the [Spec, PP] escape hatch for either kind of movement.

¹⁵³ Koopman 1997 discusses only Dutch, which doesn't have pseudopassives. The issue therefore doesn't come up in her paper.

¹⁵⁴ Note that Koopman's (1997) approach in terms of the Left Branch Condition does make a prediction concerning P-stranding under sluicing (Generalization [4]). As discussed in Fox and Lasnik 2003; Kennedy and Merchant 2000; Merchant 1999, to appear ellipsis in general, and sluicing in particular, can repair violations of the Left Branch Condition. Koopman would therefore predict that the ban against P-stranding should not hold under sluicing counter to fact.

I now turn to Hornstein and Weinberg's (1981) theory of P-stranding in terms of literal reanalysis. Hornstein and Weinberg's proposal has received an enormous amount of attention and a mostly bad press (Baltin and Postal 1996; Couper-Kuhlen 1979; Cruz and Saamenho 1996; Donaldson 1982; Duarte 1994; Inada 1981; Levine 1984; Maling and Zaenen 1985a, b; Newmeyer 1998; Salles 1997; Siegel 1983; Takami 1988, 1992). There are two important points the paper makes. Hornstein and Weinberg criticize a theory like van Riemsdijk's, Bresnan's (1978) and the present one as insufficient, because the possibility to have P-stranding is related (only) to the internal syntax of the PP. This, they argue, is inadequate because the same PP may allow extraction when it shows up in one context and may not allow extraction when it shows up in another. Hornstein and Weinberg (1981:59 ex. 19) give the following example to illustrate their point. (The judgments are theirs.)

- (326) a. ✓Who did you speak to Harry about yesterday?
 b. *Who did you speak to Harry yesterday about?

Although many of Hornstein and Weinberg's judgments seem overly restrictive to many speakers of English (see Takami 1988, 1992 for many such cases), the general point is an important one: the environment where a PP occurs plays an important role in determining whether stranding is possible or not. For other categories this is now considered a truism. Whether a DP allows extraction or not depends not only on its internal syntax but also on its environment. The same DP may allow extraction in direct object position but disallow it in subject position (Chomsky 1986a; Kayne 1981). It may allow extraction in object position but disallow it when Heavy NP Shifted, etc. Whether a

CP allows extraction or not depends largely on its position. Extraction is allowed mainly from the complement position of bridge verbs but not from adjunct position.

Consider again example (326) and compare it to example (327) discussed in chapter 3 to exemplify the ban against movement out of Heavy NP Shifted objects. If Hornstein and Weinberg's judgments are real, then the situation would actually seem to be roughly parallel. But if that is so, it is a mistake to attribute as Hornstein and Weinberg do the ill-formedness of (326b) to some condition applicable only to P-stranding and (327) to some other condition.¹⁵⁵

- (327) a. ✓Which Post office will Mary send a picture of to inspector Smithers.
b. *Which Post office will Mary send to inspector Smithers a picture of?

It seems reasonable to try to factor such conditions on extraction out of the core theory of P-stranding as I have done and relegate them to a theory of Conditions on Extraction Domains (Huang 1982). From this perspective, Hornstein and Weinberg conflate two issues that should be kept separate.

The actual proposal that Hornstein (1981) make crucially involves three ingredients. First there is the filter (328) (Hornstein and Weinberg 1981:60 ex. 20), which

¹⁵⁵ Incidentally, Hornstein and Weinberg's theory seems to predict that (327) is ungrammatical. Hornstein and Weinberg state (p. 58) that "one can show that preposition stranding is limited to those cases where the relevant preposition [sic!] is immediately dominated by VP [footnote omitted, K.A.]." Taking "preposition" to mean PP, this seems to indicate that immediate domination by VP is a necessary condition for the possibility to strand. The footnote to this sentence starts as follows: "In English, domination by the verb seems to be a sufficient condition for the application of the Reanalysis rule." The Reanalysis rule is what enables P-stranding. It applies in the base. Immediate domination by VP in the base is then a necessary and sufficient condition on the possibility to strand P. This rules out (327b) where the PP headed by *of* is not immediately dominated by VP and thus not subject to reanalysis.

rules out empty NPs marked with oblique Case.¹⁵⁶ Then there is the reanalysis rule (329) (Hornstein and Weinberg 1981:60 ex. 21). Rule (329) is assumed to operate in the base (before movement and Case marking applies) and take a verb and material that is right adjacent and reanalyse it as a new verb.¹⁵⁷

- (328) *_[NP e]
oblique
(329) V → V* (where V c-commands all elements in V*)

Let's consider a concrete example. Hornstein and Weinberg can account for the contrast between examples (326a) and (326b) repeated here. They assume that *yesterday* is base generated either at the left edge or at the right edge of VP.

- (326) a. ✓Who did you speak to Harry about yesterday?
b. *Who did you speak to Harry yesterday about?

¹⁵⁶ The filter is explicitly assumed to apply not only to traces. This point will become relevant only later on.

¹⁵⁷ As we saw in footnote (155) the conditions on the reanalysis rule must be more stringent than stated in (329). The more stringent requirements are needed to account for, among other things, examples like (i). Assuming with Hornstein and Weinberg that *at what time* in (ii) can be base generated in the position before the embedded subject, rule (329) would allow reanalysis of the string *think that at* as shown in (i'). The trace of *what* could then be assigned Case by the matrix verb. It would not be an oblique trace and the sentence is predicted to be grammatical. Thus the more restrictive formulation discussed in fn 155 is needed. The formulation demands that the reanalyzed PP be immediately dominated by VP, which is not the case in (i). On the more restrictive formulation, reanalysis blocks in this case.

- (i) *What time do you think that at John arrived.
(i') What time do you {think that at} _{t_{what}} John arrived.
(ii) ✓You think that at eight o'clock John arrived.

In example (326a) the string *Speak to Harry about* is reanalysed in the base. Then *who* is moved out and finally the trace of *who* is case marked with objective Case by the governing verb *Speak to Harry about*. If reanalysis hadn't applied, the trace of *who* would have been marked oblique by the preposition and a violation of filter (328) would have ensued. Example (326b) cannot be generated. If reanalysis applies in the base, the verb *Speak to Harry about* would be created. To derive the word order in (326b), either *yesterday* would have to move into this verb or *about* would have to move out. Movement into and out of words is assumed to be impossible. This derivation is therefore blocked. On the other hand, if reanalysis does not apply, the correct word order can be generated, but the trace of *who* will be case marked oblique by the preposition in violation of (328).

To account for the significant differences between P-stranding under A- and under A'-movement illustrated in (330), Hornstein and Weinberg (p. 65) assume that due to general conditions the reanalysed word in the case of A-movement must be a 'semantic word'. This condition does not apply to reanalysis in the case of A'-movement.

- (330) a. ✓Who did Sam talk to Harry about?
b. *John was talked to Harry about.

There are great difficulties in clarifying the notion of *semantic word*, but I set this problem aside since the conditions under which pseudopassivization is allowed remain quite elusive under any approach. Consider the famous contrast between (331) and (332), which has no satisfactory explanation as far as I know. (See Takami 1992 for conditions

under which (332) may become acceptable and Ramchand and Svenonius 2002 for an approach to particle constructions that might help with the pseudopassives.)¹⁵⁸

- (331) ✓This bed has been slept in.
(332) *New York has been slept in.

Ignoring the issue of 'semantic words' there are still grave problems for Hornstein and Weinberg's reanalysis proposal as discussed in the literature (Baltin and Postal 1996; Couper-Kuhlen 1979; Cruz and Saameno 1996; Donaldson 1982; Duarte 1994; Inada 1981; Levine 1984; Maling and Zaenen 1985a, b; Newmeyer 1998; Salles 1997; Siegel 1983; Takami 1988, 1992 among others). One class of problem (noted already in Hornstein and Weinberg 1981:74 fn. 21 and attributed there to Henk van Riemsdijk and Edwin Williams) has to do with cases of multiple P-stranding. Consider the examples in (333). To derive (333a) the string *talk to Harry about* must be reanalysed as a verb. To derive (333b), *talk to* must be reanalyzed as a verb.

- (333) a. Who did you talk to Harry about?
b. Harry has been talked to about this issue.
c. Which problems has Harry been talked to _{t_{Harry}} about _{t_{which problems}}?
d. Who would you like to be sung to _{t_{you}} by _{t_{who}}?

This is not a problem per se, of course, but to derive (333c), both reanalyses must be done simultaneously and, what's worse, *Harry* must move out of the reanalysed word *talk to Harry about*. Hornstein and Weinberg's footnote (1981:74 fn. 21) discuss the question whether it is possible to have several competing reanalyses in the same phrase marker. They endorse the conclusion that this is possible but seem to miss the real point of

¹⁵⁸ I follow essentially Newmeyer 1998 in that syntax puts constraints on what is possible (P-stranding, CED, ...) but that further considerations from other domains may

the example, namely that movement out of a reanalyzed string must be allowed. The same problem arises for a derivation of (333d).

Hornstein and Weinberg note (Hornstein and Weinberg 1981:70 fn. 19) that their theory predicts that extraposition out of reanalyzed strings ought to be impossible. They offer the examples in (334) to demonstrate that the expectation is borne out. The crucial example is (334d), where PP extraposition of *about Rembrandt* would have to occur out of the reanalysed string *put a book about Rembrandt on*. The trouble with this argument according to Levine (1984 and others) is that there are also examples like (335) (Newmeyer 1998:5 ex. 21), which demand extraposition out of a reanalysed string. In any case, even the non-extraposed version of examples like (335) given in (336) appear to pose insurmountable problems for Hornstein and Weinberg's analysis because all of them necessarily involve the application of transformations (relativization in (336a) and (336b), passivization and relativization in (336c)) within the reanalysed string, but to achieve any empirical coverage, as noted in Hornstein and Weinberg, reanalysis must apply in the base.

- (334) a. ✓John put a book about Rembrandt on all the desks yesterday.
 b. ✓John put a book on all the desks yesterday about Rembrandt.
 c. ✓Which desks did John put a book about Rembrandt on yesterday?
 d. *?Which desks did John put a book on yesterday about Rembrandt?
 e. ✓On which desks did John put a book yesterday about Rembrandt?
- (335) ✓What did you talk to the guy ____ about t_{what} who was here yesterday?
- (336) a. ✓What did you talk to the guy [_{CP} who [_{IP} t_{who} was here yesterday]] about?
 b. ✓What did you talk to the guy who Peter kissed t_{who} on the face about?
 c. ✓What did you talk to the guy who was kissed on the face about?

impact the status of particular examples.

A further problem noted in Koster 1986 (see also Baltin and Postal 1996; Newmeyer 1998 and others) is that the reanalyzed strings do not behave as words for various processes such as gapping. It must be possible to reanalyze *look at* to derive example (337b). The question is then why the verb *look_at* cannot be gapped (337a).

- (337) a. John looked at Mary and Bill ____ *(at) Sue.
 b. Who did you look at?

Finally, consider again the filter (328). It is assumed to be universal and its applicability is not restricted to traces: "This filter states that noun phrases with no lexical material (e.g. traces) which are marked oblique by the Case-marking conventions are to be ruled ungrammatical." (Hornstein and Weinberg 1981:60) To rule out P-stranding in French, French prepositions must assign oblique Case and the filter must apply in French. The trouble is that the filter also rules out the null pronominals that we find in French orphan-P constructions. This suggests a much narrower formulation of the filter, namely (338). Hornstein and Weinberg (1981:79) briefly discuss a theory containing (338) and their reanalysis rule. They reject this theory on the basis of its inelegance and its limited applicability. Filter (338) is of course what I derive as a theorem for non-stranding languages. Extraction out of PP is mainly governed by a theory of CED in stranding languages on the present view.

- (338) *_[PP P t]

Notice that the present view entails that the reasons why P-stranding in non-stranding languages is impossible are quite different from the reasons why it is sometimes impossible in stranding languages.

This predicts an asymmetry between stranding and non-stranding languages for sluicing. In cases where P-stranding is impossible in stranding languages, this impossibility is an island effect whereas in non-stranding languages this impossibility is a violation of the identity condition on deletion. As discussed above, island effects can be ameliorated under sluicing but not violations of the identity condition. Thus the non-stranding cases should ameliorate under sluicing in English. This prediction seems to be borne out.¹⁵⁹

- (339) a. #Who did you destroy a book about?
b. ✓John destroyed a book about somebody, but I can't remember who.
- (340) a. %Which inning did the Yankees lose the ball-game in?
b. ✓The Yankees lost the ball-game in some inning, but I can't remember which (inning).
- (341) a. *Which operations does your computer hang during?
b. ✓My computer hangs during some operations, but I don't know which (operations).
- (342) a. *Which professors does Luisa seem to be a genius?
b. ✓Luisa seems to some professors to be a genius, but I can't remember which (professors).

I conclude with the bulk of the literature cited above, that literal reanalysis just does not work.¹⁶⁰ This discussion has brought out an important point though. The internal structure/phase-nature of P° is not the be all and end all of extraction from PP.

¹⁵⁹ The crosshedge '#' in example (339a) indicates that speakers will generally reject the example in an out of the blue context. Takami 1992 shows that such examples become acceptable with sufficient contextual support. The '%' mark in front of (340a) is intended to signify an example that is assigned a different status by different speakers.

¹⁶⁰ Stowell 1982a suggests a more restricted version of literal reanalysis, which is in general better supported by the facts. Stowell's analysis links the availability of pseudopassives to the availability of V-particle-NP constructions, i.e. to a version of Generalization [5] (see Sugisaki et al. 2001; Sugisaki and Snyder 2001 for recent discussion). Stowell's generalization is supported by the facts of language acquisition,

Finally, let me briefly remark on government based approaches to P-stranding. Kayne, who pioneered this approach to P-stranding, rejects Hornstein and Weinberg's reanalysis approach in favor of an approach in terms of government (Kayne 1984:XII).¹⁶¹ He suggests a system where two parameters are involved.

- (343) a. P structurally governs NP.
b. P governs NP only in the sense of subcategorization.
- (344) a. P assigns structural accusative Case.
b. P assigns structural oblique Case.

The parameter in (344) is a sub-parameter of (343) in the sense that it is active (or relevant?) only in case (343a) is chosen as the value for the first parameter.

(343b) is the value of the parameter found in non-P-stranding languages. Kayne rules out P-stranding in this case by appealing to the Empty Category Principle (ECP). Kayne assumes that the ECP is defined as in (345). The crucial notion of Percolation Projection is defined in (346).

but Stowell's analysis has nothing to say about A'-stranding. Presumably, Stowell's analysis would run into the same kinds of problems with the typological facts (Generalization [3]) as van Riemsdijk's (1978a). Note also that the explanation for the facts from acquisition given in (Sugisaki et al. 2001) conflicts with the findings in Ramchand and Svenonius 2002; Svenonius 2002 concerning the base order for verb-particle constructions in English. The facts are still too unclear to warrant a conclusion one way or another at this point.

¹⁶¹ Over a number of papers Kayne moves further and further away from the reanalysis approach. The formulations used in the process are sometimes misleading, thus consider the following quote: "We shall accept the existence of a Reanalysis rule in English, while following Vergnaud's (1979) suggestion that [...] what is involved is not so much reanalysis qua constituent as reanalysis in terms of government [...]" (Kayne 1984:114-115). 'Reanalysis in terms of government' boils down to co-superscripting under government for the formation of a 'percolation projection' (Kayne 1984:58).

- (345) *Empty Category Principle* (ECP) (Kayne 1984:58)
 An empty category β must have an antecedent α such that (1) α governs β or (2) α c-commands β and there exists a lexical category X such that X governs β and α is contained in some percolation projection of X.
- (346) *Percolation Projection* (Kayne 1984:57)
 A is a percolation projection of B if A is a projection of B, or A is a projection of C, where C bears the same superscript as B and governs a projection of B, or a percolation projection of B.

Kayne assumes that ‘reanalysis in terms of government’ amounts to co-superscripting the two categories that are reanalyzed. He further proposes that “reanalysis between two lexical categories is possible only if the two govern in the same way” (Kayne 1984:116). Verbs govern structurally.

It now follows that in a language that chooses parameter setting (343b) preposition stranding is ruled out. This is illustrated by the ungrammatical French example (347). *Voté* bears superscript i, *pour* bears superscript k. They cannot be cosuperscripted because they do not govern in the same way. Therefore the percolation projection of the preposition stops at the PP boundary. It can never go higher. But then the antecedent of the trace, *qui*, is not contained within the same percolation projection as the governor of the trace *pour* and the structure is correctly ruled out by the ECP.¹⁶²

- (347) Qui as tu [VP votéⁱ [PP pour^{k/i} t_{qui}^{k/i}]ⁱ]
 who have you voted for ?

¹⁶² The discussion concentrates only on the relevant aspect here. Of course we also have to make sure that the government superscript can percolate up from the VP to the sentence level. If such percolation were impossible, simple object questions would be ruled out incorrectly. To solve this Kayne assumes that the sentence is a projection of the verb.

Two comments are in order here. First of all, Kayne must assume that PPs do not have escape hatches. Otherwise the structure could be as in (348). Still the PP cannot be co-superscripted with the verb, but now t_{qui} satisfies the ECP. It is governed by the preposition and its antecedent, t'_{qui} is contained within the percolation projection of t_{qui} 's governor, PP. t'_{qui} also satisfies the ECP since it is governed by the verb¹⁶³ and its antecedent *qui* is contained in the percolation projection of the verb.¹⁶⁴

- (348) Qui as tu [VP votéⁱ [t' _{qui} [pour^{k/i} t_{qui}^{k/i}]^{k/i}]ⁱ]
 ?

The second remark concerns the consequences these assumptions have for Generalization [2]. Since a percolation projection stops at the PP level and does not go up to the verb in non-P-stranding languages (347) and since escape hatches cannot be generally available as we just saw, all extraction from PP is blocked.¹⁶⁵

¹⁶³ This kind of government from head to the specifier of the complement is crucially appealed to in Kayne's account of prepositional complementizers.

¹⁶⁴ Kayne factually seems to assume that there is an additional parameter. Namely van Riemsdijk's [+/- escape hatch] parameter. This can be seen from the following quote:

“Van Riemsdijk's [(1978a)] analysis of Dutch preposition stranding seems essentially correct, and does not involve reanalysis. (That is, the R-pronoun ‘escape hatch position’ obviates the need for reanalysis; the ECP still imposes government [...].)” (Kayne 1984:116)

This amounts to a substantial weakening of the predictions the system makes. Having argued that Dutch does not allow P-stranding, I will ignore the consequences this additional parameter has for Kayne's system.

¹⁶⁵ The essential transitivity of the notion of percolation projection is preserved under Chomsky's (1986a) formulation so that, everything else being equal, the same false prediction concerning Generalization [2] is made.

For a P-stranding language the other value of parameter (343) is chosen, i.e. P structurally governs NP. This allows co-superscripting of PP with V. Given this assumption the English equivalent of (347) is admissible with co-superscripting. As a result the percolation projection of P may be the entire sentence, which is, therefore, predicted to be grammatical in English.

(349) who have you [_{VP} votedⁱ [_{PP} for^{k/i} t_{qui}]^{k/i}]ⁱ?

I will not discuss in detail the workings of the second parameter (344). It is a late addition to the system intended to capture the fact that while Icelandic (and some other languages) allow P-stranding under A'-movement it does not do so under A-movement, i.e. it accounts for Generalization [3].

Let's turn to Generalization [4]. Does Kayne's system predict the fact that sluicing, while ameliorating island effects does not ameliorate P-stranding violations in non-P-stranding languages? And does it predict that P-stranding violations can be ameliorated under sluicing in English but not in non-P-stranding languages as we saw in example (314)? Given that the discovery of Generalization [4] substantially postdates Kayne's work, he does not discuss these matters. I will refrain from constructing straw men. As far as I can see, incorporating Generalization [4] into Kayne's system is a non-trivial task. The Identity Solution proposed above does not translate straightforwardly.¹⁶⁶

¹⁶⁶ Kayne explicitly makes a number of predictions about cross-linguistic variation:

- (i) All languages that allow double-accusative constructions will also allow pseudopassives.

4.5.4 Conclusion

In this section I have briefly discussed the syntax of P-stranding as opposed to the lack thereof. I briefly discussed the main generalizations concerning P-stranding that emerge from the literature. I have shown how my own account of the phenomenon in terms of two parameters can account for some of them. The parameters involved the phase nature of the head P° and the question whether P° obligatorily assigns Case or not.

There appears to be an implication from the availability of P-stranding to the availability of verb-particle constructions. I left this apparently real but mysterious connection as an issue for further research. Although my account of Generalization [4] does not follow without further assumptions from the theory developed so far, the present theory which reduces successive cyclicity to Attract Closest appears to make better predictions than a theory that treats successive cyclicity (the Head Constraint) and Attract Closest as two distinct principles of locality.

-
- (ii) All languages that have prepositional complementizers of the English (Case marking) type also allow pseudopassives.

These predictions are discussed in King and Roberge 1990; Sugisaki 2002; Zhang 1990; Sugisaki 2002; Sugisaki and Snyder 2001. Prediction (i) is quite certainly false (see King and Roberge 1990; Sugisaki 2002; Zhang 1990). Prediction (ii) has no crosslinguistic support (see King and Roberge 1990; Mensching 2000:162 on Brazilian Portuguese and Tallerman 1998 for relevant discussion of the Welsh complementizer *i* see also Borsley 1986; Harlow 1992; Hendrick 1988; Rouveret 1990, 1994) but appears to be supported by data from the acquisition of English (Sugisaki 2002; Sugisaki and Snyder 2001).

I then went on to discuss previous approaches to P-stranding. Van Riemsdijk's approach and related approaches seemed ill-suited in principle to account for Generalization [2] and Generalization [3] in particular. Hornstein and Weinberg's approach was shown to be fraught with so many problems that I did not pursue the matter. I illustrated on the basis of sluicing facts that the absence of P-stranding in certain contexts in English and the absence of P-stranding in general in non-P-stranding languages are not the same phenomenon, because sluicing leads to an amelioration effect in English, as I show, but not in non-P-stranding languages, as Merchant has shown. This lends additional support to the analysis of the sluicing facts developed in section 5.4.2.

Finally I discussed Kayne's government based approach to P-stranding. Again it is ill-equipped to account for Generalization [2]. Other predictions of Kayne's approach are either not borne out or the facts are at best unclear (see fn. 166).

4.6 Conclusion

In this chapter I have applied the logic of anti-locality to prepositions and prepositions stranding. I argued that in non-P-stranding languages extraction of the complement of P° is disallowed, while extraction out of the complement of P° may be allowed. This follows on the present theory. I then turned to the issue of how to diagnose P-stranding and suggested the possibility to have comparatives on the complement of P° as a novel diagnostic, which I applied to Frisian, concluding that Frisian does have P-stranding. I had to leave dangling the issue of Papiamentu.

I then argued that German and Dutch do not allow P-stranding. As a further consequence of the analysis, it was shown that non-P-stranding languages are predicted

not to allow pronominal clitics as the complement of prepositions. Finally I turned to the issue of P-stranding, summarizing the main results from previous work and suggesting an analysis. I concluded by briefly reviewing the three main previous approaches arguing that the present approach fares better with respect to core properties of P-stranding.

Chapter 5: Head Movement and Anti-locality*

This chapter is about Head Movement. We saw in chapter 2 that Head Movement which obeys the Travis's 1984 Head Movement Constraint are ruled out in general by Last Resort. A head H° is in the closest feature satisfaction relation with its sister αP : mutual total c-command. αP is a feature bearer of the features of α° .¹⁶⁷ Therefore H° can never attract α° . On the other hand, we know that there are facts such as subject auxiliary inversion that are typically accounted for in terms of head movement.

I will adapt the system developed in Brody (2000 – see Abels 2000 for elaboration of the system, albeit with a different set of assumptions about phrase structure in mind) for my own purposes. Brody treats what is traditionally analyzed as head movement as a reflex of morphological word formation and linearization. One of the main differences between Brody's system and the standard system is that in the standard system Head Movement always goes upward whereas Brody's system allows what appears to be downward Head Movement. I argue that this is a prediction that one would

* Special thanks go to Željko Bošković, Howard Lasnik, Andrew Nevins, and, of course, Luisa Martí. I am also indebted to the audiences at a UConn Linglunch (spring 2001), the 24th GLOW Colloquium, the first TiLT workshop, the 26th UPenn Colloquium, and the XII Colloquium on Generative Grammar, where versions of this material were presented.

¹⁶⁷ I assumed that there are certain features on phase heads that do not percolate up from the head to the phrase. These features are, by assumption, also non-attractable features. Therefore the fact that in certain cases a feature of α is not on αP does not have

actually want to make to solve some long standing problems. But I am getting ahead of myself.

The story starts and ends with generalization (350) from Sag 1978. (350) says that in English an adverb or a floated quantifier cannot immediately precede the site of VP-ellipsis (VPE) or VP-fronting (VPF) – summarily *E-site*. For discussion see Baker (1971; 1981); Ernst (1983); Sag (1978; 1980). The basic facts illustrating (350) for VPE are given in (351) and (352) and for VPF in (353) and (354).

- (350) Sag's Generalization: $\ast\{Q/Adv\} - E\text{-site}$
- (351) Fred has never been rude to Grandfather, but John {✓always has; \ast has always}.
- (352) Some of them are working hard. In fact, they {✓all are; \ast are all}.
- (353) He said he would study karate, and study karate he {✓surely has; \ast has surely}.
- (354) They said they would study karate, and study karate they {✓all have; \ast have all}.

This chapter suggests an account of these facts, which deeply implicates Head Movement (HM). The present theory of HM has two crucial non-standard properties: (i) downward HM is allowed (see also Brody 2000; Cormack and Smith 1998) and (ii) a head's position depends largely on its (local) syntactic environment.

The chapter is organized as follows. After discussing a problem with the standard theory of HM, subsection 5.1 introduces the theory assumed in the rest of this chapter. Section 5.2 investigates Sag's generalization in some detail, pointing out problems both with the generalization and existing accounts of it. Section 5.3 shows how the data can be accounted for under the present theory.

an impact on my point here. α° will not be attracted on account of a feature it bears which did not percolate because that feature is not an attractable feature.

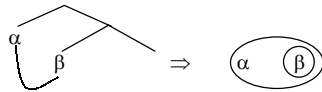
5.1 Head Movement

5.1.1 *The The-higher-the-bigger-Theorem*

Standard theories of HM assume (i) that HM proceeds upward and upward only,¹⁶⁸ (ii) that HM obeys Travis's (1984) Head Movement Constraint,¹⁶⁹ and (iii) that successive HM does not excorporate.¹⁷⁰ Jointly (i) and (iii) entail (355).

(355) *The-higher-the-bigger-Theorem:*

For all heads α , β that are members of a single head-chain, if α c-commands β , then the features of β are a proper subset of those of α .



Theorem (355) makes a prediction: Information can migrate up the tree but not down. Thus, tense information, which is base generated in T° , should appear in T° or higher but not lower. Similarly, information about clause type is generated in C° . This information should thus never appear below C° .

¹⁶⁸ (i) is often deduced from some version of Fiengo's (1974; 1977) Proper Binding Condition or the bottom-up nature of syntactic derivations (cf. e.g. Epstein 2001).

¹⁶⁹ (ii) is often viewed as an instance of Rizzi's (1990) Relativized Minimality.

¹⁷⁰ In Chomsky's (1995b) system (iii) follows from the uniformity condition on chains. Roberts 1991, Boskovic 1994b, Matushansky 2002 among others assume that HM allows excorporation. I can't address their arguments here and adopt the standard view. Cases of apparent excorporation must be analyzed in one of two ways: long head movement or remnant movement. I cannot address the issue here.

The Higher The Bigger Theorem does not appear to encode a true generalization about language. English finite verbs famously violate (355): Although V does not raise to T° in English, V shows agreement and tense information. This violation of (355) is often treated by invoking Affix Hopping Chomsky 1957 or morphological merger under 'adjacency' (Bobaljik 1994, 1995b) (356). On the standard view, 'adjacency' tolerates intervening adverbs but not negation or arguments (for discussion cf. Abels 2001a; Stepanov 2001).

$$(356) \quad \alpha_{\text{affix}} \beta \rightarrow \beta + \alpha$$

Affix Hopping is suspect on several grounds. First, it is largely redundant with HM. Both processes target heads and both give rise to the same order: the higher head follows the lower one. Second, rules in natural languages are generally structure dependent. Affix Hopping, being a linear rule, then poses a learnability problem. Third, Bošković 2001b shows that Prosodic Inversion (which is homologous to Affix Hopping) fails to account for the facts concerning South Slavic clitics that it was originally designed for (example (???) from chapter 4.1.1 exemplifies this failure). It fails precisely because of its linear as opposed to structure dependent character.

But the problems do not end here. Some violations of (355) are not amenable to an Affix Hopping treatment. Consider the Vata examples (357)–(358) from Koopman (1984:74 ex76, 61 ex. 46). (357) shows that Vata has a head-final CP. The verb moves to T° . According to (355) information residing in C° cannot show up in T° . Yet, the fact that (358) is a relative clause is marked on the verb (REL in the gloss). The verb is not

adjacent to C° ; it is sandwiched between its arguments; Affix Hopping fails (cf. also Tuller 1986 for Hausa).

(357) àlÓ Ò wà sá ká lá
 who he-R want rice WH-Comp
 Who wants some Rice?

(358) kÓ` (mômÓ`) Ò lī -dā -bo zué sá ká,...
 man HIM-HIM he-R eat -PT -REL yesterday rice
 the man who was eating rice yesterday, ...

To conclude, Affix Hopping is neither a desirable mechanism nor can it account for all the violations of the The-higher-the-bigger-Theorem.¹⁷¹

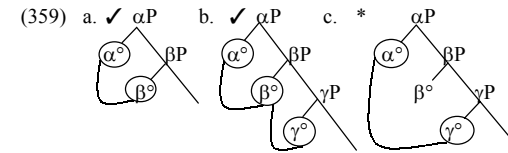
5.1.2 Head Movement in Mirror Theory Brody 2000

To solve the problems from section 1.1, we assume that HM is post-syntactic and modular Abels 2001a; Brody 2000.¹⁷² The two modules producing HM effects are (i) Morphological Word Formation (MWF)—words are formed under the necessary but not sufficient condition that one head be the head of the complement of the other, and (ii) Positioning—a unique linear position for elements with multiple hierarchical positions is determined. We examine the modules in turn (for details Brody 1997, 2000; Abels 2000).

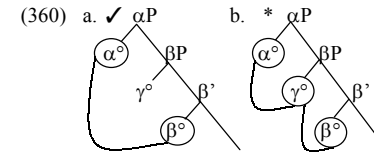
The heads joined by an arc form a word. In (359a) and (359b), β° is the head of the complement of α° , and in (359b) γ° is the head of the complement of β° and MWF

¹⁷¹ Chomsky (1993) voids theorem (355) by adopting strict lexicalism. This move avoids the problems with Affix Hopping, but it also denies the possibility of giving a syntactic account of Baker's (1985) Mirror Principle (see Brody 2000 for discussion).

can take place. The condition that one head be the head of the complement of the other can be understood transitively. However, MWF of α° with γ° is blocked in (359c), because the intervening head β° is skipped. This captures the HMC and the ban against excorporation.



In (360a) and (360b) β° is again the head of the complement of α° . γ° is the specifier of βP . Therefore, MWF is allowed in (360a) but not in (360b).



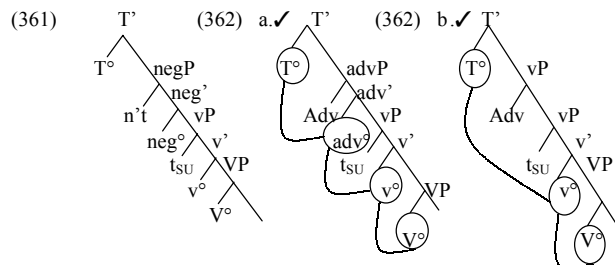
Above we characterized the structural configuration head-of-complement-of as a necessary but not sufficient condition. To see why, consider some examples. Structure (359a) could be instantiated by taking $\alpha = v^\circ$ and $\beta = say$. These two can form a word: ✓ *say+v*. (359b) could be instantiated by taking $\alpha = T^\circ$, $\beta = v^\circ$, and $\gamma = say$. These three heads can form a word: ✓ *say+v+T*. The same configuration would also be instantiated by taking $\alpha = v^\circ$, $\beta = say$, and $\gamma = that$. These cannot form a word by morphological fiat:

¹⁷² HM has always been late: Chomsky 1957; Fillmore 1965; Baker 1971. For recent discussion see also Boeckx and Stjepanović Boeckx and Stjepanović 2001;

*that+say+v. Implicit in these examples is the assumption that the top to bottom order of syntactic heads determines the right to left order of morphemes. This captures the Mirror Principle (Baker 1985, 1988; Brody 2000).

Usually, Affix Hopping is assumed to be blocked by intervening negation but not by adverbs. This fact can be re-described in the present theory as a statement about MWF. Assuming that (361) - with $n't$ as the specifier of NegP - is the correct syntactic structure, we simply say that neg^o can form a word with auxiliaries and T^o , but not with main verbs V: $*V+v+neg$.

Example (362) depicts the situation for adverbs. If adverbs are specifiers of dedicated functional heads (Alexiadou 1997; Cinque 1999), then these heads must be able to form a word with T^o , v^o and V^o (362a). Alternatively, if they are adjuncts, then no complications arise since no heads intervene (362b).

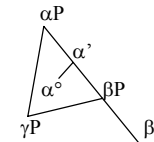


Let's turn to the second module of the present theory of HM: Positioning. Labels aside, the structure assumed for a complex word like (360a) has an important property in common with structures created by movement (363). The similarity between the different

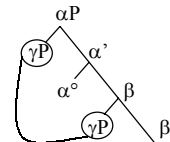
Chomsky 2000; 2001a.

structures comes out more clearly optically if we assume chains. Thus (364) is the chain-style representation of (363).

(363) Movement as Multidomination



(364) Movement as Chain



The only thing that is crucial about these pictures at the moment is that they contain objects that have multiple mothers. The Morphological Word in (362a) has multiple mothers (T^o , adv^o , v^o , V^o), the moved constituent γP in the multi-domination structure (363) has multiple mothers (β , αP), and the chain in (364) has multiple mothers (β , αP). Multi-headed words and movement chains give rise to the same problem: What is the (unique) linear position of an item occupying several hierarchical positions simultaneously? Assuming that positions in a word/chain can be strong (s) or weak (w), the linearization problem has the general solution (365) in a single output syntax (Bobaljik 1995a; Brody 1995; Gärtner 2002; Groat and O'Neil 1996; Kayne 1998).¹⁷³

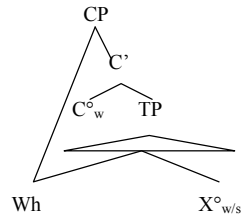
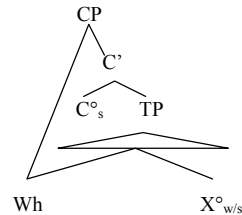
¹⁷³ Notice that the present system with its strong commitment to the extension condition on Merge must be a single cycle/single output system. A question arises what

(365) Positioning Algorithm

Pronounce an element E (a word or a chain) in the lowest position P such that all higher positions P' of E are weak.

Algorithm (365) says that an element is pronounced in the highest strong position.

If there is no strong position, it is pronounced in the lowest position, i.e. the base position (cf. Brody 2000; Gärtner 1999; Abels 2001a). The formulation allows for multiple strong positions within a chain or word.

(366) a. *wh*-in situ(366) b. *wh*-movement

The difference between *wh*-in situ and *wh*-movement can be captured easily.

(366a) is the case of a *wh*-in situ language: the *wh*-element moves to SpecCP, but the high position is weak. This is indicated on C°, the relevant attractor which determines whether we are dealing with strong or weak attraction. According to the Positioning

the exact status of covert movement is in the system developed in chapter 2. Given that features can be satisfied at a distance, is there a necessity for covert movement and weak members in a chain. Pesetsky 2000 argues that indeed we need to distinguish agreement at a distance from covert phrasal movement. If this is true then complex words and chains are fully symmetric. If covert phrasal movement can be eliminated (as I tacitly assumed in chapter 2), then there is an asymmetry between complex words which can contain more than one weak position and chains which would contain at most one weak position (the foot). I will assume in the text that complex words and movement chains are indeed fully parallel, i.e. movement chains can have strong and weak members/links.

Generalization, the *wh*-chain is realized in the base position in (366a). In (366b) on the other hand, the high position is strong and the *wh*-chain is phonologically realized in the high position.

The difference in terms of V-to-T movement between English and French (see Emonds 1978; Pollock 1989b but see Williams 1994 for a dissenting view – the example given here just serves to illustrate how the system works, not to endorse a particular analysis) is captured in the same way. In both language the verb and T form a word: #V+v+T#. In French T° is strong (#V+v+T_s#), but in English T° is weak (#V+v_s+T_w# – following Koizumi 1995; Lasnik 1995c, I assume that v° is strong in English). This is the only relevant difference between the languages.

Note how this view solves the problems for theorem (355) discussed above. MWF obeys the strict locality usually associated with HM, MWF obeys the non-excorporation condition, but Positioning, crucially, does not give rise to the The-higher-the-bigger-theorem (see also Zwart 2001 for a similar proposal).

We can summarize the results so far as follows. The Mirror Theoretical view of HM does not entail the problematic theorem (355), i.e., downward HM is allowed. HM arises as the combined effect of post syntactic MWF (a possibility inherent in the standard theory, see Chomsky and Lasnik 1993) and post syntactic positioning.

5.2 *{Q/Adv} – E-site

This section discusses Sag's generalization (350) in some detail. After noting a problem for (350), we discuss Sag's (1995) explanation of (350) and show that they run into

additional problems. Before that though, we will broaden the array of data under consideration somewhat.

The auxiliary adverb order is not the only factor interacting with VPE. Baker (1971; 1981) notes that the stress level of auxiliaries also interacts with ellipsis and word order (see also Wilder). The fact is that unstressed (tensed) auxiliaries never follow adverbs (367) and (368).¹⁷⁴

(367) John always {✓h[æ]s; * h[P]s; * 's} been rude to grandpa.

(368) He said he would study karate, and study karate he surely {✓h[æ]s; * h[P]s; * 's}.

A theory of auxiliary adverb word order has to account for this fact along with Sag's generalization.

¹⁷⁴ There is some discussion whether the converse also holds, i.e., whether stressed auxiliaries can precede adverbs. The answer seems to be 'yes' (Baker 1981 with reservations; Ernst 1983; Sag 1980; Wilder - contra Baker 1971: (i)-(ii).

- (i) John {✓h[æ]s; ✓h[P]s; ✓ 's} always been rude to grandpa.
- (ii) ✓She said that they'd all read *Moby Dick*, and they HAD all read it.

5.2.1 Sag's Generalization

Recall Sag's generalization (350), which was exemplified above in (351)-(354). Is Sag's generalization empirically correct? Superficial counterexamples are easy to construct (369). Examples like (369) are only apparent counterexamples, however. Sag's generalization says that that adverbs and floated quantifiers cannot precede the E-site. In (369) the adverb actually follows the (silent) E-site (370). The fact that (370) is the correct analysis of (369) can be shown by replacing *slowly* by an adverb that cannot appear in VP-final positions such as *hardly* (371). Example (372), which is parallel to (369) is ungrammatical. Example (369) is therefore not a true counterexample to Sag's generalization.

(369) ✓John writes fast, but Peter does slowly.

(370) John writes fast, but Peter –s ~~write~~ slowly

(371) Ian's theory has {*surprised me hardly; ✓hardly surprised me}.

(372) *Ed's theory has utterly surprised me, but Ian's theory has hardly.

(373) ✓Ed's theory has utterly surprised me, but Ian's theory hardlyhas.

The next set of examples (from Baker 1981) are true counterexamples to Sag's generalization. Example (375) shows that *always* cannot appear clause finally. In (374) *always* thus immediately precedes the E-site in violation of (350). The same is shown for floated quantifiers in (376) and (377). The crucial factor seems to be the presence of negation in (374) and (376). Examples (378)-(380) show that matrix interrogatives can also violate Sag's generalization.

(374) ✓ Fred has sometimes been rude to Grandpa, but he hasn't always.

(375) ?? Fred has sometimes been rude to Grandfather, but he hasn't been rude to Grandfather always.

(376) ? Some of them are working on the assignment, but they aren't all.

(377) * Some of the boys are working on the assignment, but they aren't working on the assignment all.

- (378) ✓ Some of the boys are working on the assignment. - Are they all?
 (379) John hasn't gotten along with Grandpa lately. — ✓ Has he ever?
 (380) John hasn't gotten along with Grandpa lately. — ?* Has he gotten along with Grandpa ever?

Sag's generalization is too strong. Informally speaking, the auxiliary can be drawn away from its position immediately preceding the E-site by negation and by the interrogative complementizer.

5.2.2 The Radical Emptiness Account of Sag's Generalization

Sag and Fodor (1995) attempt to derive Sag's generalization. Although coming from very different theoretical viewpoints, the explanation is the same: Sag's generalization holds because (at the relevant level of representation) the E-site is literally empty. If the E-site is empty, there is no structure the adverb and floated quantifiers could possibly adjoin to. Since adverbs and floated quantifiers need a host, they must adjoin higher, placing them to the left of the auxiliary. The account is appealingly simple.

The counterexamples to Sag's generalization involving negation and matrix interrogatives, noted above, remain problematic. There are several further classes of examples that pose a challenge to the account of Sag's generalization in terms of radical emptiness.

The first problem is posed by pseudogapping (cf. Levin 1978, 1979/1986).

Pseudogapping has been analyzed as a kind of VPE (Jayaseelan 1990; Johnson 1996;

Lasnik 1995c).¹⁷⁵ In typical pseudogapping examples like (381a), there is clearly some structure present in the E-site: *her essay*. Yet the auxiliary must follow the adverb (381b).

- (381) a. ✓Joe quickly ripped up his paper, and Sue slowly did her essay.
 b. *Joe quickly ripped up his paper, and Sue did slowly her essay.

This paradigm cannot be accounted for by appealing to radical emptiness. To avoid this problem, proponents of radical emptiness would have to offer a convincing alternative analysis of pseudogapping.

The next two sets of examples (from Abels 2001a) show that radical emptiness comes at the cost of enriching the theory of adjunction by ad hoc assumptions. Adverbs like *completely* cannot usually precede tensed auxiliaries (382). In fact Jackendoff (1972) claims that *completely* is within VP (cf. also Lasnik in press-b; Oku 1998). However, under VPE the order *completely* > auxiliary becomes not only possible but obligatory as shown in (383). Under the radical emptiness assumption, (383b) is expected to be ungrammatical. But why is (383a) acceptable? If the range of categories an adverb can adjoin to is stable across constructions, the grammaticality of (383a) remains mysterious under the radical emptiness approach. A construction specific theory of adjunction sites must be invoked.

- (382) a. *Sue completely has finished her essay.
 b. ✓Sue has completely finished her essay.
 (383) a. ✓Joe partially revised his paper, and Sue completely did.
 b. ?*Joe partially revised his paper, and Sue did completely.

¹⁷⁵ Not all speakers of English accept pseudogapping with equal ease.

Some adverbs receive different readings depending on their position, as shown for *happily* in (384). Crucially the order *adv*>*aux* forces the speaker oriented reading (384a). The VPE example (385) is a counterexample to the general pattern, since both readings are available. Radical emptiness demands that *happily* in (385) is adjoined in the same position as in (384a). This makes the false prediction that (385) only has the speaker oriented reading.

- (384) a. *Adv* > *Aux* John happily will return to his village.
 ✓ speaker oriented reading, ?* manner reading
 b. *Aux* > *Adv* John will happily return to his village.
 ? speaker oriented reading, ✓ manner reading
 (385) John will return to his village and Bill happily will, too.
 ✓ speaker oriented reading, ✓ manner reading

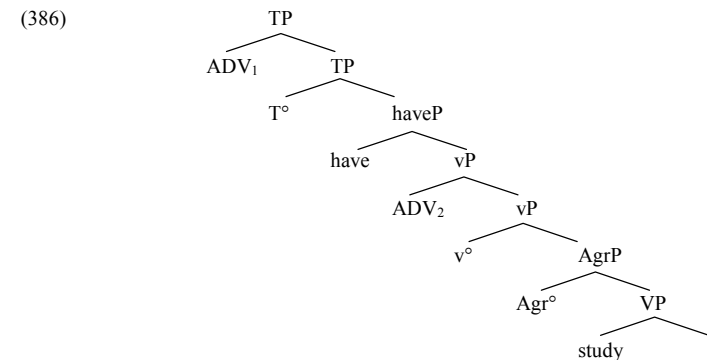
All of this suggests that the radical emptiness approach is wrong. Moreover, the data make sense if the base position of the auxiliary is above the relevant adverbs. The auxiliary would then move below the position of those adverbs in VPE contexts including pseudogapping, but could be attracted to a higher position if negation or the interrogative complementizer *C_Q* is present. The next section develops such an account.

5.3 The Account

The theory from section 1 accounts for the facts if the following additional assumptions are made: (i) Abstract heads need to be licensed. I will assume that one licensing mechanism for abstract heads is to form a word with overt material. Some principle regulating the distribution of empty heads is necessary if syntax is to have empirical content. Under the name of ECP, the principle regulating the occurrence of abstract elements was at the center of syntactic debate in the eighties. The interest has

unfortunately subsided, but I will assume that word formation is part of an ultimate theory of licensing of abstract elements. (ii) Ellipsis and movement target only phrasal categories. (iii) Every phrase has a head (endocentricity).

Following Williams 1994 I will assume that adverbs which can appear before or after the auxiliary can attach in several different sites, TP and vP (presumably also haveP – though I do not show this in the tree to keep it from becoming too cluttered). I leave all specifiers out of the tree to keep it simple. The structure is simplified in yet another way. The structure underlying simple clauses with such adverbs is shown in (386). Based on (386) two words will be formed as shown in (387).



- (387) #have+T°# #study+Agr°+v°#

As we will see, this is not the only possible pattern of word formation. Under ellipsis of various sizes, I will also invoke the morphological words indicated in (388), where *strikeout* indicates elision.

- (388) a. #v°+have+T°# ~~study~~ Agr° AgrP ellipsis
 b. #Agr°+v°+have+T°# ~~study~~ VP ellipsis

Of course, we now have a question on our hands. If ellipsis does not apply, is the word formation pattern in (387) forced? Or is it in principle possible to find the non-elliptical patterns in (389) alongside that in (387)? As we will see shortly, I must assume that (387) is forced and that the patterns in (389) are ruled out as indicated by the asterisks. The next question to ask then is why that should be so. I will give a brief answer to this question before turning to the ellipsis facts. There are two ways one could go, a representational way and a derivational way.

- (389) a. * #v°+have+T°# #study Agr°#
 b. * #Agr°+v°+have+T°# #study#

On the representational approach the structures in (389) could be ruled out because there is something wrong with the words formed – the only words that appear in (389) and that I don't evoke independently are the combinations #study+Agr°# and #study#. We could claim that these combinations violate some requirement X. Indeed, in Abels 2003a I suggest just that. Essentially, the assumption was that a root, here *study*, cannot be realized unless it is conferred a category, which is the job that only v° can do (see Marantz 2001). The problem with this idea is, that I do evoke words like the ones in (389). In chapter 3 I crucially assumed following Wurmbrand 2001 that bare VPs without embedding vPs above them are possible. This assumption is heavily involved in Wurmbrand's explanation of the long passive in German. I therefore must reject this option.

This leaves us with the second option, that there is nothing wrong with the structures per se. Rather they cannot be derived. This would follow if we make the following assumptions: (i) Morphological word formation applies after ellipsis; (ii)

morphological word formation is a bottom up process; (iii) the process that creates morphological words maximizes these words, that is it keeps lumping heads together in a sequence as long as possible. Starting at the bottom this algorithm must lump together *study* and Agr° and v° because it can. It must stop there, because the morpheme *have* never forms part of an independent main verb in English. Assumption (i) seems independently plausible. If morphological word formation applied before ellipsis, we would end up under certain circumstances with XP elisions that elide parts of words already formed. This must be excluded empirically as we will see, but it is also a plausible constraint on XP deletion that it cannot elide parts of words. Assumption (ii) appears to be the most computationally plausible way of implementing morphological word formation (Ed Stabler, p.c.). Assumption (iii) is empirically justified. Together these assumptions derive the fact that in the absence of ellipsis (387) describes the only possible situation.

We now turn to positioning in the case without ellipsis and then to ellipsis.

Consider again the structure in (387). What is the result of positioning?

- (387) #have+T°# #study+Agr°+v°#

Since T° and *have* are both weak (by assumption) #have+T°# is pronounced in the lower of the two positions. Recall that I follow Koizumi 1995; Lasnik 1995c and assume that v° is strong in English. Therefore, #study+Agr°+v°# will be pronounced in the

position of v° . Depending on where the adverb attaches *has* precedes or follows the adverb. None of this is very exciting so far.¹⁷⁶

We now turn to VPE. Suppose that VPE is PF-deletion. The underlying structure is still the tree in (386). Suppose that deletion can apply freely to vP , $AgrP$, or VP . If vP is elided, any adverb that appears must have been attached higher than vP , i.e. it must have been attached to $haveP$ or TP . Since the word $\#have+T^\circ\#$ is realized as before in the position of *have* it will necessarily follow all adverbs under ellipsis. Recall though, that there are adverbs that do not attach as high as $haveP$ or TP . An example discussed above was *completely* (382). The surprising fact was that just in case ellipsis takes place, *completely* can all of a sudden precede the auxiliary. We can account for both facts easily by assuming that *completely* attaches to vP but not to higher XPs. In non-elliptical structures, the position where the auxiliary occurs is the position of *have* in the tree. The auxiliary therefore necessarily precedes *completely*. If *completely* is to survive ellipsis at all, an XP must be elided that does not include *completely*, e.g. $AgrP$. Because of the bottom up process of morphological word formation *have* must now form a word not only with T° but also with v° : $\#v^\circ+have+T^\circ\#$. Inclusion of v° in this word is independently insured by the condition that abstract heads must be licensed through MWF.

This gives us an account of Sag's generalization.

¹⁷⁶ In Abels 2001a, 2003a I analyze somewhat more intricate patterns. The analysis there relies on a set of assumptions that is only partially compatible with the assumptions made here.

The counterexamples to Sag's generalization (negation – (374) and (376) and interrogatives – (378) and (379)) can be explained if we assume that T° is strong when its complement is $NegP$ and that C_Q° is strong.¹⁷⁷ The structure for interrogatives is shown in (390a). Where other theories posit HM, the current theories posits word formation, i.e. T-to-C movement is modeled as word formation of T with C . Since vP is elided in (379), X° must form a word upward. The resulting word $\#X^\circ+have+T^\circ+C^\circ\#$ correctly predicted to be pronounced in the highest strong position by (365), i.e. in C° .

- (390) a. $[_{CP} C_{Q,S} [_{TP} he T^\circ_W [_{haveP} \text{always} [_{haveW} [_{vP} \dots]]]]]$
 b. $[_{TP} he T^\circ_S [_{NegP} n't Neg^\circ_W [_{haveP} \text{always} [_{haveW} [_{vP} \dots]]]]]$

(390b) shows the structure of an example with negation, namely (374). Since vP is elided, X° must again form a word upwards with *have*, Neg° , and T° . By assumption Neg° renders T° strong, thus the resulting word $\#X^\circ_W+have_W+Neg^\circ_W+T^\circ_S\#$ is pronounced in T° .

The assumption that the strength or weakness of a head-position is not an inherent property of that head but is determined by its local context needs justification. Although this assumption is non-standard, it is not particularly strange. In fact for XP-chains it is the standard assumption. Whether a position in a *wh*-chain, for example, counts as weak

¹⁷⁷ The notation C_Q° might be misleading a little. I am not contradicting the position developed in chapter 2 here that C° enters the syntactic derivation with the features such as 'clause type' unspecified. Remember that I am assuming a single cycle model of the grammar and that the phonology/morphology has the entire hierarchical structure to work with. At the point where MWF applies, the relevant features will have

or strong is determined by the host not the *wh*-element itself: the attractor determines strength.¹⁷⁸

I assume generally that whether a head position counts as strong or weak is determined under MWF locally by an inherent property of the head of its complement. Thus if α° and β° form a word and β° is the head of the complement of α° , then β° determines whether α° counts as strong. This leaves the lowest head in a word without a value for strength. This is irrelevant though, as a glance at (365) reveals. Whether the lowest position in a word or a chain is strong or weak never matters. (365) asks whether positions higher than P are weak or strong. Whether P itself is weak or strong does not matter.

We now turn to pseudogapping. The pseudogapped version of (353) is *Fred surely has karate*. *Karate* can survive deletion only if vP and AgrP survive deletion, i.e. if only studyP is deleted. Since v° and Agr° are abstract, they need a licenser and word formation proceeds bottom up. The following word formation is then forced: $\#\text{Agr}^\circ + v^\circ_S + X^\circ_W + \text{have}_W + T^\circ_{W\#}$.¹⁷⁹ Together with the Positioning Algorithm (365) this gives the correct result: adverb > aux > object.

been specified and are available. There is then no problem in subscripting C° with Q for the purposes of this discussion.

¹⁷⁸ For additional discussion of this point cf. Abels 2000.

¹⁷⁹ In light of the preceding discussion, Agr° does not have a value for strength. Agr° makes the next higher head (v°) strong, hence, verb > object order. In fact, on slightly different assumptions Case adjacency can be made to follow (see Abels 2000).

The account presented here makes a further prediction. If the root of the verb is not elided but moved, low adverbs like *completely* should always move along and never stay behind. Higher adverbs should have more freedom. Both expectations are borne out as (391) and (392) show.

- (391) He said he would completely solve the problem, and (✓completely) solve the problem he (?*completely) did.
 (392) He said he would always love her, and (✓always) love her he (✓always) did.

Sag's generalization, the counterexamples to it, and even the problems for radical emptiness all fall under the present, simple theory.

5.4 Conclusion

Two main hypotheses make the account offered in this chapter possible: (i) Theorem (355) of HM is rejected; and (ii) MWF and Positioning are post-syntactic processes. Affix hopping was shown to be inadequate to the task of dealing with violation of the The-higher-the-bigger-Theorem. The idiosyncrasies of Affix Hopping are reinterpreted as morphological selection.

The fast argument against the existence of traces of movement and ellipsis Sag and Fodor 1995, i.e. radical emptiness, is dispelled. There is necessarily *some* abstract structure present at the ellipsis site. Most importantly, the present theory makes precise what it may mean for HM to be a PF-phenomenon Chomsky 2000 without duplicating syntactic movement operations in PF.

The emerging picture is pleasing. Head Movement is ruled out in the syntax, but possible in the phonology. This split allows me to derive the fact that downward

‘movement’ is possible under head movement but not, by the extension condition, in syntax. Sag’s generalization falls out as a near automatic consequence.

Chapter 6: Conclusion

The leitmotif of this thesis has been the Stranding Generalization, which I introduced as number (1) on the first page of chapter 1. But the basso continuo has been playing quite a different theme: How is long distance movement implemented in the grammar? Can we decide between theories that postulate uniform paths and those that postulate punctuated paths? And how?

This thesis begins to answer these questions by providing a two-pronged argument for the punctuated path hypothesis, i.e. for the hypothesis that long distance movement stops over in certain designated positions along the path and skips the remaining positions.

In chapter 2, section 2.1.3, I present the first part of the argument. I argue there that certain positions along the path of movement are not available as reconstruction sites. If the argument is correct and if more comparable cases are found, then long distance movement paths must be punctuated. Until further corroborating facts are discovered, this part of the argument remains somewhat programmatic.

The second line of argument comes from the Stranding Generalization: Phase heads cannot be stranded by their complements. We know independently (from cyclicity effects such as complementizer agreement in Irish and other languages, Q-float in West Ulster English, floated focus particles in Dutch, etc.) that certain positions are prone to exhibiting cyclicity effects and other are not. The ones that show cyclicity effects have also proven useful in defining locality domains: phases and their heads. If the Stranding

Generalization indeed proves to be correct, then we have a third kind of effect involving phases. Such clustering of properties cannot be accidental. In chapter 2 I therefore construct a theory that links up all three properties of phases, reducing them to a common core: Movement out of a phase must pass through the specifier or an adjoined position of that phase. Within such a theory, the Stranding Generalization becomes an argument for the punctuated path hypothesis. Without it, the clustering of properties remains mysterious.

Much work remains to be done. Further cases of lacking reconstruction sites ought to be sought. It should be shown carefully that cyclicity effects and reconstruction actually cluster around the same positions. Cyclicity effects should be demonstrated with adpositions. The diagnostics distinguishing traces of movement from resumptive pronouns should be strengthened....

The least I hope to have achieved is to point out certain empirical differences between theories of punctuated paths and theories of uniform paths. The question how long distance movement is implemented in the grammar is important enough to receive a well-reasoned empirical answer.

References

- Abels, Klaus. 2000. Move? Ms. University of Connecticut, Storrs.
 —. 2001a. On a modular view of Head Movement. Ms. University of Connecticut, Storrs.
 —. 2001b. The Predicate Cleft Construction in Russian. In *Formal Approaches to Slavic Linguistics 9*, eds. Steven Franks and Michael Yadroff, 1-19. Bloomington, IN: Michigan Slavic Publications.
 —. 2002a. 'Expletive Negation' in Russian: A Conspiracy Theory. Ms. Storrs, CT
 —. 2002b. On an Alleged Argument for the Proper Binding Condition. In *Proceedings of HUMIT 2001. MIT Working Papers in Linguistics 43*, eds. Tania R. Ionin, Heejeong Ko and Andrew Nevins, 1-16. Cambridge, MA.
 —. 2003a. Auxiliary Adverb Word Order Revisited. In *Proceedings of the 26th UPenn Colloquium*.
 —. 2003b. A note on clitics and P-stranding. In *Proceedings of FASL 11*. Ann Arbor: Michigan Slavic Publishers.
 —. in press. *[P clitic]! - Why? In *Proceedings of Formal Description of Slavic Languages 4*. Potsdam, Germany.
 Abney, Steve. 1987. The English noun phrase in its sentential aspect. Doctoral dissertation. MIT, Cambridge, MA.
 Abraham, Werner. 1995. *Deutsche Syntax im Sprachvergleich. Grundlegung einer typologischen Syntax des Deutschen*: Studien zur deutschen Grammatik 41. Tübingen: Gunter Narr Verlag.
 Åfarli, T. 1985. Norwegian Verb Particle Constructions as Causative Constructions. *Nordic Journal of Linguistics* 8:75-98.
 Alexandre, Nélia, and Tjerk Hagemeijer. 2002. Pronomes Resumptivos e Abandono de Preposição nos Crioulos Atlânticos de Base Portuguesa. *Working Papers of the Portuguese Linguistics Society*.
 Alexiadou, Artemis. 1997. *Adverb Placement: A case study in antisymmetric syntax*. vol. 18: *Linguistik Aktuell*. Amsterdam, Philadelphia: John Benjamins.
 Bach, Emmon, and George M. Horn. 1976. Remarks on 'Conditions on Transformations'. *Linguistic Inquiry* 7:265-299.
 Baker, Carl Leroy. 1971. Stress level and auxiliary behavior in English. *Linguistic Inquiry* 2:167-181.
 —. 1981. Auxiliary-Adverb Word Order. *Linguistic Inquiry* 12:309-315.
 Baker, Mark. 1985. The Mirror Principle. *Linguistic Inquiry* 16:373.
 —. 1988. *Incorporation: A theory of grammatical function*. Chicago: University of Chicago press.
 Baltin, Mark, and Paul Postal. 1996. More on Reanalysis Hypotheses. *Linguistic Inquiry* 27:127-145.
 Barbiers, Sjef. 1995. The Syntax of Interpretation. Ph. D. Dissertation. Rijksuniversiteit Leiden, Leiden, NL.

- Barss, Andrew. 1986. Chains and anaphoric dependence: On reconstruction and its implications. Doctoral dissertation. MIT, Cambridge, MA.
- Bayer, Josef. 1990. Interpretive Islands. In *Scrambling and Barriers*, eds. Günther Grewendorf and Wolfgang Sternefeld, 341-421. Amsterdam: Benjamins.
- Beck, Sigrid. 1996. Quantified structures as barriers for LF movement. *Natural Language Semantics* 4:1-56.
- . 2000. The Semantics of *different*: Comparison Operator and Relational Adjective. *Linguistics and Philosophy* 23:101-139.
- Beck, Sigrid, Toshiko Oda, and Koji Sugisaki. 2002. Comparative Constructions in Japanese vs. English. Paper presented at *NELS 33*, Cambridge, MA.
- Behagel, Otto. 1932. *Deutsche Syntax. Band 4: Wortstellung. Periodenbau*. Heidelberg: Carl Winters Universitätsbuchhandlung.
- Belletti, Adriana. 1988. The Case of unaccusatives. *Linguistic Inquiry* 19:1-34.
- Bennis, Hans, Norbert Corver, and Marcel den Dikken. 1997. Predication in Nominal Phrases. *Journal of Comparative Germanic Linguistics* 1:85-117.
- Bennis, Hans, and Teun Hoekstra. 1984. Gaps and Parasitic Gaps. *The Linguistic Review* 4:29-87.
- Besten, Hans den. 1978. On the Presence and Absence of *Wh*-Elements in Dutch Comparatives. *Linguistic Inquiry* 9:641-671.
- . 1981. Marking *Wh*-Movement in Afrikaans. In *Generative Studies in Creole Languages*, ed. Pieter Muysken, 141-179. Dordrecht: Foris.
- . 1985. The ergative hypothesis and free word order in Dutch and German. In *Studies in German Grammar*, ed. Jindra Toman, 23-65. Dordrecht: Foris.
- Besten, Hans den, and Gert Webelhuth. 1990. Stranding. In *Scrambling and Barriers*, eds. Günther Grewendorf and Wolfgang Sternefeld, 77-92. Amsterdam and Philadelphia: Academic Press.
- Birmingham, J.C. 1970. The Papiamentu language of Curaçao. Ph.D. thesis. University of Virginia.
- Blatz, Friedrich. 1896. *Neuhochdeutsche Grammatik mit Berücksichtigung der historischen Entwicklung der Deutschen Sprache*. vol. 2. Karlsruhe: Lang.
- Blevins, James P. 1990. Syntactic Complexity: Evidence for Discontinuity and Multidomination. Ph.D. dissertation. University of Massachusetts, Amherst, MA.
- Bobaljik, Jonathan D. 1994. What does adjacency do? In *MIT Working Papers in Linguistics*, 1-32. Cambridge, MA: MITWPL, Department of Linguistics and Philosophy, MIT.
- . 1995a. In terms of Merge: Copy and head movement. In *Papers on minimalist syntax, MIT working papers in linguistics*, eds. Rob Pensalfini and Hiroyuki Ura, 41-64. Cambridge, MA: MITWPL, Department of Linguistics and Philosophy, MIT.
- . 1995b. Morphosyntax: The syntax of verbal inflection. Doctoral dissertation. MIT, Cambridge, MA.
- . 1999. Adverbs: the hierarchy paradox. *GLOT International* 4:27-28.
- Boeckx, Cedric. 1999. Conflicting c-command requirements. *Studia Linguistica* 53:227-250.

- . 2001. Mechanisms of Chain Formation. Doctoral Dissertation. University of Connecticut, Storrs, CT.
- . 2002. Agree or Attract? A Relativized Minimality solution to a Proper Binding Condition puzzle. In *Theoretical Approaches to Universals*, ed. Artemis Alexiadou, 41-64. Amsterdam: John Benjamins.
- Boeckx, Cedric, and Sandra Stjepanović. 2001. Heading Toward PF. *Linguistic Inquiry* 31:345-355.
- Borsley, Robert D. 1986. Prepositional complementizers in Welsh. *Journal of Linguistics* 22:67-84.
- . 2001. What do 'Prepositional Complementizers' do? *Probus* 13:155-171.
- Borsley, Robert D, and Ewa Jaworska. 1988. A note on prepositions and Case marking in Polish. *Linguistic Inquiry* 19:685-691.
- Bošković, Željko. 1994a. D-structure, theta-criterion, and movement into theta-positions. *Linguistic Analysis* 24:247-286.
- . 1994b. Participle movement in Serbo-Croatian and related issues. Storrs: University of Connecticut.
- . 1995. Participle movement and second position cliticization in Serbo-Croatian. *Lingua* 96:245-266.
- . 1997a. On certain violations of the superiority, AgrO, and economy of derivation. *Journal of Linguistics* 33:227-254.
- . 1997b. *The syntax of nonfinite complementation: An economy approach*. Cambridge, Mass.: MIT Press.
- . 2001a. Floating Quantifiers, THETA-Role Assignment, and Economy. In *Proceedings of NELS 31*, eds. Minjoo Kim and Uri Strauss. Amherst, MA.: GLSA.
- . 2001b. *On the Nature of the Syntax-Phonology Interface: Cliticization in South Slavic*. London: Elsevier North-Holland.
- . 2002a. Be careful where you float your quantifiers. Ms. University of Connecticut, Storrs.
- . 2002b. A-movement and the EPP. *Syntax* 5:167-218.
- . in press. On left branch extraction. In *Proceedings of the 4th European Conference on Formal Description of Slavic Languages*.
- . no year. On Right Node Base Generation. Ms. University of Connecticut, Storrs, CT.
- Bošković, Željko, and Howard Lasnik. 2002. On the Distribution of Null C. Ms. University of Connecticut, Storrs and University of Maryland, College Park
- Bošković, Željko, and Daiko Takahashi. 1998. Scrambling and last resort. *Linguistic Inquiry* 29:347-366.
- Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24:591-656.
- Bresnan, Joan W. 1974. The position of certain clause-particle in phrase structure. *Linguistic Inquiry* 5.
- . 1977. Variables in the theory of transformations. In *Formal syntax*, eds. Peter Culicover, Thomas Wasow and Adrian Akmajian, 157-196. New York: Academic Press.

- . 1978. A Realistic Transformational Grammar. In *Linguistic Theory and Psychological Reality*, eds. Moris Halle, G. A. Miller and Joan W. Bresnan. Cambridge, Mass.: MIT Press.
- Bresnan, Joan W., and Jonni M. Kanerva. 1992. The thematic hierarchy and locative inversion in UG. In *Syntax and Semantics 26: Syntax and Lexicon*, eds. Eric Wehrli and Tim Stowell, 53-101. New York: Academic Press.
- Brody, Michael. 1995. *Lexico-Logical Form. A radically minimalist approach*. Cambridge, MA: MIT Press.
- . 1997. Mirror Theory. In *UCL Working Papers in Linguistics 9*, 179-223. London: Department of Phonetics and Linguistics, University College London.
- . 2000. Mirror Theory. Syntactic Representation in Perfect Syntax. *Linguistic Inquiry* 31:29-57.
- Bruening, Benjamin. 2001. Syntax at the Edge: Cross-Clausal Phenomena and The Syntax of Passamaquoddy. Doctoral dissertation. MIT, Cambridge, MA.
- Campos, Héctor. 1991. Preposition Stranding in Spanish? *Linguistic Inquiry* 22:741-750.
- Cardinaletti, Anna, and Michal Starke. 1999. The Typology of Structural Deficiency. In *Clitics in the Languages of Europe*, ed. Henk van Riemsdijk, 145-234. Berlin and New York: Mouton de Gruyter.
- Castillo, J. C., John Drury, and Kleanthes K. Grohmann. 1999. The status of Merge over Move preference. Ms. University of Maryland, College Park.
- Cattell, Ray. 1976. Constraints on movement rules. *Language* 52:18-50.
- Čavar, Damir, and Chris Wilder. 1994. Clitic third in Croatian. In *Eurotyp Working Papers, Grammatical Models Selection*, 19-61.
- Chametzky, Robert. 1996. *A theory of phrase markers and the extended base*. Albany, NY: State University of New York Press.
- . 2000. *Phrase Structure: From GB to Minimalism*. Malden, MA: Blackwell.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
- . 1970. Remarks on nominalization. In *Readings in English transformational grammar*, eds. Roderick A. Jacobs and Peter S. Rosenbaum, 184-221. Waltham, MA: Ginn and Co.
- . 1973. Conditions on transformations. In *Festschrift for Morris Halle*, eds. Stephen R. Anderson and Paul Kiparsky, 232-286. New York: Holt, Rinehart and Winston.
- . 1977. On *wh*-movement. In *Formal syntax*, eds. Peter Culicover, Thomas Wasow and Adrian Akmajian, 71-132. New York: Academic Press.
- . 1981. *Lectures on government and binding*. Dordrecht: Foris.
- . 1986a. *Barriers*. vol. 13: Linguistic Inquiry Monograph. Cambridge, MA: MIT Press.
- . 1986b. *Knowledge of language*. New York: Praeger.
- . 1993. A minimalist program for linguistic theory. In *The view from Building 20*, eds. Kenneth Hale and Samuel J. Keyser, 1-52. Cambridge, MA: MIT Press.
- . 1995a. Bare phrase structure. In *Government and binding theory and the minimalist program*, ed. Gert Webelhuth, 383-439. Oxford: Basil Blackwell.
- . 1995b. Categories and transformations. In *The minimalist program*, 219-394. Cambridge, MA: MIT Press.
- . 1995c. *The minimalist program*. Cambridge, MA: MIT Press.

- . 2000. Minimalist inquiries: The framework. In *Step by Step: Essays on minimalism in honor of Howard Lasnik*, eds. Roger Martin, David Michaels and Juan Uriagereka, 89-155. Cambridge, MA: MIT Press.
- . 2001a. Beyond Explanatory Adequacy. Ms. MIT, Cambridge, MA.
- . 2001b. Derivation by Phase. In *Ken Hale: a Life in Language*, ed. Michael Kenstowicz, 1-52. Cambridge, MA: MIT Press.
- Chomsky, Noam, and Howard Lasnik. 1993. The theory of principles and parameters. In *Syntax: An international handbook of contemporary research*, eds. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld and Theo Vennemann, 506-569. Berlin: Walter de Gruyter.
- Chung, Sandra. 1998. *The Design of Agreement: Evidence from Chamorro*. Chicago, IL: University of Chicago Press.
- Chung, Sandra, William Ladusaw, and James McCloskey. 1995. Sluicing and Logical Form. *Natural Language Semantics* 3:239-282.
- Cinque, Guglielmo. 1990. *Types of A'-dependencies*. Cambridge, MA: MIT Press.
- . 1999. *Adverbs and Functional Heads - A Cross-Linguistic Perspective*. New York and Oxford: Oxford University Press.
- Citko, Barbara. 2002. ATB questions and the nature of merge. Paper presented at *NELS* 33, Cambridge, Mass.
- Collins, Chris. 1993. Topics in Ewe syntax. Doctoral dissertation. MIT, Cambridge, MA.
- . 1994. Economy of derivation and the generalized proper binding condition. *Linguistic Inquiry* 25:45-61.
- . 1997. *Local economy*. Cambridge, Mass.: MIT Press.
- . 2002. Eliminating Labels. In *Derivation and Explanation in the Minimalist Program*, eds. Samuel David Epstein and T. David Seely, ??? Malden, MA and Oxford UK: Blackwell.
- Conway, Laura. 1996. NPs in disguise. Ms. University of Connecticut, Storrs.
- Cormack, Annabel, and Neil Smith. 1998. Negation, polarity and V positions in English. *UCL Working Papers in Linguistics* 10.
- Corver, Norbert. 1990. The Syntax of Left Branch Extractions.
- . 1991. The internal syntax and movement behavior of the Dutch 'wat voor'-construction. *Linguistische Berichte* 133:190-228.
- Couper-Kuhlen, Elizabeth. 1979. *The prepositional passive in English: A semantic-syntactic analysis, with a lexicon of prepositional verbs*. Tübingen: Max Niemeyer Verlag.
- Craig, Colette Grinevald. 1977. *The Structure of Jacalteco*. Austin and London: University of Texas Press.
- Cruz, Juan de la, and Emilio Saameno. 1996. The Middle English Prepositional Passive: Analogy and GB. *Neuphilologische Mitteilungen* 97:169-186.
- Culicover, Peter W. 1999. *Syntactic Nuts*. Oxford: Oxford University Press.
- Davis, A. 1984. Behind the for-to filter: for-to infinitives in Belfast English and the theory of government. *Scheffeld Working Papers in Linguistics* 1:56-71.
- Dekydtspotter, Laurent. 1992. The syntax of predicate clefts. Paper presented at *Proceedings of the North Eastern Linguistic Society*, Amherst.

- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA.: MIT Press.
- Dijkhoff, Marta B. 1983. Movement Rules and the Resumptive Pronoun Strategy in Papiamenty. Unpublished M.A. thesis. Rijksuniversiteit Groningen.
- Dikken, Marcel den. 1992. Particles. Ph.D. Thesis. Leiden.
- . 1995. Binding, expletives, and levels. *Linguistic Inquiry* 26:347-354.
- Donaldson, Anne. 1982. On Restructuring and Preposition Stranding. *Cahiers linguistiques d'Ottawa* 10:81-100.
- Doron, Edith. 1982. On the syntax and semantics of resumptive pronouns. *Texas Linguistic Forum*:1-48.
- Duarte, Yara. 1994. Constructions with Unaccompanied Prepositions in English. *Revista de Documentacao de Estudos em Linguistica Teorica e Aplicada* 10:409-420.
- Embick, David, and Rolf Noyer. 2001. Movement Operations after Syntax. *Linguistic Inquiry* 32:555-595.
- Emonds, Joseph. 1972. Evidence that indirect object movement is a structure-preserving rule. *Foundations of Language* 8:546-561.
- . 1978. The verbal complex V'-V" in French. *Linguistic Inquiry* 9:151-175.
- . 1985. *A Unified Theory of Syntactic Categories*. Dordrecht: Foris.
- Epstein, Samuel David. 1999. Un-principled syntax and the derivation of syntactic relations. In *Working Minimalism*, eds. Samuel David Epstein and Norbert Hornstein, 317-345. Cambridge, MA: MIT Press.
- . 2001. Deriving the Proper Binding Condition. Paper presented at *TiLT*, Utrecht, NL.
- Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima, and Hisatsugu Kitahara. 1998. *A derivational approach to syntactic relations*. Oxford: Oxford University Press.
- Epstein, Samuel David, and T. Daniel Seely. 1999. SPEC-ifying the GF "subject"; eliminating A-Chains and the EPP within a Derivational Model. Ms. Univ. of Michigan
- . 2002. Rule Applications as Cycles in a Level-Free Syntax. In *Derivation and Explanation in the Minimalist Program*, eds. Samuel David Epstein and T. Daniel Seely, 65-89. Malden, Mass.: Blackwell.
- Ernst, Thomas. 1983. More on adverbs and stressed auxiliaries. *Linguistic Inquiry* 14:542-549.
- . 2002. *The Syntax of Adverbs*. Cambridge, UK: Cambridge University Press.
- Fanselow, Gisbert. 1983. Zu einigen Problemen von Kasus, Rektion und Bindung in der deutschen Syntax. M.A. Thesis. Universität Konstanz.
- . 1991. Minimale Syntax. Habilitationsschrift. Universität Passau.
- Fiengo, Robert. 1974. Semantic conditions on surface structure. Doctoral dissertation. MIT, Cambridge, MA.
- . 1977. On trace theory. *Linguistic Inquiry* 8:35-62.
- Fiengo, Robert, and Robert May. 1994. *Indices and identity*. Cambridge, MA.: MIT Press.
- Fillmore, Charles J. 1965. *Indirect Object Constructions in English and the Ordering of Transformations*. The Hague: Mouton.

- Fleischer, Jürg. 2002. Preposition Stranding in German Dialects. In *Syntactic Microvariation*, eds. Sjef Barbiers, Leonie Cornips and Susanne van der Kleij. Amsterdam: Meertens Institute.
- Fox, Danny, and Howard Lasnik. 2003. Successive cyclic movement and island repair: The difference between Sluicing and VP Ellipsis. *Linguistic Inquiry* 34:143-154.
- Fox, Danny, and David Pesetsky. 2003. Cyclic Linearization and the Typology of Movement. Paper presented at *University of Massachusetts*.
- Frampton, John. 1999. The Fine Structure of *Wh*-Movement and the Proper Formulation of the ECP. *The Linguistic Review* 16:43-61w.
- Frampton, John, and Sam Gutmann. 1997. Cyclic computation, a computationally efficient minimalist syntax: Northeastern University.
- . 2002. Crash-proof Syntax. In *Derivation and Explanation in the Minimalist Program*, eds. Samuel David Epstein and T. Daniel Seely, 90-105. Malden, MA and Oxford, UK: Blackwell.
- Frank, Robert. 2002. *Phrase Structure Composition and Syntactic Dependencies*: Current Studies in Linguistic Theory. Cambridge, Mass.: MIT Press.
- Franks, Steven. 1998. Clitics in Slavic. Paper presented at *Comparative Slavic Morphosyntax Workshop*, Bloomington, ID.
<http://www.indiana.edu/~slavconf/linguistics.index.html>.
- Franks, Steven, and Tracy Holloway King. 2000. *A handbook of Slavic clitics*. New York: Oxford University Press.
- Franks, Steven, and Ljiljana Progovac. 1994. On the placement of Serbo-Croatian clitics. Paper presented at *Proceedings of the 9th Biennial Conference on Balkan and South Slavic Linguistics, Literature, and Folklore (=Indiana Linguistic Studies 7)*, Bloomington.
- Gärtner, Hans Martin. 1999. Phrase-Linking Meets Minimalist Syntax. *Proceedings of the West Coast Conference on Formal Linguistics* 18:159-169.
- . 2002. *Generalized Transformations and Beyond: Reflections on Minimalist Syntax*. Berlin: Akademie Verlag.
- Gawron, Jean Mark, and Andrew Kehler. to appear. The semantics of respective readings, conjunction, and filler gap dependencies. *Linguistics and Philosophy*.
- Goodall, Grant. 1987. *Parallel Structures in Syntax*. Cambridge: Cambridge University Press.
- Grewendorf, Günther. 1990. *Ergativity in German*. Dordrecht: Foris.
- Grimshaw, Jane. 1991. Extended Projection. Ms. Waltham, MA. Brandeis University
- . 2000. Locality and Extended Projection. In *Lexical Specification and Insertion*, eds. Peter Coopmans, Martin Everaert and Jane Grimshaw, 115-135. Amsterdam: John Benjamins.
- Groat, Erich, and John O'Neil. 1996. Spell-Out at the LF Interface. In *Minimal Ideas: Syntactic studies in the Minimalist framework*, eds. Samuel David Epstein, Höskuldur Thráinsson and Jan-Wouter C. Zwart, 113-139. Amsterdam: John Benjamins.
- Grohmann, Kleanthes K. 2000. Proliferous Peripheries: A Radical View from the Left. Doctoral Dissertation. University of Maryland, College Park.

- Hagemteijer, Tjerk. 2000. Serial Verb Constructions in São-Tomense. MA Thesis. University of Lisbon, Lisbon.
- Harlow, Stephen. 1992. Finiteness and Welsh sentence structure. In *Structure de la Phrase et Théorie du Liage*, eds. Hans-Georg Obenauer and Anne Zribi-Hertz, 93-119. Paris: Presses Universitaires Vincennes.
- Hauge, Kjetil Rå. 1999. *A Short Grammar of Contemporary Bulgarian*. Bloomington, ID: Slavica Publishers.
- Hellan, Lars, and Christer Platzack. 1999. Pronouns in Scandinavian Languages: An Overview. In *Clitics in the Languages of Europe*, ed. Henk van Riemsdijk, 123-142. Berlin: Mouton deGruyter.
- Hendrick, Randall. 1988. *Anaphora in Celtic and Universal Grammar*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Herslund, Michael. 1984. Particles, prefixes and preposition stranding. In *Topics in Danish syntax*, eds. Finn Sørensen and Lars Heltoft, 34-71. Copenhagen: Akademisk Forlag.
- Hiraiwa, Ken. 2002. Raising and Indeterminate-Agreement. Ms. MIT, Cambridge, Mass.
- Hoekstra, Jarich. 1991. Oer it beklamjen fan ferhâldingswurden yn it Frysk, it Hollânsk en it Ingelsk. *Us Wurk* 40:67-103.
- . 1995. Preposition Stranding and Resumptivity in West Germanic. In *Studies in Comparative Germanic Syntax*, eds. Hubert Haider, Susan Olsen and Sten Vikner, 95-118: Kluwer Academic Publishers.
- Holmberg, Anders. 1986. Word order and syntactic features in the Scandinavian languages and English. Ph.D. Dissertation. University of Stockholm, Stockholm.
- . 1999. Remarks on Holmberg's generalization. *Studia Linguistica* 53:1-39.
- Holmberg, Anders, and Thorbjörg Hróarsdóttir. 2003. Agreement and movement in Icelandic raising constructions. *Lingua* 113:997-1019.
- Hornstein, Norbert. 1998. Movement and Chains. *Syntax* 1:99-127.
- . 1999. Movement and Control. *Linguistic Inquiry* 30:69-96.
- Hornstein, Norbert, and Jairo Nunes. 1999. On asymmetries between parasitic gap and across-the-board constructions. Ms. University of Maryland and College Park and Universidade Estadual de Campinas
- Hornstein, Norbert, and Amy Weinberg. 1981. Case theory and preposition stranding. *Linguistic Inquiry* 12:55-91.
- Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation. MIT, Cambridge, MA.
- . 1993. Reconstruction and the structure of VP: Some theoretical consequences. *Linguistic Inquiry* 24:103-138.
- . 1997. On lexical structure and syntactic projection. *Chinese Languages and Linguistics* 3:45-89.
- Hudson, Richard. 1976. Conjunction reduction, gapping, and right node raising. *Language* 52:535-562.
- Inada, Toshiaki. 1981. Problems of reanalysis and preposition stranding. *Studies in English Linguistics* 9:120-131.

- Ishii, Toru. 1997. An asymmetry in the composition of phrase structure and its consequences. Doctoral dissertation. University of California, Irvine.
- . 1998. Derivational selectional restriction and reconstruction. *English Linguistics* 15:28-49.
- . 1999. Cyclic Spell-Out and the *That-t* Effects. In *WCCFL 18 Proceedings*, eds. S Bird, Andrew Carnie, J Haugen and P Norquest, 220-231. Somerville, MA: Cascadia Press.
- Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA.: MIT Press.
- . 1973. The Base Rules for Prepositional Phrases. In *A Festschrift for Morris Halle*, eds. Stephen R Anderson and Paul Kiparsky, 345-356. New York: Holt, Rinehart and Winston.
- Jayaseelan, Karattuparambil A. 1990. Incomplete VP deletion and gapping. *Linguistic Analysis* 20:64-81.
- Johnson, David, and Shalom Lappin. 1997. A Critique of the Minimalist Program. *Linguistics and Philosophy* 20:273-333.
- Johnson, Kyle. 1996. When Verb Phrases go missing. *GLOT International* 2:2-9.
- Joshi, Avarind K., Leon Levi, and Masako Takahashi. 1975. Tree adjunct grammars. *Journal of the Computer and System Sciences* 10:136-163.
- Kayne, Richard. 1981. ECP extensions. *Linguistic Inquiry* 12:93-133.
- . 1983. Connectedness. *Linguistic Inquiry* 14:223-249.
- . 1984. *Connectedness and binary branching*. Dordrecht: Foris.
- . 1994. *The antisymmetry of syntax*. Cambridge, MA.: MIT Press.
- . 1998. Overt vs. covert movement. *Syntax* 1:128-191.
- . 1999. Prepositional complementizers as attractors. *Probus* 11:39-73.
- Kennedy, Christopher. 2002. Comparative Deletion and Optimality in Syntax. *Natural Language and Linguistic Theory* 20:553-621.
- Kennedy, Christopher, and Jason Merchant. 2000. Attributive comparative deletion. *Natural Language and Linguistic Theory* 18:89-146.
- King, Ruth, and Yves Roberge. 1990. Preposition Stranding in Prince Edward Island French. *Probus* 3:351-369.
- King, Tracy Holloway. 1996. Slavic clitics, long head movement, and prosodic inversion. *Journal of Slavic Linguistics* 4:274-311.
- Kitahara, Hisatsugu. 1997. *Elementary operations and optimal derivations*. Cambridge, MA: MIT Press.
- Koizumi, Masatoshi. 1995. Phrase structure in minimalist syntax. Doctoral dissertation. MIT, Cambridge, MA.
- Koopman, Hilda. 1984. *The syntax of verbs: From verb movement rules in the Kru languages to Universal Grammar*. Studies in Generative Grammar; 15. Dordrecht: Foris.
- . 1997. Prepositions, Postpositions, Circumpositions and Particles: The Structure of Dutch PPs. Ms. UCLA, Los Angeles, CA.C:\My Documents\Aufsaetze von anderen\Koopman Dutch PPs.pdf

- Koot, Hans van de, and Ad Neeleman. 2002. The Configurational Matrix. *Linguistic Inquiry* 33:529-574.
- Koster, Jan. 1986. *Domains and Dynasties: The radical autonomy of syntax*. Studies in Generative Grammar 30. Dordrecht: Foris.
- Kratzer, Angelika. 1996. Severing the verb from its external argument. In *Phrase Structure and the Lexicon*, eds. Johan Rooryck and Laurie Zaring, 109-137. Dordrecht: Kluwer Academic Publishers.
- Kroch, Anthony. 1989. Asymmetries in long distance extraction in a tree adjoining grammar. In *Alternative conceptions of phrase structure*, eds. Mark Baltin and Anthony Kroch. Chicago, IL: University of Chicago Press.
- Kroch, Anthony, and Aravind Krishna Joshi. 1985. The linguistic relevance of Tree Adjoining Grammar. Ms. Philadelphia, Pennsylvania
- Kuno, Masakazu. 2003. A New Perspective on "How-Likely" Puzzle. Ms. Harvard University, Cambridge, MA
- Länzlinger, C. and Ur Shlonsky. 1997. Pronoun Clustering and Adjacency Effects: A Comparative Study of German and Hebrew Weak Pronouns. *Studia Linguistica* 51:154-185.
- Lasnik, Howard. 1995a. Case and expletives revisited: On Greed and other human failings. *Linguistic Inquiry* 26:615-633.
- . 1995b. Last resort. In *Minimalism and linguistic theory*, eds. Shosuke Haraguchi and Michio Funaki, 1-32. Tokyo: Hituzi Syobo.
- . 1995c. A note on pseudogapping. In *Papers on minimalist syntax, MIT working papers in linguistics*, eds. Rob Pensalfini and Hiroyuki Ura, 143-163. Cambridge, Mass.: MITWPL, Department of Linguistics and Philosophy, MIT.
- . 1999a. Chains of arguments. In *Working minimalism*, eds. Samuel D. Epstein and Norbert Hornstein, 189-215. Cambridge, MA.: MIT Press.
- . 1999b. Pseudogapping puzzles. In *Fragments: Studies in ellipsis*, eds. Elabbas Benmamoun, Hiroto Hoshi and Shalom Lappin. Oxford: Oxford University Press.
- . 2001. When can you save a structure by destroying it? In *Proceedings of NELS 31*, eds. Minjoo Kim and Uri Strauss, 301-320. Amherst, MA.: GLSA.
- . 2002. On Clause-Mate Conditions. Paper presented at *University of Connecticut*, May 14, 2002.
- . in press-a. Feature Movement or Agreement at a Distance? In *Remnant movement, F-movement, and the T-model*, eds. Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers and Hans Martin Gärtner. Amsterdam: John Benjamins.
- . in press-b. Patterns of verb raising with auxiliary 'be'. Paper presented at *Proceedings of the 1995 UMass Conference on African American English*, Amherst.
- Lasnik, Howard, and Mamoru Saito. 1991. On the subject of infinitives. Paper presented at *Papers from the 27th regional Meeting of the Chicago Linguistic Society Part One: The General Session*, Chicago, Ill.
- . 1992. *Move α : Conditions on its application and output*. Cambridge, MA: MIT Press.
- Law, Paul. 1998. A unified Analysis of P-stranding in Romance and Germanic. In *Proceedings of NELS 28*, eds. Pius N. Tamanji and Kiyomi Kusumoto. Amherst, MA: GLSA.

- Lebeaux, David. 1995. Where does the Binding Theory apply? In *Maryland Working Papers in Linguistics*, 63-88. College Park, MD: University of Maryland.
- Legate, Julie Anne. 2003. Some Interface Properties of the Phase. *Linguistic Inquiry* 34:506-515.
- Lenerz, Jürgen. 1995. Klammerkonstruktionen. In *Syntax, Vol. II*, eds. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld and Theo Vennemann, 1266-1276. Berlin: Mouton de Gruyter.
- Leu, Thomas. 2003. What moves where and *what for*. An account of Swiss German *was für*. Ms. New York University
- Levin, Beth, and Malka Rappaport Hovav. 1995. *Unaccusativity: At the Syntax - Lexical Semantics Interface*. Linguistic Inquiry Monograph Twenty-Six. Cambridge, MA: MIT Press.
- Levin, Nancy. 1978. Some identity-of-sense deletions puzzle me. Do they you. Paper presented at *Proceedings of the Fourteenth Annual Meeting of the Chicago Linguistic Society*, Chicago, Ill.
- . 1979/1986. Main verb ellipsis in spoken English. Doctoral dissertation. Ohio State University, Columbus.
- Levine, Robert D. 1984. Against Reanalysis Rules. *Linguistic Analysis* 14:3-30.
- . 1985. Right node (non-)raising. *Linguistic Inquiry* 14.
- Lunt, Horace G. 1959. *Old Church Slavonic Grammar*. Slavistic Printings and Reprintings. 'S-Gravenhage: Mouton & Co.
- Mahajan, Anoop. 1990. The A/A-bar distinction and movement theory. Doctoral dissertation. MIT, Cambridge, MA.
- Maling, Joan. 1972. On gapping and the order of constituents. *Linguistic Inquiry* 3:101-108.
- Maling, Joan, and Annie Zaenen. 1978. The Nonuniversality of a Surface Filter. *Linguistic Inquiry* 9:475-497.
- . 1985a. Preposition-Stranding and Oblique Case. *Cornell Working Papers in Linguistics* 7:149-161.
- . 1985b. Preposition-Stranding and Passive. *Nordic Journal of Linguistics* 8:197-209.
- . 1990. Preposition-stranding and passive. In *Syntax and Semantics 24: Modern Icelandic syntax*, eds. Joan Maling and Annie Zaenen. New York: Academic Press.
- Manzini, Maria Rita. 1992. *Locality: A theory and Some of Its Empirical Consequences*. Linguistic Inquiry Monograph 19. Cambridge, MA: MIT Press.
- Marantz, Alec. 2001. Words. Paper presented at *WCCFL XX*, USC, LA.
- Martin, Roger. 1992. Case theory, A-chains, and expletive replacement. Storrs: University of Connecticut.
- . 1996. A minimalist theory of PRO and control. Doctoral dissertation. University of Connecticut, Storrs.
- Matushansky, Ora. 2002. Head-movement in Linguistic Theory. Ms. MIT, Cambridge, MA.
- . 2003. Going through a phase. Paper presented at *ISP Workshop: EPP and Phases*, MIT, Cambridge, Mass.

- May, Robert. 1985. *Logical Form: Its structure and derivation*. Cambridge, MA: MIT Press.
- McCawley, James D. 1987. Some additional evidence of discontinuity. In *Syntax and Semantics Vol. 20: Discontinuous Constituency*, eds. Geoffrey Huck and Almerindo E. Ojeda. Chicago: University of Chicago Press.
- McCloskey, James. 1979. *Transformational Syntax and Model Theoretic Semantics: A Case Study in Modern Irish*. Synthese Language Library. Dordrecht, Boston, and London: D. Reidel Publishing Company.
- . 1986a. Inflection and Conjunction in Modern Irish. *Natural Language and Linguistic Theory* 4:245-281.
- . 1986b. Right Node Raising and Preposition Stranding. *Linguistic Inquiry* 17:183-186.
- . 1990. Resumptive Pronouns, A'-Binding, and Levels of Representation in Irish. In *The Syntax of the Modern Celtic Languages. Syntax and Semantics 23*, ed. Randall Hendrick, 199-248. San Diego: Academic Press.
- . 2000. Quantifier Float and *Wh*-Movement in an Irish English. *Linguistic Inquiry* 31:57-84.
- McCloskey, James, and Ken Hale. 1984. On the syntax of person number marking in Modern Irish. *Natural Language and Linguistic Theory* 1:487-533.
- McGinnis, Martha. 1998. Reflexive External Arguments and Lethal Ambiguity. In *Proceedings of WCCFL XVI*, eds. Emily Curtis, James Lyle and Gabriel Webster, 303-317. Stanford: CSLI Publications.
- Mehlhorn, Grit. 2001. Kontrastierte Konstituenten im Russischen. Experimentelle Untersuchungen zur Informationsstruktur. Universität Leipzig.
- Mensching, Guido. 2000. *Infinitive Constructions with Specified Subjects: A Syntactic Analysis of the Romance Languages*. Oxford Studies in Comparative Syntax. Oxford: Oxford University Press.
- Merchant, Jason. 1999. The Syntax of Silence - Sluicing, Islands, and Identity of Ellipsis. Ph.D. dissertation. UCSC, Santa Cruz, CA.
- . to appear. Variable Island Repair under Ellipsis. In *Topics in Ellipsis*, ed. Kyle Johnson. Cambridge: Cambridge University Press.
- Moro, Andrea. 2000. *Dynamic Antisymmetry*. Linguistic Inquiry Monograph 38. Cambridge, MA: MIT Press.
- Müller, Gereon. 1998. *Incomplete Category Fronting: A Derivational Approach to Remnant Movement in German*. Dordrecht, Boston: Kluwer Academic Publishers.
- . 2000. Das Pronominaladverb als Reparaturphänomen. *Linguistische Berichte* 182:139-178.
- Munn, Alan. 1991. Binding in gerunds and the Leftness Condition. In *MIT Working Papers in Linguistics*, 163-178. Cambridge, MA: MITWPL, Department of Linguistics and Philosophy, MIT.
- Murasugi, Keiko, and Mamoru Saito. 1995. Adjunction and cyclicity. In *Proceedings of the West Coast Conference on Formal Linguistics*, eds. Raul Aranovich, William Byrne, Susanne Preuss and Martha Senturia, 302-317. Stanford, CA: CSLI Publications, Stanford University.

- Muysken, Pieter. 1977. Movement rules in Papiamentu. *Amsterdam Creole Studies* 1:80-102.
- . 1980. Papiamentu prepositions stranding and the typology of empty categories. Paper presented at LSA, Albuquerque.
- Neeleman, Ad. 1994. Complex Predicates. Led, Utrecht.
- . 2001. Particle Placement. Ms. University College London.
- Nemoto, Naoko. 1991. Scrambling and conditions on A-Movement. Paper presented at *Proceedings of the West Coast Conference on Formal Linguistics*, Stanford, Calif.
- Newmeyer, Frederick J. 1998. Preposition Stranding: Parametric Variation and Pragmatics. *Languages and Linguistics* 1:1-24.
- Nilsen, Øystein. 2001. Adverb order in Type Logical Grammar. In *Proceedings of the Amsterdam Colloquium 2001*, eds. A. van Rooy and Martin Stokhof.
- . 2003. *Eliminating Positions: Syntax and semantics of sentence modification*. Utrecht, NL: LOT.
- Nissenbaum, John. 2001. Investigations of Covert Phrasal Movement. Ph. D. Dissertation. MIT, Cambridge, MA.
- Nomura, Masashi. 2001. Extraposition or Scattered Deletion? Ms. Storrs, CT.
- Noonan, M. 1997. Functional Architecture and *Wh*-Movement: Irish as a Case in Point. *Canadian Journal of Linguistics* 42:111-139.
- Nunes, Jairo. 1995. The copy theory of movement and linearization of chains in the Minimalist Program. Doctoral dissertation. University of Maryland, College Park, MD.
- . 1996. On why traces can't be phonetically realized. *Proceedings of NELS 26*. Amherst, MA: GLSA. 211-225. In *Proceedings of the North East Linguistics Society 26*, 211-225. Amherst, MA: GLSA.
- . 1999. Linearization of Chains and Phonetic Realization of Chain Links. In *Working Minimalism*, eds. Samuel David Epstein and Norbert Hornstein, 217-250. Cambridge, MA: MIT Press.
- . to appear. Linearization of Chains and Sideward Movement. Ms. Universidade Estadual de Campinas.
- Nunes, Jairo, and Juan Uriagereka. 2000. Cyclicity and Extraction Domains. *Syntax* 3:20-43.
- Obenauer, Hans-Georg. 1976. *Etudes de syntaxe interrogative de français*. Tübingen: Niemeyer.
- Oku, Satoshi. 1998. A theory of selection and reconstruction in the minimalist perspective. Ph.D. Dissertation. University of Connecticut, Storrs, CT.
- Oppenrieder, Wilhelm. 1991. Preposition Stranding in Deutschen? - Da will ich nichts von hören! In *Strukturen und Merkmale Syntaktischer Kategorien*, eds. Gisbert Fanselow and Sascha Felix, 159-173. Tübingen: Gunter Narr.
- Ormazabal, Javier. 1995. The syntax of complementation: On the connection between syntactic structure and selection. Doctoral dissertation. University of Connecticut, Storrs.
- Pafel, Jürgen. 1996. Die syntaktische und semantische Struktur von *was für*-Phrasen. *Linguistische Berichte* 161:37-67.

- Pesetsky, David. 1987. *Wh*-in situ: Movement and unselective binding. In *The representation of (in)definiteness*, eds. Eric J. Reuland and Alice G. B. ter Meulen, 98-129. Cambridge, MA: MIT Press.
- . 2000. *Phrasal Movement and Its Kin*: Linguistic Inquiry Monograph 37. Cambridge, MA: MIT Press.
- Pesetsky, David, and Esther Torrego. 2001. T-to-C Movement: Causes and Consequences. In *Ken Hale: a Life in Language*, ed. Michael Kenstowicz, 355-426. Cambridge, MA: MIT Press.
- . to appear. Tense, Case, and the Nature of Syntactic Categories. In *The Syntax of Time*, eds. Jacqueline Guéron and Jacqueline Lecarme. Cambridge, MA: MIT Press.
- Pollard, Carl, and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar. Volume 1: Fundamentals*: CSLI Lecture Notes Series No. 13. Stanford, CA: Center for the Study of Language and Information.
- Pollock, Jean-Yves. 1989a. Opacity, genitive subjects, and extraction from NP in English and French. *Probus* 1:151-162.
- . 1989b. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20:365-424.
- Postal, Paul M. 1972. On some rules that are not successive cyclic. *Linguistic Inquiry* 3:211-222.
- . 1974. *On raising: One rule of English grammar and its theoretical implications*. Cambridge, MA: MIT Press.
- . 1993. Some defective paradigms. *Linguistic Inquiry* 24:357-364.
- . 1998. *Three investigations of extraction*: Current Studies in Linguistics Series 29. Cambridge, Mass: MIT Press.
- Price, Susan. 1990. *Comparative constructions in Spanish and French syntax*. London: Routledge.
- Progovac, Ljiljana. 1996. Clitics in Serbian/Croatian: Comp as the Second Position. In *Approaching Second: Second Position Clitics and Related Phenomena*, eds. A. Halpern and A. Zwicky, 411-428. Stanford: CSLI Publications.
- Radanović-Kocić, Vesna. 1988. The grammar of Serbo-Croatian clitics: A synchronic and diachronic perspective. Doctoral dissertation. University of Illinois, Urbana-Champaign.
- Ramchand, Gillian, and Peter Svenonius. 2002. The Lexical Syntax and Lexical Semantics of the Verb-Particle Construction. In *Proceedings of WCCFL 21*, eds. L. Mikkelsen and Chris Potts, 101-114. Summerville: Cascadia Press.
- Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657-720.
- Richards, Norvin W. III. 1997. What Moves Where When in Which Language? Ph.D. dissertation. MIT, Cambridge, MA.
- Riemsdijk, Henk van. 1975. A Case for a Trace. *Linguistics in the Netherlands 1972-1973*:194-200.
- . 1978a. *A case study in syntactic markedness: The Binding Nature of Prepositional Phrases*. Lisse: The Peter de Ridder Press.

- . 1978b. On the Diagnosis of *Wh* Movement. In *Recent Transformational Studies in European Languages*, ed. Samuel Jay Keyser, 189-206. Cambridge, MA: MIT Press.
- . 1990. Circumpositions. In *Unity in Diversity. Papers Presented to Simon Dik on his 50th Birthday*, eds. Harm Pinkster and Inge Genee, 229-241. Dordrecht: Foris.
- . 1998. Head Movement and Adjacency. *Natural Language and Linguistic Theory* 16:633-678.
- Riemsdijk, Henk van, and Edwin Williams. 1985. *Introduction to the theory of grammar*. Cambridge, MA: MIT press.
- Rizzi, Luigi. 1980. Violations of the *wh*-island constraint and the Subjacency condition. *Journal of Italian Linguistics* 5:157-195.
- . 1990. *Relativized minimality*. vol. 16: Linguistic Inquiry Monograph. Cambridge, MA: MIT Press.
- . 1997. The Fine Structure of the Left Periphery. In *Elements of Grammar: Handbook in Generative Syntax*, ed. Liliane Haegeman, 281-337. Dordrecht, Netherlands: Kluwer.
- Roberge, Yves. 1998. Les prépositions orphelines dans diverses variétés de français d'Amérique du Nord. In *Français d'Amérique: variation créolisation, normalisation*, ed. P. Brasseur. Avignon: Presses de l'Université d'Avignon.
- Roberge, Yves, and Nicole Rosen. 1999. Preposition Stranding and Que-Deletion in Varieties of North American French. *Linguistica Atlantica* 21:153-168.
- Roberts, Ian. 1991. Excorporation and minimality. *Linguistic Inquiry* 22:209-218.
- Rosenbaum, Peter S. 1967. *The grammar of English predicate complement constructions*. Cambridge, MA: MIT Press.
- Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral dissertation. MIT, Cambridge, MA.
- . 1969. Guess who? Paper presented at *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*, Chicago, Ill.
- Rouveret, Alain. 1990. X-bar theory, minimality, and barrierhood in Welsh. In *The Syntax of the Modern Celtic Languages: Syntax and Semantics 23*, ed. Randall Hendrick, 27-79. New York: Academic Press.
- . 1991. Functional Categories and Agreement. *The Linguistic Review* 8:353-387.
- . 1994. *Syntaxe du Gallois: Principes Généraux et Typologie*. Paris: CNRS éditions.
- Sabbagh, Joseph A. 2003. Right Node Raising is Sometimes Movement Before and Sometimes Movement after Spellout. Paper presented at *WCCFL 22*.
- Sag, Ivan A. 1978. Floated quantifiers, adverbs, and extraction sites. *Linguistic Inquiry* 9:146-150.
- . 1980. A further note on floated quantifiers, adverbs, and extraction sites. *Linguistic Inquiry* 11:255-257.
- Sag, Ivan A., and Janet D. Fodor. 1995. Extraction without traces. In *The Proceedings of the 13th West Coast Conference on Formal Linguistics*, eds. Raul Aranovich, William Byrne, Susanne Preuss and Martha Senturia, 365-384. Stanford, CA: Center for the Study of Language and Information.

- Saito, Mamoru, and Keiko Murasugi. 1999. Subject predication within IP and DP. In *Beyond Principles and Parameters: Essays in Memory of Osvaldo Jaeggli*, eds. Kyle Johnson and Ian Roberts, 167-188. Dordrecht: Kluwer Academic Publisher.
- Salles, Heloisa Maria Moreira-Lima. 1997. Prepositions and the Syntax of Complementization. Ph.D. Dissertation. University of Wales, Bangor.
- Schütze, Carson. 1994. Serbo-Croatian second position clitic placement and the phonology-syntax interface. In *MIT Working Papers in Linguistics*, 373-473. Cambridge, MA: MITWPL, Department of Linguistics and Philosophy, MIT.
- Sekerina, Irina A. 1997. The Syntax and Processing of Scrambling Constructions in Russian. State University of New York at Stony Brook.
- Sharvit, Yael. 1999. Resumptive Pronouns in Relative Clauses. *Natural Language and Linguistic Theory* 17:587-612.
- Shima, Etsuro. 2000. A preference for Move over Merge. *Linguistic Inquiry* 31:375-385.
- Shlonsky, Uri. 1997. *Clause Structure and Word Order in Hebrew and Arabic: An Essay in Comparative Semitic Syntax*. New York: Oxford University Press.
- Siegel, Muffy E. A. 1983. Problems in Preposition Stranding. *Linguistic Inquiry* 14:184-188.
- Siewierska, A. 1984. Phrasal Discontinuity in Polish. *Australian Journal of Linguistics* 4:57-71.
- Siewierska, A., and L Uhlířová. 1998. An overview of word order in Slavic languages. In *Constituent Order in the Languages of Europe*, ed. A Siewierska, 105-149. Berlin, New York: de Gruyter.
- Sigurðsson, Halldór Ármann. 1996. Icelandic Finite Verb Agreement. *Working Papers in Scandinavian Syntax* 57:1-46.
- Sobin, Nicholas. 1987. The Variable Status of Comp-Trace Phenomena. *Natural Language and Linguistic Theory* 5:33-60.
- Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19:425-449.
- Starke, Michal. 2001. Move Dissolves into Merge: a Theory of Locality. Doctoral dissertation. University of Geneva.
- Stepanov, Arthur. 2001. Late Adjunction and Minimalist Phrase Structure. *Syntax* 4:94-125.
- Stjepanovic, Sandra. 1998. On the placement of Serbo-Croatian clitics: Evidence from VP ellipsis. *Linguistic Inquiry* 29:527-537.
- . 1999. What do second position cliticization, scrambling, and multiple wh-fronting have in common? Ph.D dissertation. University of Connecticut, Storrs, CT.
- Stowell, Timothy. 1981. Origins of phrase structure. Doctoral dissertation. MIT, Cambridge, MA.
- . 1982a. Conditions on Reanalysis. *MIT Working Papers in Linguistics* 4:245-269.
- . 1982b. The tense of infinitives. *Linguistic Inquiry* 13:561-570.
- Sugisaki, Koji. 2002. Innate Constraints on Language Variation: Evidence from Child Language. Doctoral Dissertation. University of Connecticut, Storrs, CT.
- Sugisaki, Koji, Howard Lasnik, and William Snyder. 2001. Obligatory Object Shift in Child English and its Consequences. Ms. University of Connecticut, Storrs, CT

- Sugisaki, Koji, and William Snyder. 2001. Preposition Stranding in the Acquisition of English and the Theory of Parameters. Ms. University of Connecticut, Storrs, CT
- Svenonius, Peter. 2002. Limits on P: *filling in holes* vs. *falling in holes*. Ms. Trmsoø
- . 2003. Phases at the Interface. Paper presented at *ISP Workshop: EPP and Phases*, MIT, Cambridge, Mass.
- Takahashi, Daiko. 1994. Minimality of movement. Doctoral dissertation. University of Connecticut, Storrs.
- Takami, Ken-ichi. 1988. Preposition Stranding: Arguments against Syntactic Analyses and an Alternative Functional Explanation. *Lingua* 76:299-335.
- . 1992. *Preposition Stranding: From Syntactic to Functional Analyses*. Berlin: Mouton de Gruyter.
- Takano, Y. 2000. Illicit Remnant Movement: An Argument for Feature-Driven Movement. *Linguistic Inquiry* 31:141-156.
- Tallerman, Maggie. 1998. The uniform Case-licensing of subjects in Welsh. *The Linguistic Review* 15:69-133.
- Terzi, Arhonto. 2001. Complex Prepositions as Small Clauses. Paper presented at *The 22nd Linguistics meeting of the University of Thessaloniki*, Thessaloniki, GR.
- Tomić, Olga Miseska. 1996. The Balkan Slavic clausal clitics. *Natural Language and Linguistic Theory* 14:811-872.
- Travis, Lisa. 1984. Parameters and effects of word order variation. Doctoral dissertation. MIT, Cambridge, MA.
- Trissler, Susanne. 1993. P-Stranding im Deutschen. In *Extraktion im Deutschen I*, eds. Franz-Josef d' Avis, Sigrid Beck, Uli Lutz, Jürgen Pafel and Susanne Trissler, 248-291. Universität Tübingen, Universität Stuttgart.
- Tuller, L A. 1986. Bijective Relations in Universal Grammar and the Syntax of Hausa. Ph.D. Dissertation. University of California, Los Angeles.
- Ura, Hiroyuki. 1993. On feature-checking for *wh*-traces. In *Papers on Case and Agreement I - MIT Working Papers in Linguistics*, eds. Jonathan D. Bobaljik and Colin Phillips, 215-242. Cambridge, MA: MITWPL, Department of Linguistics and Philosophy, MIT.
- Uriagereka, Juan. 1999. Multiple Spell-Out. In *Working minimalism*, eds. Norbert Hornstein and Samuel D. Epstein, 251-282. Cambridge, MA: MIT Press.
- Vikner, Sten. 1995. *Verb movement and expletive subjects in Germanic languages*. Oxford: Oxford University Press.
- Vukic, Sasa. 2003. Ph.D. Dissertation. University of Connecticut, Storrs, Conn.
- Webelhuth, Gert. 1992. *Principles and parameters of syntactic saturation*. New York: Oxford University Press.
- Wexler, Kenneth, and Peter Culicover. 1980. *Formal principles of language acquisition*. Cambridge, MA: MIT Press.
- Wilder, Chris. 1997a. English Finite Auxiliaries in Syntax and Phonology. In *Clitics, Pronouns and Movement*, eds. James R Black and Virginia Motapanyane. Amsterdam/Philadelphia: John Benjamins.

- . 1997b. Some properties of ellipsis in coordination. In *Studies in Universal Grammar and Typological Variation*, eds. Artemis Alexiadou and Tracy Hall. Amsterdam: John Benjamins.
- . 1998. Transparent Free Relatives. In *Proceedings of WCCFL XIII*.
- . 1999. Shared Constituents and the LCA. Colloquium talk given at the University of Connecticut, March 12 1999.
- Wilder, Chris, and Damir Čavar. 1994. Long head movement? Verb movement and cliticization in Croatian. *Lingua* 93:1-58.
- Williams, Edwin. 1994. A reinterpretation of evidence for verb movement in French. In *Verb movement*, eds. David Lightfoot and Norbert Hornstein, 189-205. Cambridge: Cambridge University Press.
- . 2002. *Representation Theory*. Cambridge, Mass.: MIT Press.
- Willis, David. 2000. On the Distribution of Resumptive Pronouns and wh-trace in Welsh. *Journal of Linguistics* 36:531-573.
- Wunderlich, Dieter. 1984. Zur Syntax der Präpositionalphrasen im Deutschen. *Zeitschrift für Sprachwissenschaft* 3:65-99.
- Wurmbrand, Susanne. 2001. *Infinitives. Restructuring and Clause Structure*: Studies in Generative Grammar 55. Berlin and New York: Mouton de Gruyter.
- Zabrocki, T. 1984. on the Nature of Movement Rules in English and Polish. In *Contrastive Linguistics: Prospects and Problems*, ed. Jacek Fisiak. Amsterdam: Mouton.
- Zhang, Shi. 1990. Correlations between the Double Object Construction and Preposition Stranding. *Linguistic Inquiry* 21:312-316.
- Zribi-Hertz, Anne. 1984. Orphan Prepositions in French and the Concept of 'Null Pronoun'. Bloomington, IN: Indiana University Linguistics Club.
- Zwart, Jan-Wouter C. 2001. Syntactic and Phonological Verb Movement. *Syntax* 4:34-62.
- Zwarts, Frans. 1978. Extractie uit Prepositionale Woordgroepen in het Nederlands. In *Proeven van Neerlandistiek*, eds. A. van Berkel and et al., 303-399.