

Multiple dominance and interface operations

by

Chris O'Brien

Submitted to the Department of Linguistics & Philosophy
in partial fulfillment of the requirements for the degree of

Ph.D. in Linguistics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2017

© Massachusetts Institute of Technology 2017. All rights reserved.

Signature redacted

Author

.....

Department of Linguistics & Philosophy

September 8, 2017

Signature redacted

Certified by..

.....

David Pesetsky

Ferrari P. Ward Professor of Linguistics

Thesis Supervisor

Signature redacted

Certified by.

.....

Danny Fox

Anshen-Chomsky Professor of Language and Thought

Thesis Supervisor

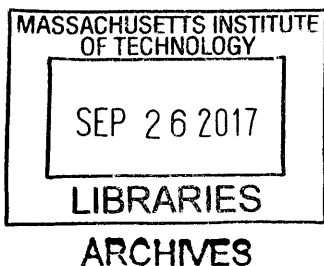
Signature redacted

Accepted by

.....

David Pesetsky

Head, Department of Linguistics & Philosophy



Multiple dominance and interface operations

by

Chris O'Brien

Submitted to the Department of Linguistics & Philosophy
on September 8, 2017, in partial fulfillment of the
requirements for the degree of
Ph.D. in Linguistics

Abstract

This dissertation explores the consequences of multidominance in syntactic theory, with a particular focus on how multidominance interacts with interpretation at the interfaces. In particular, I explore how interpretation is sensitive to *complete dominance*, in which a phrase dominates every position containing another phrase. I argue that complete dominance plays a crucial role in the resolution of two puzzles: The right-edge restriction on right-node raising and selective island effects in A'-movement. I develop a linearization algorithm which is locally sensitive to complete dominance, and show how, when applied to right-node raising structures, it predicts the right-edge effect. I also explore how, following Bachrach & Katzir (2009, 2017), complete dominance plays a role in cyclic Spellout. Bachrach & Katzir argue that PF Spellout of incompletely dominated material is *delayed*. I extend Bachrach & Katzir's delayed Spellout model to both PF and LF, and show how, when combined with Johnson's (2012, 2014) model of movement, it predicts the range of selective island effects (Cinque 1990, Postal 1998). Finally, I explore a puzzle concerning the PF theory of islands and the question of whether island constraints are active at LF.

Thesis Supervisor: David Pesetsky
Title: Ferrari P. Ward Professor of Linguistics

Thesis Supervisor: Danny Fox
Title: Anshen-Chomsky Professor of Language and Thought

Acknowledgments

This dissertation would not have been possible without the support, intellectual and otherwise, of my advisors, Danny Fox and David Pesetsky. The countless hours they spent helping me work out my ideas and constantly pushing me to develop analyses that were both careful and bold have been immeasurably helpful. Their sense of dedication will always be an inspiration to me. While I only began working with my third committee member, Roni Katzir, this past year, reading his work with Asaf Bachrach as a first year student played a major role in shaping what ultimately became my thesis, and I have been delighted to have him around while dissertating.

Of course, graduate school would have scarcely been survivable without the numerous fellow students and visitors who populate the 7th-9th floors of the Stata Center. I'm not going to name names, since you know who you are, but you all have made the last five years on the whole really quite fun. Thanks to all of you.

And of course, thanks to Juliet, my sister Elizabeth, and my parents, Cyndy and Tim.

Contents

1	Introduction	11
1.1	What this dissertation is about	11
1.2	Syntactic assumptions	15
1.2.1	Syntax with and without multiple dominance	15
1.3	Bachrach & Katzir	21
1.4	Putting the pieces together	22
2	Deriving the right-edge restriction on RNR	23
2.1	A restriction on right-node raising and its consequences	23
2.1.1	Multidominance and RNR	25
2.1.2	Linearization and compositionality	25
2.1.3	The cross-linguistic perspective	27
2.1.4	Comparison with alternatives	27
2.2	Arguments for multidominance in RNR	29
2.3	Data: The right edge restriction and right-node wrapping	32
2.4	A compositional theory of linearization	35
2.4.1	Dominance and complete dominance	35
2.4.2	Linearization: The basics	37
2.4.3	Complete dominance and linearization	39
2.4.4	Deriving the right-edge restriction and the right-node wrapping generalization	41
2.4.5	Left-node raising	43
2.4.6	Internal merge	44

2.5	Bachach & Katzir (2009, 2017)	45
2.6	Conclusion	51
3	How to get off an island	53
3.1	Multi-gap dependencies and selective islands	53
3.1.1	Symbiotic and parasitic gaps	53
3.1.2	Selective islands	55
3.1.3	The proposal	57
3.2	Data	60
3.2.1	Symbiotic gaps	60
3.2.2	Selective islands	64
3.3	A theory of movement	67
3.3.1	Internal merge and its discontents	67
3.3.2	Parallel Merge: A review of how it works	69
3.3.3	<i>Wh</i> -movement	70
3.3.4	Successive cyclicity	73
3.3.5	non-DP movement	74
3.4	Delayed Spellout	75
3.4.1	Cyclic Spellout and islandhood	75
3.4.2	LF and PF Spellout	76
3.4.3	Complete dominance and delayed Spellout	77
3.4.4	A sample derivation: DP movement	79
3.4.5	ATB-movement	81
3.5	How to get off an island	83
3.5.1	Anti-pronominality	86
3.5.2	DP/non-DP asymmetry	89
3.5.3	Scope reconstruction	89
3.5.4	Parasitic gaps	93
3.5.5	Delayed Spellout and single-gap sentences	96
3.5.6	Just-in-time Delivery	97

3.6	More issues	102
3.6.1	Some reconstruction puzzles	102
3.6.2	A syntactic question	104
3.6.3	<i>wh</i> -in-situ	106
3.7	Conclusion	110
3.7.1	The resumptive pronoun analysis revisited	110
3.7.2	Parasitic gap formation and ATB-movement: A unified analysis?	111
3.7.3	Selective islands and RNR: A comparison with the results of the previous chapter	113
4	Islands at PF and LF	117
4.1	Introduction	117
4.2	Islands at PF and LF	119
4.2.1	An argument for the PF theory of islands	119
4.2.2	An argument for the LF theory of islands	123
4.2.3	Summary	127
4.3	Analysis	128
4.3.1	PF and LF Constraints	128
4.3.2	<i>Wh</i> -in-situ and multiple questions	128
4.3.3	Predictions I: Deriving the ‘contradictory’ pattern	130
4.4	Predictions II: When deletion provides no salvation	132
4.4.1	Intervention effects and sluicing	132
4.4.2	Extending the prediction: Binding theory	138
4.5	Sluicing and scope parallelism	139
4.5.1	More on indefinites and islands	141
4.6	Two theories of LF islandhood	143
4.6.1	Bachrach & Katzir revisited, again	143
4.6.2	Cyclic linearization at LF	145
4.7	Conclusion	147

Chapter 1

Introduction

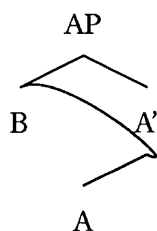
1.1 What this dissertation is about

A hypothesis has recently come to occupy the attention of many syntacticians that a single phrase may be simultaneously dominated by two or more mothers. Within contemporary merge-based theories (Chomsky, 1995), this *multidominant* view of syntax has proven fruitful in the analysis of various displacement phenomena (Wilder, 1999; Citko, 2005; Bachrach & Katzir, 2009, 2017; Johnson, 2012, 2014)¹. This thesis explores this hypothesis, focusing particularly on the interpretation of multidominant syntactic structures at the PF and LF interfaces. I argue that constraints on the interpretation of multidominant structures predict a variety of observed restrictions on displacement.

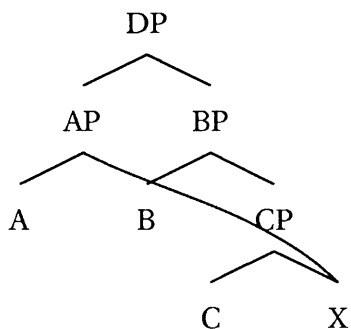
Minimalist theories typically allow Merge to generate at least one of two kinds of multidominant structures. First are internal Merge structures (Chomsky, 1995), like (1a), where a phrase is merged in two positions, one c-commanding the other. Second are *parallel Merge* structures (Citko, 2005), also known as sharing structures, in which a phrase is merged in two non-c-commanding positions (1b):

¹For some pre-minimalist antecedents, see Sampson (1975); Peters & Ritchie; Engdahl (1986); McCawley (1982).

(1) a. **Internal Merge**



b. **Parallel Merge**



These two varieties of multidominant structures are typically invoked in the analysis of different sorts of displacement. For instance, parallel Merge often forms part of the analysis of right-node raising, along with multi-gap dependencies like across-the-board movement and parasitic gap formation. Internal merge appears in the analysis of different sorts of movement phenomena (raising, *wh*-movement, etc.). However, this thesis will also explore the idea, introduced by Johnson (2012, 2014), that movement structures also involve parallel merge. According to Johnson's conception of the grammar, then, there is no principled difference between different displacement processes like right node raising and processes like movement.

I will focus on the interpretation of multidominant structures at the interfaces. Particularly important will be the observation that multidominance allows us to define a novel structural relation: In addition to ordinary dominance, we can now define a notion of *complete dominance*. A phrase X completely dominates a phrase Y when X dominates every position which Y occupies within some larger structure. Within (1b), for instance, DP completely dominates X, since it dominates both positions in which X is merged.

I will argue that interpretation at PF and LF is crucially sensitive to complete dominance. This idea will play a role in two distinct sorts of phenomena. First, I will argue

that it plays a role in the analysis of the *right-edge restriction* on right-node raising. A gap site² in an RNR construction must correspond to the rightmost position within the conjunct that it appears in. This is illustrated in (2) below:

- (2) a. ✓ The president gave an award to _ and then publicly praised **the Olympic gold medalists**.
- b. * The president gave _ an award and then publicly praised **the Olympic gold medalists**.

I will argue that this results from an interaction between multidominance and linearization. Linearization, I will argue, proceeds compositionally. That is, the linearization of any given node is typically sensitive only to material contained within that node. However, building on insights of Wilder (1999), I show that linearization is *locally* sensitive to complete dominance. I show that this allows us to derive the right-edge restriction without some of the problems associated with Wilder's original proposal.

Second, I explore island obviation effects in multi-gap dependencies and their sensitivity to *selective island effects*, such as *anti-pronominality* (Postal, 1993, 1994, 1998), illustrated below:

(3) **An anti-pronominal context**

Blake painted his house ✓ blue/*it.

Postal uses effects like anti-pronominality as a tool to explore the typology of English movement types. Postal shows that only certain types of extraction are subject to anti-pronominality. For instance, parasitic gap formation is (4a), but ordinary ATB-movement is not (4b):

- (4) a. * That's the color that Blake criticized _ after painting his house _.
(PG construction)
- b. ✓ That's the color that Bill criticized _ but Blake painted his house _.
(ATB-movement)

²I use 'gap site' as a theory-neutral term for the position in any displacement construction (e.g. movement, RNR) in which material is merged but not pronounced.

Postal argues, based on this sort of contrast, that ATB-movement and parasitic gap formation rely on distinct syntactic mechanisms. I argue against this position, and for a system where both processes involve sharing structures. Key to the contrast in (4) is not the type of multi-gap dependency, but the presence versus absence of an island. Building on work of Bachrach & Katzir (2009, 2017), I argue **in chapter 3** that this paradigm follows if PF and LF interpretation is sensitive to complete dominance. Bachrach & Katzir argue for a delayed Spellout mechanism, where a phrase that is merged inside of a Spellout domain K but also merged in a position outside of K may avoid Spellout in K. I show that this mechanism, when combined with the assumption that islandhood is determined at PF Spellout and with a model of movement developed by Johnson (2012, 2014), predicts selective island effects.

Finally, I explore a puzzle associated with the Spellout theory of islands. Kotek (2014) argues, based on data from multiple-*wh*-constructions in English, that phrasal movement, even when covert, is island sensitive. This appears to contradict an important assumption of my own analysis: that islandhood is determined by constraints on PF Spellout. I argue that this problem can be resolved if we assume that parallel island constraints apply at both PF and LF. These two constraints are stated below:

(5) **Constraints on extraction from islands**

a. **PF**

A phrase originating in an island may not be phonologically interpreted at any position outside of that island (except when ellipsis enables salvation by deletion)

b. **LF**

A phrase originating in an island may not be semantically interpreted in a position outside of that island

I discuss a few ways of deriving these constraints from assumptions compatible with the theories discussed in this dissertation.

Before continuing to the main body of the thesis (**chapters 2-4**), the rest of this chapter will provide some more preliminary information. First I outline the syntactic as-

sumptions that will inform the rest of the dissertation. Second, I briefly outline the relationship between the work in this thesis and the theory developed by Bachrach & Katzir (2009, 2017). Finally I discuss some remaining tensions between the dissertation's analysis of right-node raising and its analysis of selective islands.

1.2 Syntactic assumptions

This section summarizes this dissertation's basic assumptions about syntax. Throughout, I will be assuming a Merge-based theory, as first proposed by Chomsky (1995). Such theories are typically derivational. Beginning with a set of primitives (lexical items), a derivational step is determined by where application of Merge is permissible. In the first subsection, I examine exactly what elements a Merge-based conception of the derivation requires in order to admit multidominant structures. As we will see, the required elements are non-trivial, and this raises the question of whether multidominance is correct in the first place. This dissertation stands as an extended vindication of the hypothesis in the face of the reader's natural skepticism. Besides hopefully being inherently interesting, a subsidiary goal of this section is to introduce the reader to specific assumptions about the mechanics of syntactic derivations that will prove important in **chapter 3**.

1.2.1 Syntax with and without multiple dominance

Following Chomsky (1995, a.o.), I assume that Merge (6), an operation that simply takes two well-formed syntactic objects and forms a third, is the fundamental structure building operation.

$$(6) \text{ Merge}(A,B) = \{C, \{A,B\}\} = [{}_C A B]^3$$

Next I assume that derivations take place within *workspaces*. A derivation begins with an initial workspace, commonly called a *numeration*, that contains only *tokens* of lexical

³I use C as a placeholder for the label assigned to the entire structure. This thesis does not make any commitments with respect to the appropriate theory of labeling.

items, in other words ordered pairs of lexical items and natural numbers (so that any lexical item has an infinite number of tokens).

A simple theory, which allows neither internal nor parallel Merge, can be defined as follows:

(7) **Recursive definition of a language (without internal or parallel Merge)**

Let LEX be a set of lexical items.

Let \mathbb{N} be the set of natural numbers.

The set of possible workspaces \mathcal{W} is the minimal set such that

1. for all $N \subseteq \text{LEX} \times \mathbb{N}$, $N \in \mathcal{W}$
(Numerations)
2. If $\{A, B, C_1, \dots, C_n\} \in \mathcal{W}$, then $\{\text{Merge}(A, B), C_1, \dots, C_n\} \in \mathcal{W}$
(external merge)

The set of well-formed syntactic objects \mathcal{S} is the union of all workspaces in \mathcal{W}
($\mathcal{S} = \bigcup \mathcal{W}$)

The statement (7.1) states that any set of lexical item tokens is a possible workspace. This determines the set of possible numerations. A possible derivational step is defined in (7.2). This says that if a workspace contains any two mergeable objects A and B, then the workspace derived by replacing these two objects with $\text{Merge}(A, B)$ is also a workspace. Finally, the set of well-formed syntactic objects is defined as the set of everything that can appear inside of a workspace.

Notice that this definition of the grammar derives a particularly strong version of Chomsky's (1995) extension condition, stated below:

(8) EXTENSION CONDITION:

Both arguments of Merge must be elements in some workspace

This definition not only rules out iterations of Merge which apply countercyclically, but also operations which don't involve independent syntactic objects. In other words, it rules out internal Merge. Indeed, this grammar has an even stronger property: Merger

of any two syntactic objects must not share any of the same lexical item tokens as constituents. I'll call this the disjointness condition:

(9) DISJOINTNESS CONDITION

Both arguments of Merge are disjoint

(i.e. they contain no identical lexical item tokens)

This follows because any derivation begins with a single set of lexical item tokens. Since each token of a lexical item is treated as a distinct object, and since within the course of a derivation Merge may only apply to a given lexical item token once, it follows that if the same lexical item appears in two phrases within a workspace, they must be distinct tokens. This disjointness property rules out the possibility of parallel merge.

A more liberal grammar can be developed which loosens up the extension and disjointness conditions slightly, by allowing merger of two objects where one of the objects contains the other. This sort of grammar generates both external and internal Merge structures:

(10) **Recursive definition of a language (with external and internal Merge)**

Let LEX be a set of lexical items.

Let \mathbb{N} be the set of natural numbers.

The set of possible workspaces \mathcal{W} is the minimal set such that

1. for all $N \subseteq \text{LEX} \times \mathbb{N}$, $N \in \mathcal{W}$

(numerations)

2. If $\{A, B, C_1, \dots, C_n\} \in \mathcal{W}$, then $\{\text{Merge}(A, B), C_1, \dots, C_n\} \in \mathcal{W}$

(external Merge)

3. If $\{A, C_1, \dots, C_n\} \in \mathcal{W}$ and A properly contains⁴ some B, then $\{\text{Merge}(A, B), C_1, \dots, C_n\} \in \mathcal{W}$

\mathcal{W}

(internal Merge)

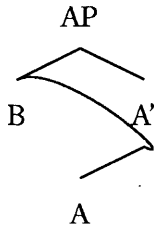
The set of well-formed syntactic objects \mathcal{S} is the union of all workspaces in \mathcal{W}

($\mathcal{S} = \bigcup \mathcal{W}$)

⁴A properly contains B iff A dominates B and $A \neq B$.

This definition extends the one in (7) to include structures like (1a), repeated below:

(11) **Internal Merge**⁵



(12) $\{A_1, B_1\} \in \mathcal{W}$, via (10.1)

(13) $\{[A' A_1 B_1]\} \in \mathcal{W}$, via (10.2)

(14) $\{[_{AP} B_1 [_{A'} A_1 B_1]]\} \in \mathcal{W}$, via (10.3)

This new definition allows Merge to produce multidominant structures, as long as one of the positions *c*-commands the other. In other words, this derives a slightly different version of the extension condition:

(15) EXTENSION CONDITION (REVISED)

Both arguments of Merge must be (i) elements of a single workspace or (ii) in a proper containment relationship

This definition still rules out the parallel merge structure. To define this sort of structure, I propose that, under certain circumstances, sub-derivations may proceed independently up to a particular point, at which point they combine into a single derivation. I model this using the notion *Collect* (after Fox & Pesetsky 2007), defined in (16.3) below, which says that two workspaces can combine into a single workspace. A grammar with this property is provided below:

(16) **Recursive definition of a language (with parallel Merge)**

Let LEX be a set of lexical items.

Let \mathbb{N} be the set of integers.

The set of of workspaces \mathcal{W} is defined as the minimal set such that:

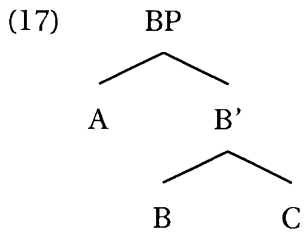
⁵Throughout the dissertation, non-terminal nodes are labelled according to \bar{X} -theoretic conventions. This is done entirely for ease of presentation.

1. If $N \subseteq \text{LEX} \times \mathbb{N}$, then $N \in \mathcal{W}$
(numerations)
2. If $\{A, B, C_1, \dots, C_n\} \in \mathcal{W}$, then $\{\text{Merge}(A, B), C_1, \dots, C_n\} \in \mathcal{W}$
(external merge)
3. If $W_1, W_2 \in \mathcal{W}$, then $W_1 \cup W_2 \in \mathcal{W}$
(collect)

The set of well-formed syntactic objects \mathcal{S} is the union of all workspaces in \mathcal{W}
($\mathcal{S} = \bigcup \mathcal{W}$)

Like the earlier definitions, any set of lexical item tokens is a possible workspace (16.1), and new workspaces can be derived via external Merge (16.2). The novelty is the derivational step defined in (16.3). This says that any two workspace may be combined together to form a new workspace.

This definition will admit all of the same structures as the external Merge-only definition in (7). As an instance, the proof that this definition admits the simple structure in (17) is provided below:



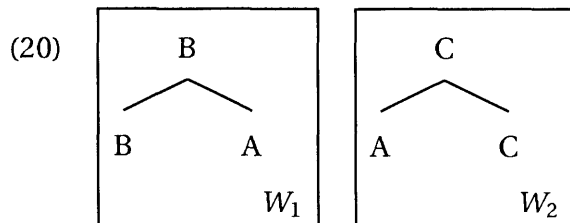
- (18)
- a. $\text{LEX} = \{A, B, C\}$
 - b. $\{A_1, B_1, C_1\} \in \mathcal{W}$
via (18a), (16.1)
 - c. $\{A_1, [B' B_1 C_1]\} \in \mathcal{W}$
Via (18b), (16.2), definition of Merge
 - d. $\{[A_1 [B' B_1 C_1]]\} \in \mathcal{W}$
Via (18c), (16.2), definition of Merge

- e. $[A_1 [B_1 C_1]] \in \mathcal{S}$
 via (18.d), (16)

It also licenses internal Merge structures:

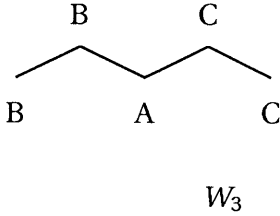
- (19) a. $LEX = \{A, B\}$
 b. $\{A_1, B_1\} = \mathcal{W}1$
 via (19a), (16.1)
 c. $\{B_1\} = \mathcal{W}2$
 via (19a), (16.1)
 d. $\{[A_1, B_1]\} = \mathcal{W}1$
 via (19b), (16.2)
 e. $\{B_1, [A_1, B_1]\} = \mathcal{W}3$
 via (19c,d), (16.3)
 f. $\{[B_1 [A_1, B_1]]\} = \mathcal{W}3$
 via (19e), (16.2)

However, parallel Merge structures are also now admitted. To see how, it's helpful to illustrate the derivation with syntactic objects represented as graphs and workspaces as boxes surrounding them. Suppose that we have a lexicon containing the items A, B, and C. The grammar will allow us to have the numerations $W_1 = \{B_1, A_1\}$ and $W_2 = \{C_1, A_1\}$. Crucially, each workspace contains the same token of A. If merge of the two objects in each workspace applies (20), and if Collect subsequently applies, combining the two workspaces together, the result is a single workspace (W_3) which contains two objects. However, each object contains the same token of A (21). This is the derivational sequence responsible for parallel Merge configurations, and is illustrated below:



$\xrightarrow{\text{COLLECT}}$

(21)



- (22)
- a. $\text{LEX} = \{A, B, C\}$
 - b. $\{A_1, B_1\}, \{A_1, C_1\} \in \mathcal{W}$
 - c. $\{[BP A_1 B_1]\}, \{[CP A_1, C_1]\} \in \mathcal{W}$
 - d. $\{[BP A_1 B_1], [CP A_1, C_1]\} \in \mathcal{W}$
 - e. $\{[[BP A_1 B_1] [CP A_1, C_1]]\} \in \mathcal{W}$

What's interesting about this proposal is that it can derive a number of multidominant structures via Merge operations that locally obey the same extension and disjointness conditions that held for the most restrictive external merge-only grammar. What makes this possible is the assumption that just as Merge combines syntactic objects, entire workspaces can also be combined together using the operation Collect. As we will see in **chapter 3**, this is not simply a novel technical device, but can also be made to do empirically testable work for us.

1.3 Bachrach & Katzir

In many respects, this dissertation is an extended response to the system developed in Bachrach & Katzir (2009, 2017). Their theory starts from a very similar set of concerns as this dissertation, but reaches quite different conclusions in many ways. In this section, I wish to highlight how the comparison between Bachrach & Katzir's own system and the one developed here will be distributed within the body of the dissertation.

Since their work is largely concerned with deriving the edge restriction on RNR structures and right-node wrapping, the same empirical concerns as **chapter 2**, that chapter

will include an extensive discussion of their proposal (in **section 2.5**). **Chapter 3** utilizes another important piece of technology from Bachrach & Katzir, the assumption that incomplete dominance at a given Spellout domain entails delayed Spellout. However, as I discuss in that chapter, my use of this technology differs considerably from Bachrach & Katzir in empirically testable ways. Finally, in **chapter 4, section 4.6.1** I show that Bachrach & Katzir's system provides a way of resolving the puzzle discussed in that chapter, though in a way that is incompatible with the world view developed in this dissertation.

1.4 Putting the pieces together

As mentioned above, the next two chapters crucially rely on a notion of complete dominance. However, the notion is deployed in distinct ways in each chapter and the reader will notice a certain, unfortunately irreducible, tension between them.

The source of this tension can be diagnosed as follows. To prefigure the discussion somewhat, the analysis of right-node raising in **chapter 2** will rely crucially on the assumption that linearization is strictly *local*. That is, the linearization of any given node depends exclusively on information contained within that node. Conversely, the analysis of selective islands **chapter 3** explores the possibility that, when a node *N* is multiply dominated at a position within a given Spellout domain *K* and also at some position outside of *K*, then *N* can avoid interpretation inside of *K*. This is the delayed Spellout mechanism originally proposed by Bachrach & Katzir. Because this notion of delayed Spellout depends crucially on information outside of *K*—more precisely, other positions of *N* outside of *K*—the interpretation mechanism is necessarily non-local. Since the theoretical and empirical domain will hopefully be clearer to the reader after reading these two chapters, I will more fully discuss, though by no means resolve, this puzzle in the conclusion of **chapter 3**. Nonetheless, it may be helpful for the reader to bear it in mind while reading the chapters.

Chapter 2

Deriving the right-edge restriction on RNR

2.1 A restriction on right-node raising and its consequences

A right-node raising (RNR) sentence canonically¹ involves a coordinate structure in which a particular phrase—the pivot—is associated with a position in each conjunct. In the example below, the pivot **the #1 hit song** is the VP object in each conjunct:

(23) **Right-node raising**

Dolly wrote _ and Whitney sang **the #1 hit song**.

RNR constructions are subject to a *right-edge restriction*. A gap site must correspond to the rightmost position within the conjunct that it appears in. This is illustrated in (24) below:

¹I ignore the phenomenon of non-coordinate, or Hudson-style RNR sentences (Hudson, 1976; Postal, 1994; Phillips, 1996). Intriguingly, though, unlike coordinate RNR (25c), this variety of RNR does not permit right-node wrapping:

- (1) a. ✓ Garth should polish _ before giving to Merle **the slide guitar**.
b. * Garth should polish _ before giving **the slide guitar** to Merle.

- (24) a. ✓ The president gave an award to _ and then publicly praised **the Olympic gold medalists**.
- b. * The president gave _ an award and then publicly praised **the Olympic gold medalists**.

In the grammatical example (24a), the gap site occupies the rightmost position in the left conjunct. In the ungrammatical (24b), the gap site appears in a non-rightmost position in the left conjunct.

It turns out, however, that while this restriction holds for all *non-final* conjuncts, it can be obviated in final conjuncts. This is sometimes termed the *right-node wrapping* phenomenon (Wilder, 1999; Whitman, 2009; Bachrach & Katzir, 2009, 2017). In the acceptable sentences below, the gap site appears within an intermediate position in the final conjunct:²

- (25) **Right-node wrapping**
- a. “I’ve got friends in low places, where the whiskey drowns _ and the beer chases **my blues** away.”³
- b. I defiled _ and then turned **the homework assignment** into a paper airplane.
- c. Garth should polish _ and then give **the slide guitar** to Merle.

This chapter argues that the right edge restriction and its obviation in right-node wrapping contexts provides insight into the relationship between syntactic structure and the sub-component of the grammar responsible for linearization. In particular, if we assume that RNR involves multidominant syntactic representations, in which the pivot is simultaneously merged in each conjunct, these data fall out from a model in which lin-

²Right-node wrapping is subject to variation in acceptability. For instance, some speakers reject the sentences in (25). Moreover, even for speakers who accept (25), not all RNR constructions allow wrapping. For instance, when the pivot is embedded inside of an embedded clause, the sentence becomes markedly worse:

- (1) a. ✓ John lost _ but told Mary that he returned to the library **the book about right-node raising**.
- b. * John lost _ but told Mary that he returned **the book** to the library.

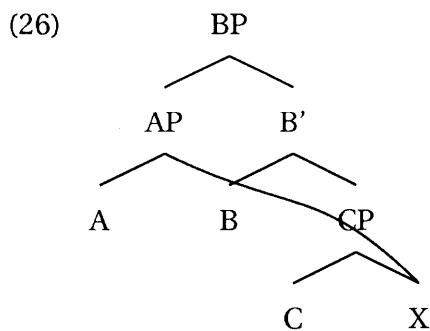
Although I offer no account of this variation, Bachrach & Katzir attribute it to restrictions on cyclic Spell-out.

³From “Friends in low places” (Brooks, 1990). The significance of this lyric for syntactic theory was first pointed out by Whitman (2009). The term *right-node wrapping* is also due to Whitman.

earization is determined *compositionally*. That is, the linearization of any given phrase is determined exclusively by information contained within that phrase.

2.1.1 Multidominance and RNR

Multidominance is utilized in many *pivot-internal* analyses of right-node raising (McCawley 1982, Wilder 1999, a.o.). In this analysis, the pivot is assumed to be in-situ, and is simply parallel Merged in each conjunct. This is schematically illustrated below, where a phrase X is simultaneously dominated by AP and CP:



A number of arguments have been proposed for this account of RNR. These arguments are reviewed in **section 2.2**.

2.1.2 Linearization and compositionality

Following Chomsky (1995) and others, I will assume that the output of the syntactic component of the grammar is unordered⁴, and that linearization is a part of PF interpretation. The central claim of this chapter is that the linearization procedure which applies at PF is compositional, in the sense that the linearization of a given node is blind to information outside of that node.

As is discussed in the introduction, I propose that linearization is sensitive to complete dominance. The proposal depends on one crucial property of complete dominance: the notion that “X *completely dominates* Y” can be defined only with respect to some larger structure. In (26), for instance, AP does not completely dominate X within

⁴This idea was originally proposed by Curry (1961) and Šaumjan & Soboleva (1963). For early critical discussions, see Chomsky (1965) and Hall (1965). It was later developed within the GPSG/HPSG tradition by Gazdar et al. (1985) and others.

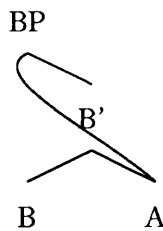
BP, since X is simultaneously merged in a second position within BP (i.e. CP) which is not in turn contained by AP. However, AP *does* completely dominate X with respect to the smaller structure AP itself, just CP completely dominates X with respect to CP itself and B'.

Although linearization applies compositionally, I will assume that it is *locally* sensitive to complete dominance. In a structure like (26), this means that the linearization of the elements in AP will order A and X, because each is completely dominated by AP with respect to the AP node itself. However the linearization of AP and CP within the BP will not involve X, because neither phrase completely dominates X with respect to BP. As we will see, this gives rise to a right edge effect.

I argue that the compositional linearization of multidominant structures only predicts the right-edge restriction to be active in *non-final* conjuncts. The model predicts the right-node wrapping effect, where the pivot may appear non-finally within the final conjunct.

Now consider internal Merge structures:

(27) **Internal Merge**



There is a tension between this sort of multidominant structure and the principle of compositionality. In a structure like (27), a phrase 'A' is merged both in a complement and specifier position. Supposing the rules of linearization active for a given language typically place complements to the right of heads, thus placing B before A, and specifiers to the left of the heads, placing A before B, then we might naively expect the linearization component of the grammar to deliver contradictory instructions to PF. In **section 2.4.6** below, I outline some solutions to this problem and how they fit into the present architecture.

2.1.3 The cross-linguistic perspective

Japanese has been shown to display a phenomenon of *left-node raising* in which the pivot shows up in the initial, rather than the final conjunct (Yatabe, 2001). As Bachrach & Katzir (2017) observe, Japanese left-node raising sentences are subject to an edge restriction, such that the pivot must be merged at the leftmost position of every non-initial conjunct, and also permit wrapping effects. As we will see (section 2.4.5), the theory presented in this chapter can be extended to the Japanese data.

2.1.4 Comparison with alternatives

This chapter is a contribution to a line of research inaugurated by Wilder (1999), who proposes that the right-edge restriction (and its obviation in RNW contexts) falls out from an interaction between multidominant syntactic structures and a theory of linearization based on Kayne's (1994) *linear correspondence axiom* (LCA). Within Wilder's variant, the LCA is crucially dependent on complete dominance: for each pair of nodes A and B where A asymmetrically c-commands B, the set of terms which A completely dominates will be ordered before all terms that B completely dominates. Wilder shows that this suffices to allow for linearization of the sort of multidominant structures assumed by the pivot-internal analysis of RNR.

Like the system presented above, the crucial feature of Wilder's proposal is that, while it avoids contradictory orderings by invoking complete dominance, shared material is still linearizable within each conjunct. More specifically, if a shared node asymmetrically c-commands any material within a non-final conjunct, a contradictory ordering will be produced. Wilder's theory therefore predicts a version of the right-edge restriction; however, as Sabbagh (2007) points out, the version of the right-edge restriction predicted by Wilder's theory overgenerates.

(28) **Sabbagh's observation**

Wilder's system predicts that the pivot in a RNR sentence may not *asymmetrically c-command* any node in a non-final conjunct.

It follows that embedding a pivot should preserve asymmetry and obviate the right-edge restriction. As we see with (29), this prediction appears to be incorrect.

- (29) ✖ The president gave an award [to _] at the ceremony and then publicly praised **the Olympic gold medalists**.

In this sentence, the pivot is embedded within a PP in the left conjunct, and therefore shouldn't c-command the material to its right.

Another example, adapted from Sabbagh, is given below:

- (30) a. ✓ John edited [one review of _], and will edited another review of **my new syntax textbook**.
b. ✖ John edited [one review of _] for Blackwell, and will edited for MIT press another review of **my new syntax textbook**.

Again, since the pivot is embedded in the initial clause in a position where it does not c-command any material, this should be licit.

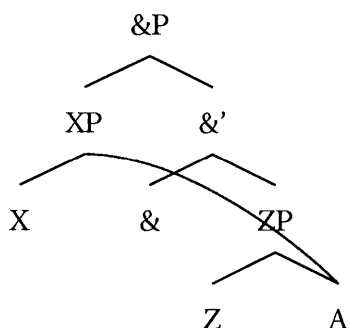
In part due to Sabbagh's observation, both Sabbagh and Bachrach & Katzir (2009, 2017) develop alternatives to Wilder's proposal, Sabbagh by abandoning the pivot-internal analysis and Bachrach & Katzir by radically weakening the constraints on linearization. I have retained Wilder's central hypothesis that linearization is sensitive to complete dominance. The central contribution of this chapter is to show that the right-edge restriction and right-node wrapping generalization is a consequence not of the LCA but of the assumption that linearization is a strictly compositional operation. Among other things, this avoids the overgeneration problem pointed out by Sabbagh. A detailed comparison with Bachrach & Katzir's system is provided in **section 2.5**.

2.2 Arguments for multidominance in RNR

This chapter assumes a pivot-internal multidominant analysis of RNR (McCawley, 1982, a.o.), in which the pivot is shared between the position corresponding to the gap site and the position where it appears. Below, I enumerate some familiar arguments for the pivot internal analysis over its pivot-external counterpart.⁵

The in-situ analysis is illustrated with the schematic example below (31):

(31) Pivot-internal multi-dominant RNR



According to this analysis, RNR is not strictly speaking a *movement* operation, since the pivot does not raise to a position c-commanding its base-generated positions. The major rival of this analysis is the *pivot-external* analysis, in which RNR is taken to be an overt movement process along the lines of leftward across-the-board movement.

Let's consider some arguments for this analysis. First, if the pivot external analysis is correct, we might expect—all other things being equal—that RNR should be sensitive to the same constraints as (i) leftward movement processes and (ii) other forms of rightward movement. For instance, we might expect RNR to be sensitive to islands. However, as Wexler & Culicover (1980) observe, RNR is island insensitive. For instance, in the sentence below, the pivot is merged within a relative clause in each conjunct, and so should be subject to the complex NP constraint (Ross, 1967):

(32) Island insensitivity in RNR sentences

✓ John talked to a guy who admires , and Bill met someone who despises , **the 44th president of the United States.**

⁵There is an additional proposal in the literature which treats RNR as a case of backwards ellipsis (Wexler & Culicover, 1980). I ignore this analysis, since it does not seem to provide a plausible account of the right-edge restriction and right-node wrapping generalization.

While this is puzzling from the perspective of the pivot-external account I should note that Sabbagh's (2007) analysis attempts to reconcile the pivot external analysis with these data.

It's also well known that rightward movement is subject to a set of rather severe locality constraints. One example of this is the ban on preposition stranding in English rightward extraction (33a). Notably, RNR is *not* sensitive to this ban (33b):

- (33) a. *Waylon edited a review of _ for Blackwell my syntax textbook.
b. ✓ Waylon edited a review of _ and Hank blogged about _, **my syntax textbook**.

Another locality condition, sometimes termed the *right roof constraint* (Ross, 1967, a.o.), which severely restricts unbounded rightward movement, is also not found in RNR. Again, if RNR is a variety of rightward movement, then we would expect it to be subject the same sorts of locality constraints as other varieties of rightward movement.⁶

Another argument against the pivot-external analysis is based on a set of observations showing that RNR may target constituents that ordinary movement processes typically *cannot* target. For instance, unlike movement (34), RNR can apply below the word level (Booij, 1985; Toman, 1985; Wilder, 1997):

- (34) a. *What does your theory under-_
b. *Soviet business practices, I study post-_
(35) a. My theory under-_, and your theory over-**generates**.
b. Lev wrote a book comparing pre-_ and some post-**Soviet business practices** in Russia.

Furthermore, unlike rightward (36a) or leftward (36b) movement, RNR can target constituents like a single N head (36c; Chaves 2007):

- (36) a. * John met an interesting _, yesterday teacher.
b. * What is Mary an interesting _?

⁶Sabbagh's analysis also attempts to account for this within the terms of the pivot-external theory, however his analysis invokes a large number of assumptions, rendering its evaluation beyond the scope of this dissertation.

- c. ✓ Mary is both an interesting _ and an inspiring **teacher**.

Another argument for an in-situ analysis comes from Zulu. Halpert (in prep) argues that the ‘conjoint/disjoint’ alternation in Zulu reflects the presence or absence of non-verbal material within a νP . When material is evacuated, an obligatory ‘disjoint’ marker (*-ya-* in present tense) appears on the verb. When material appears within the νP , then this marker is not allowed. Halpert shows that this alternation applies to a variety of right-dislocation phenomena that are plausibly analyzed as rightward movement.

What’s important for the present discussion are sentences involving coordinated sentences in which an object is shared:

- (37) *ngi-buk-e:la* *_ futhi npi-phinde* *ngi-dlale* ***ibho:la***
 1SG.SM-watch-APPL _ and 1SG.SM-again 1SG.SM-play.SJC **AUG.5.soccer**
 ‘I watched and I also played soccer.’

Crucially, even though the DP *ibho:la* ‘soccer’ doesn’t appear within the first conjunct, the disjoint marker—unlike in rightward movement processes in Zulu—is not obligatory. This suggests that, at least in Zulu, processes like right-node raising are morphologically distinct from movement processes, a conclusion compatible with an in-situ analysis but not obviously so with a movement analysis.

Finally, let us return to the right-node wrapping data, which formed a key motivation for our analysis. Recall that within the rightmost conjunct the pivot X of a right-node wrapping sentence appears to the left of a non-shared constituent Y. If, as the pivot-external theory stipulates, X is raised to a position c-commanding the coordinate structure, then the appropriate word order will obtain only if the the non-shard constituent Y has also moved to a higher position, in violation of the coordinate structure constraint (Ross, 1967):⁷

⁷It’s well known that some coordinate structures appear to obviate the coordinate structure constraint (Ross 1967, a.o.):

- (1) That’s the turkey that John went to the store and bought _.

One might wonder whether the class of coordinate structures that admit right-node wrapping is identical to those which allow CSC-violations, which could permit one to maintain the derivation in (38b). The data below provide the crucial test case. Sentence (2a) provides a VP level coordinate structure in which the CSC is active, and (2b) right-node wrapping in the same structure.

(38) a. **rightward ATB-movement of pivot**

[_{&P} John defiled _ & turned _ into a paper airplane] **the homework assignment**

b. **CSC-violating movement of PP** [_{&P} John defiled _ & turned _] **the homework assignment**

into a paper airplane

The possibility of an RNR pivot appearing at an intermediate position is very puzzling from the pivot-external perspective, but perfectly natural under the pivot-internal analysis.

2.3 Data: The right edge restriction and right-node wrapping

The right-edge restriction (Wilder 1999; Sabbagh 2007; Bachrach 2009, 2015) which we will account for is stated below:

(39) **The right edge restriction (right-edge restriction)**

The pivot must be merged in the rightmost position of every non-final conjunct.

Here, to say that the pivot is merged in the *rightmost position* of a conjunct is just to say that the linearization of that phrase within that position would place it to the right of all other phrases within the conjunct.

There are numerous apparent counter-examples to this generalization:

(40) a. John bought _ raw and then cooked **the shrimp that Mary ate**.

b. Fred hammered _ flat and Mary tempered **the armor for Sir Lancelot**.

(2) a. * That's the language that John knows French and speaks _ to his friends.

b. ✓ John knows _ and often speaks **German** to his friends.

According to the author's judgments, (2b) is acceptable; (2a) is not, suggesting that obviation of the CSC is not relevant. In addition, Postal (1998) observes that CSC-obviating coordinate structures constitute selective islands for leftward movement, meaning that—among other things—only DPs can move out of them. If this constraint also applies to rightward movement, then a derivation like (38b) would be ruled out anyway.

In both of these sentences, the pivot is base-generated within the initial conjunct inside a small clause, which seems to predict that it should be linearized to the left of *raw* and *flat*, in violation of the right-edge restriction.

These cases are accounted for by the assumption that *conjunct-internal* rightward movement is possible. If such movement is available, then a structural representation satisfying our generalization can be generated after all:

(41) **right-edge restriction satisfied through movement**

[&P [*Conj*₁ Fred hammered flat the armor] & [*Conj*₂ Mary tempered **the armor**]]

In this representation, the pivot is base-generated within the first conjunct in a position inside a small clause, but then undergoes rightward movement to a position at the right edge of the conjunct. The final representation satisfies the right-edge restriction. Crucially, rightward movement is possible with arguments which appear within these small clause structures:

- (42) a. John bought raw the shrimp.
 b. Fred hammered flat the metal.

In order to test the right-edge restriction, we must find cases where such clause-internal movement is blocked. This is illustrated schematically below, where the pivot is base-generated in a position which would linearize it to the left of Y, and further rightward movement is blocked:

(43) **right-edge restriction violated**

[&P [*Conj*₁ X **Pivot** Y] & [*Conj*₂ Z **Pivot**]]

The first case involves ditransitive constructions. As (44) illustrates, movement of the indirect object of a ditransitive to the right of the direct object is forbidden.

(44) **Ditransitives: IO cannot be extraposed**

- a. ✓ Willie gave the book a good review.
 b. * Willie gave a good review the book.

Given this contrast, the right-edge restriction predicts sentences in which the pivot is the indirect object of a non-final clause to be ungrammatical. As we saw above, this prediction is correct:

(45) **IO pivot of ditransitive cannot appear in non-final clause**

- a. ✓ Willie should edit _ and then send **the book** to the publisher.
- b. * Willie should give _ a good review and then send **the book** to the publisher.

A similar point can be drawn from the observation that rightward movement in English disallows preposition-stranding:

(46) **Rightward movement cannot strand prepositions⁸**

- a. * Waylon edited a review of _ for Blackwell my syntax textbook.
- b. * Waylon returned the book to _ in good condition the Somerville public library.

Again, the right-edge restriction predicts that the pivot may not be Merged as the complement of a preposition in a non-final conjunct:

(47) **Preposition stranding and the right-edge restriction**

- a. ✓ Waylon returned a book to _ and then checked out a DVD from **the Somerville public library**.
- b. * Waylon returned a book to _ in good condition and checked out a DVD from **the Somerville public library**.
- c. ✓ Hank edited a review of _ and then wrote a blog post about **my new syntax textbook**.
- d. * Hank edited a review of _ for Blackwell and then wrote a blog post about **my new syntax textbook**.

Examples (47a,c) show that, unlike typical rightward movement, RNR generally allows preposition-stranding. However, when material intervenes between the gap-site and the

⁸Adapted from Sabbagh (2007).

edge of the conjunct, then the sentence is degraded (47b,d), as the right-edge restriction predicts.

As we also saw above, the right-edge restriction does *not* apply within the final conjunct. Some examples of grammatical right-node wrapping sentences are found below:

- (48) a. I defiled _ and then turned **the homework assignment** into a paper airplane.
b. Garth should read _ and then give **the book** to Merle.

Again, I will call this the right-node wrapping generalization:

(49) **Right-node wrapping generalization**

The pivot may be non-rightmost within the final conjunct

2.4 A compositional theory of linearization

2.4.1 Dominance and complete dominance

Let's first define a notion of reflexive dominance:

(50) **Reflexive dominance**

- a. If $X = Y$, then X reflexively dominates Y ,
b. If X reflexively dominates a mother of Y , then X reflexively dominates Y

This sort of definition will suffice for a grammar which generates only trees, that is only structures in which each phrase has at most one mother. However if a grammar allows a phrase to have more than one mother, then more definitions of dominance become available. In particular, we can define a notion of *complete dominance*.⁹

(51) **Complete dominance**

- a. If $X = Y$, then X completely dominates Y ,

⁹My formulation of complete dominance differs slightly from Bachrach & Katzir's, for reasons irrelevant to the present discussion, but which will be relevant in **chapter 4**. The notion of complete dominance dates back to Wilder (1999), who calls it *full dominance*. Fox and Pesetsky (2007) use the term *total dominance* to distinguish their definition, virtually identical to my own, from Bachrach and Katzir's definition. I'm sticking with this label, though, since it is used with the definition above in other works (e.g. Gracanin-Yukse 2007).

- b. If X completely dominates every mother of Y within some Z,
then X completely dominates Y

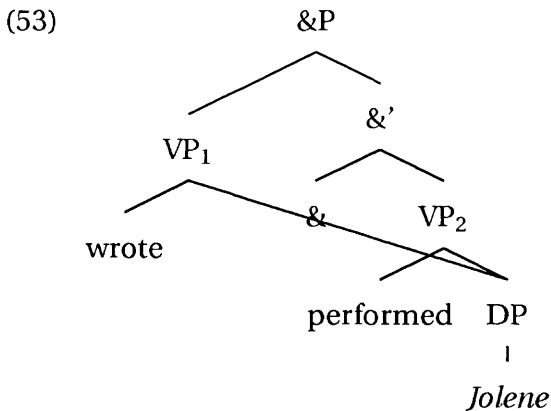
Complete dominance can also be defined using the notion of *path*. Suppose that a phrase A appears as a constituent of some larger structure S. A path P from A to the root to S is a n -tuple $\langle X_0, X_1, \dots, X_n \rangle$ where $A = X_0$, $S = X_n$, and for every member of the tuple X_m, X_{m+1} is a mother of X_m .¹⁰ Within a multidominant structure, there might be a number of paths leading from a single phrase to the root. We can now define complete dominance in terms of quantification over paths:

(52) **Complete dominance**

A completely dominates B in some structure C iff A is included in every path from B to C.

I believe that both ways of defining the notion are equivalent, but include both in case the reader finds one easier to work with than the other.

To see how complete dominance works more intuitively, consider the multidominant structure below:

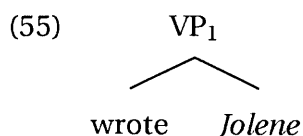


(54) VP_1 completely dominates *Jolene* in $\&P$: \mathcal{X}

Here there are two paths leading from the pivot to the root, $\langle DP, VP_1, \&P \rangle$ and $\langle DP, VP_2, \&', \&P \rangle$. Since some but not all paths includes VP_1 , VP_1 partially but does not completely dominate the pivot.

¹⁰Motherhood may in turn be defined in terms of Merge. If $C = \text{Merge}(A,B)$, then C is the mother of A and B.

Determining whether two nodes are in a complete dominance relation requires that we look at some larger structure. While VP_1 fails to completely dominate *Jolene* with respect to the entire coordinate structure, it obviously does if we restrict our attention to the substructure consisting of only VP_1 itself:



(56) VP_1 completely dominates *Jolene* in VP_1 : ✓

This can be made even clearer if we define a notion of *complete dominance domains* (CDD), the set of all nodes which a given node completely dominates with respect to some larger structure. As seen below, the definition of the CDD for a given node must taken into account two arguments, the node itself and some larger structure containing the node:

(57) *The complete dominance domain of A in B* ($CDD(A,B)$) = {*a*: A completely dominates *a* in B}

The complete dominance domain of VP_1 with respect to VP_1 will differ from that of VP_1 with respect to &P:

- (58) a. $cdd(VP_1, VP_1) = \{\text{wrote, } Jolene\}$
 b. $cdd(VP_1, \&P) = \{\text{wrote, } \cancel{Jolene}\}$

2.4.2 Linearization: The basics

Following a long tradition (Gazdar et al., 1985; Chomsky, 1995, a.o.), I assume that the rule system which determines structural relations is separate from the rule system which determines linear precedence relations. As was discussed in **section 1.2**, I assume that structural relations are determined by the rule Merge.¹¹ A linearization algorithm then applies to this unordered syntactic object.

¹¹In this chapter, it is unnecessary to invoke the fully articulated system, utilizing workspaces and the operation Collect, outlined in the last chapter. Those notions will play a role in chapter 3 however.

The linearization algorithm can be characterized by a function **Lin** which applies to a syntactic object and returns a *linear ordering* of terminal nodes within that syntactic object. A linear ordering is a binary relation which is *total*, *antisymmetric*, and *transitive*:

- (59) A relation **Lin**(SO) is *linear* iff for any distinct terms x, y, and z where SO dominates x,y,z
- a. $x < y \in \mathbf{Lin}(\text{SO})$ or $y < x \in \mathbf{Lin}(\text{SO})$
(*Totality*)
 - b. If $x < y \in \mathbf{Lin}(\text{SO})$, then $y < x \notin \mathbf{Lin}(\text{SO})$
(*Antisymmetry*)
 - c. If $x < y \in \mathbf{Lin}(\text{SO})$ and $y < z \in \mathbf{Lin}(\text{SO})$, then $x < z \in \mathbf{Lin}(\text{SO})$
(*Transitivity*)

Lin(S) must meet the following specifications. First, it identifies specific asymmetric structural relations between nodes of S. Let's designate such a relation R(A,B). For each node in this relation, the algorithm identifies some set of terminal nodes dominated by that node. Following Kayne (1994), I will call this set the *image* of a node X, and abbreviate it d(X). Finally, based on the relevant asymmetric structural relations, the algorithm determines that the image of one member of the relation is ordered before the image of the other. If R(A,B) then $d(A) \times d(B)$ must be included in **Lin**(S).

These specifications are consistent with a number of specific proposals. For instance, Kayne's (1994) *linear correspondence axiom* (LCA) designates *c-command* as the relevant structural relation. In GPSG (Gazdar et al., 1985) and other so-called ID/LP grammars, sisterhood relations between particular nodes, as stated in explicit *linear precedence* (LP) rules, are privileged. For the sake of concreteness, we will stipulate a set of such LP rules. Some representative rules are provided below:

- (60) For any pair of sisters A and B:
- a. HEAD RULE
If A is a head and B a complement, then A precedes B
 - b. SPEC RULE

If A is a specifier, then A precedes B

We are assuming a strictly compositional theory of linearization. Within a derivational framework, this amounts to the following assumption:

(61) **Compositionality of linearization**

Linearization is calculated locally at every derivational step

An algorithm consistent with this principle can be described as follows

(62) For every $C = \text{Merge}(A,B)$ in a derivation

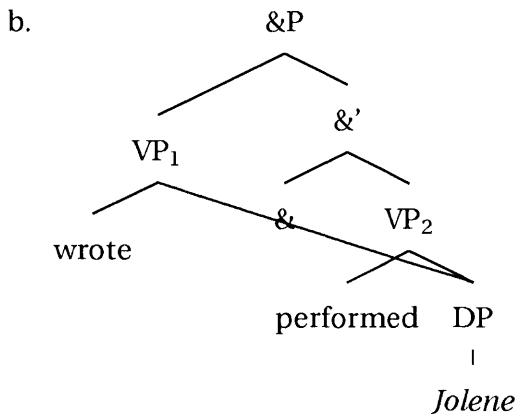
1. Add **Lin**(A) and **Lin**(B) to **Lin**(C)
2. If $A < B$ according to some LP rule, then add $d(A) \times d(B)$ to **Lin**(C)

In words, this says that for every derivational step in which A and B Merge to form C, the linearization of C consists of (i) the LP relations determined for A and B and (ii) a set of novel LP relations formed by ordering the image of A before the image of B.

2.4.3 Complete dominance and linearization

Consider again a simple RNR sentence (63a) and the structural description of its VP (63b):

(63) a. Dolly wrote _ and performed *Jolene*.



Notice that an attempt to define the *image* in terms of mere reflexive dominance will fail. To see why, suppose counterfactually that we defined the image of X as the set of all

terminal nodes which X reflexively dominates. In this case $d(VP_1) = \{\text{wrote}, \text{Jolene}\}$ and $d(VP_2) = \{\text{performed}, \text{Jolene}\}$. The linearization of &P will be:

- (64) a. $\text{Lin}(VP_1) = \{\text{wrote} < \text{Jolene}\}$
 b. $\text{Lin}(VP_2) = \{\text{performed} < \text{Jolene}\}$
 c. $\text{Lin}(\&P) =$
 $\text{Lin}(VP_1) \cup \text{Lin}(VP_2) \cup d(VP_1) \times d(VP_2) =$
 $\{\text{wrote} < \text{Jolene}\} \cup \{\text{performed} < \text{Jolene}\} \cup \{\text{wrote} < \text{performed}, \text{wrote} < \text{Jolene},$
 $\text{Jolene} < \text{performed}\}$

From the linearization of VP_2 , we decide that ‘perform’ precedes *Jolene*. From the linearization of &P, we decide that *Jolene* precedes ‘perform’, resulting in a violation of antisymmetry.

Following Wilder, suppose that we instead identify the image with all of the terms which a given node *completely dominates*:

- (65) **Wilder’s (1999) hypothesis**
 $d(X) \text{ within some } S = \text{CDD}(X,S)$

Of course, this hypothesis in itself does not provide us with a theory of the right edge restriction or right-node wrapping. While Wilder’s analysis (discussed more in **section 2.1.4**) invokes Kayne’s LCA, we will instead exploit our compositional linearization algorithm to derive these effects.

To see how this works, consider (63) again. The problem with the earlier theory was that ‘performed’ < *Jolene* was determined by the rule which says that verbs generally precede their complements, but *Jolene* < ‘performed’ was determined by the rule which places the image of the first conjunct before the image of the second, resulting in a contradiction. With Wilder’s modification, though, this problem disappears. Now the new LP relations introduced by $\text{Lin}(\&P)$ will not make reference to *Jolene* at all, since it is not completely dominated by either conjunct within the larger coordinate structure:

- (66) a. $\text{Lin}(VP_1) = \{\text{wrote} < \text{Jolene}\}$
 b. $\text{Lin}(VP_2) = \{\text{performed} < \text{Jolene}\}$

c. $\text{Lin}(\&P) =$
 $\text{Lin}(VP_1) \cup \text{Lin}(VP_2) \cup \text{cdd}(VP_1, \&P) \times \text{cdd}(VP_2, \&P) =$
 $\{\text{wrote} < \text{Jolene}\} \cup \{\text{performed} < \text{Jolene}\} \cup \{\text{wrote} < \text{performed}\}$

With this modification, our compositional linearization algorithm can successfully interpret RNR structures.

2.4.4 Deriving the right-edge restriction and the right-node wrapping generalization

In section 2.3, we stated two generalizations:

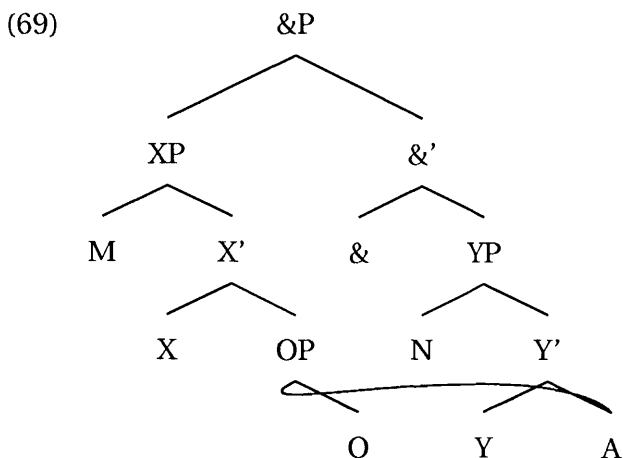
(67) **Right-edge restriction**

The pivot must be merged in the rightmost position of each non-final conjunct

(68) **The right-node wrapping generalization**

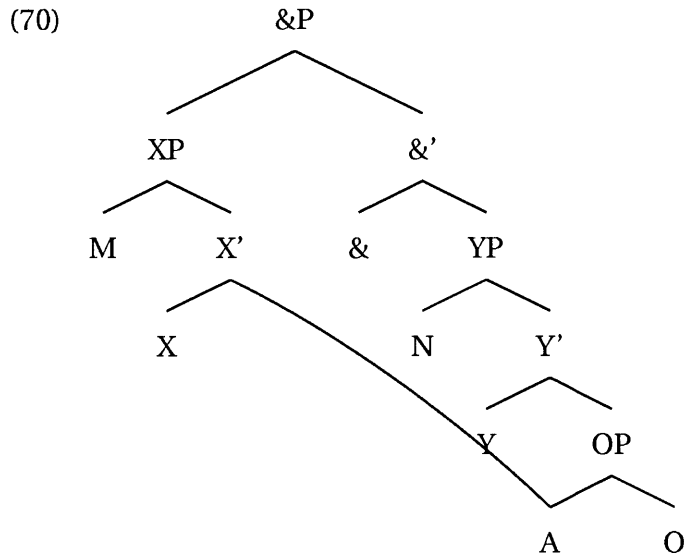
The pivot may be merged in a non-rightmost position of the final conjunct

Within the theory of linearization described above, the right-edge restriction follows because (i) the pivot is linearized within each conjunct, and (ii) all unshared material is linearized between two conjuncts. Consequently, a structure like (69) below, in which the pivot is not merged at a rightmost position within the initial conjunct, is predicted to be ungrammatical:



$A < O \in \text{Lin}(\&P)$ will follow from $\text{Lin}(OP)$. $Y < A \in \text{Lin}(\&P)$ will follow from $\text{Lin}(Y')$. $O < Y \in \text{Lin}(\&P)$ follows from the ordering of XP before YP. By transitivity, $O < Y$ and $Y < A$ entails that $O < A \in \text{Lin}(\&P)$. The result is not antisymmetric.

The proposal also derives the right-node wrapping generalization. This is illustrated by the following structure:



Assume that A is ordered before O in $\text{Lin}(OP)$. Unlike with (69), an ordering contradiction will not arise.

The difference between (69), in which the pivot is non-rightmost in a non-final conjunct, and (70), where it is non-rightmost in the final conjunct, can be summarized as follows. In (69), the pivot is merged in a position within the first conjunct (XP) that will force it to be linearized to the left of some phrase O. Within the second conjunct, the pivot is linearized to the right of material which will in turn be linearized to the right of O. This necessarily results in a linearization contradiction, deriving the right-edge restriction. Notice that the contradiction follows from the *compositional* nature of the linearization procedure. It is only because the linearization algorithm forces the pivot to be linearized within the first conjunct that a problem arises in the first place. In the right-node wrapping structure (70), on the other hand, no linearization contradiction arises, because the material to the right of the pivot is not linearized to the left of anything else.

2.4.5 Left-node raising

Yatabe (2001) argues that sentences like (71) are the left-node raising analogs of RNR sentences in languages like English (Yatabe's example (3)):

- (71) ✓ **yonde** ageta hito to _agenakatta hito ga ita
 read-GER give-PAST person and _ give-NEG.PAST person nom be-PAST
 'There were people who gave (him/her) the favor of reading (it) (to him/her) and people who didn't.'

If we follow Yatabe in assuming that this is fundamentally the same phenomenon as RNR, the theory proposed here should predict it to be subject to the same edge restriction as right-node raising, with the only difference being that now material cannot precede the pivot in any non-initial conjunct. As Bachrach & Katzir (2017) observe, this is correct. For instance, in the sentence below the pivot **omoi** 'think' is merged in a position in the right conjunct preceding the expression *sukoshi wa* 'at least a little', running afoul of this left-edge restriction:

- (72) * (sukoshi mo) **omoi**-das-an-ai no ka sukoshi wa _-das-u
 (even a little) think-get-NEG-PRES nominalizer Q at least a little _-get-PRES
 no ka ga mondai da
 nominalizer Q NOM question be-PRES
 'The question is whether (you) don't recall (it) (even a little) or whether you recall (it) at least a little.'

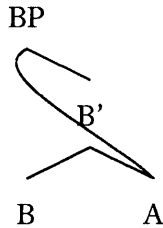
Moreover, we should also predict the availability of wrapping effects within the initial conjunct. As Bachrach & Katzir (2017) also observe, this prediction also turns out to be correct. For instance, the sentence below, in which *sukoshi wa* precedes the pivot within the initial conjunct, is well-formed:

- (73) ✓ *sukoshi wa* **omoi**-das-u no ka _-das-an-ai no ka
 at least a little think-get-PRES nominalizer Q _-get-NEG-PRES nominalizer Q
 ga mondai da
 NOM question be-PRES
 'The question is whether (you) don't recall (it) at least little or whether you recall (it).'

2.4.6 Internal merge

As mentioned in the introduction, internal Merge structures like (27), repeated below, pose a challenge to the assumption of compositional interpretation:

(74) **Internal Merge**



In this structure, A is simultaneously merged in the complement and specifier position of a phrase BP. If the LP rule governing complements places them to the right of the head and the rule governing specifiers to the left, then a compositional linearization procedure should derive an ordering contradiction for (74).

What needs to be recognized here is that movement structures pose a serious potential obstacle to interpretation at both interfaces. Any theory which allows movement structures to occur (whether via internal merge or something else) must also invoke some mechanism for letting the interfaces know that the movement structures exist. The interfaces in turn must have some set of rules for dealing with these structures and rendering them interpretable anyway.¹² At PF, this will include some means of recognizing that the internal merge structure constitutes a *chain*, and that the moved object is to be phonologically interpreted in only one position of that chain. What's crucial, from my perspective, is that these additional PF interpretation rules will have to create a structure which the linearization algorithm can successfully interpret.

In other words the structure in (74) is somewhat misleading. Despite the fact that both internal and parallel merge can be analyzed as involving multidominance, the suggestion here is that they in some important sense constitute a very different sort of operation from the multidominance involved in RNR and require additional interpretative

¹²At LF, this includes operations like predicate abstraction (Heim & Kratzer, 1998, a.o.) and, for copy theoretic structures, *trace conversion* (Fox, 1999). Just as I'm suggesting here for PF, these additional rules allow an essentially compositional semantic interpretation function to handle essentially non-local movement structures.

mechanisms to handle.

2.5 Bachrach & Katzir (2009, 2017)

As discussed in **chapter 1**, the theory proposed in this dissertation represents a response Bachrach & Katzir (2009, 2017). The purpose of this section is to outline their system. In addition to critically evaluating their analysis of the right edge restriction and right-node wrapping, I also hope this outline of the system will prepare the reader for further discussion of this comparison in chapters 3 and 4.

B&K's system resembles my own in outlining a compositional theory of linearization. It differs, however, in a number of respects. First, they propose that linearization is a process distinct from, and which applies prior to, Spellout. Furthermore, they assume that only Spellout is sensitive to complete dominance. Linearization of a phrase K will taken into account all nodes dominated by K , even if they are not completely dominated. In order to explain how successful linearization of pivot-internal structures is possible at all, they argue that the constraints on linearization are quite weak. Whenever two phrases A and B merge together, the leftmost terminal in A must be ordered before the leftmost terminal in B and the rightmost term in A before the rightmost term in B . When combined with further constraints on cyclic Spellout, this sort of grammar predicts the right edge restriction in a way which is empirically equivalent to the theory proposed in this chapter. The conceptual advantage of my own theory is that it reduces the the problem entirely to the linearization component, and not to a complex interplay between linearization and PF Spellout.

The basic architecture of their theory works as follows. Linearization is assumed to apply to all nodes in a syntactic structure, and linearity must be satisfied for every node. However, for every pair of sister nodes $C = [A B]$, the constraint governing the ordering of the nodes dominated by A and the nodes dominated by B is rather weak. Supposing that some law of precedence requires $A < B$ in C , the leftmost terminal of A is required to precede the leftmost terminal of B , and the rightmost terminal of A is required to precede the rightmost terminal of B . This forms part of their LINEARIZATION MAPPING

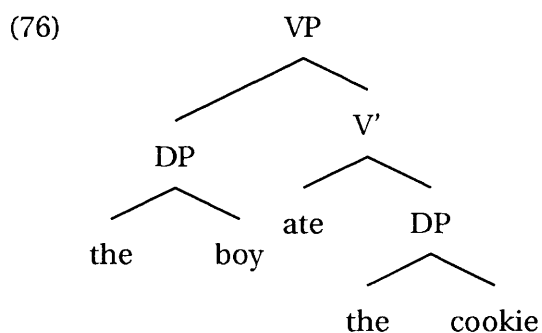
CONDITION:

(75) LINEARIZATION MAPPING CONDITION

- a. For any structure $A = [B C]$,
If $B < C$ in A according to some law of precedence,
then $\text{left-edge}(B) \leq \text{left-edge}(C)$ & $\text{right-edge}(B) \leq \text{right-edge}(C)$ ¹³
- b. For any A and B where A dominates B ,
 $\text{Lin}(B) \subseteq \text{Lin}(A)$
- c. $\text{Lin}(X)$ is *linear* (i.e. total, antisymmetric, transitive)

Other than this, the linearization component of the grammar places no restrictions on the orderings of the material within A and the material within B . Importantly, the linearization component is sensitive to reflexive dominance, not complete dominance, so shared material is subject to this constraint.

Notice that (75) alone overgenerates. Consider the structure below:



Assuming that $DP < V'$, (75a) imposes the requirement that $\text{left-edge}(DP) = \textit{the} < \text{left-edge}(V') = \textit{ate}$ and $\text{right-edge}(DP) = \textit{boy} < \text{right-edge}(V') = \textit{cookie}$. Crucially, however, no relation is established between *boy* and the terminal nodes dominated by V' . Hence linearizations like (77) will satisfy (75):

(77) * The ate boy the cookie.

To amend this problem, Bachrach & Katzir assume a theory of cyclic Spellout which requires adjacency relations to be preserved:

¹³ $\text{left-edge}(X) = x$ iff X dominates x and there is no y s.t. X dominates y and $y < x$. Vice versa for $\text{right-edge}(X)$.

(78) ADJACENCY PRESERVATION

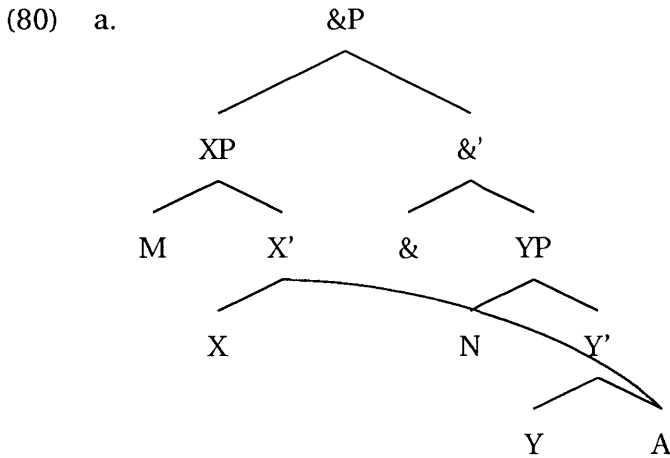
If an adjacency relation is established within a Spellout domain K, then it is preserved in every Spellout domain which contains K

In addition, a theory like Bachrach & Katzir’s must assume that all specifiers and adjuncts are Spellout domains.¹⁴ In the case of (76), the subject DP is by assumption a Spellout domain. By ADJACENCY PRESERVATION, *the* must remain adjacent to *boy*, correctly ruling out problematic orderings like (77). In addition, unlike linearization, Spellout domains are sensitive to complete dominance:

- (79) For some Spellout domain K in a structure S,
 $SOD(K) = CDD(K, S)$

Right node raising

Now consider a schematic RNR structure:



- b. $Lin(&P) = M < X < \& < N < Y < A$

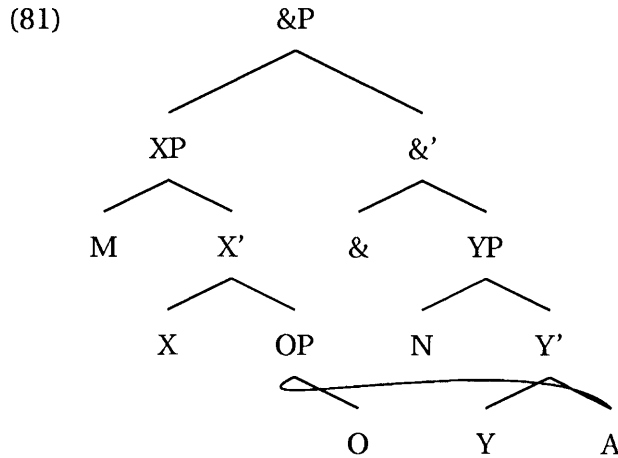
Bachrach & Katzir’s LINEARIZATION MAPPING CONDITION derives the ordering (80b). Importantly, the left edge of M precedes the left edge of &' and the right edge of both conjuncts is just A itself.

Notice that RNR is only possible because cyclic Spellout is sensitive to *complete* dominance. This means that, even if XP is a Spellout domain, ADJACENCY PRESERVATION will

¹⁴Bachrach & Katzir make the stronger assumption that For every Merge(X,Y), X or Y is a Spellout domain. But I think that the minimal necessary assumption is that all specifiers and adjuncts are SODs.

not require A to be adjacent to X in the final output string. If Spellout were sensitive to reflexive dominance not only would X have to be adjacent to A in the final string, but Y would also have to be adjacent to A, which would render an RNR structure unlinearizable.

Now consider a sentence which violates the right edge restriction:



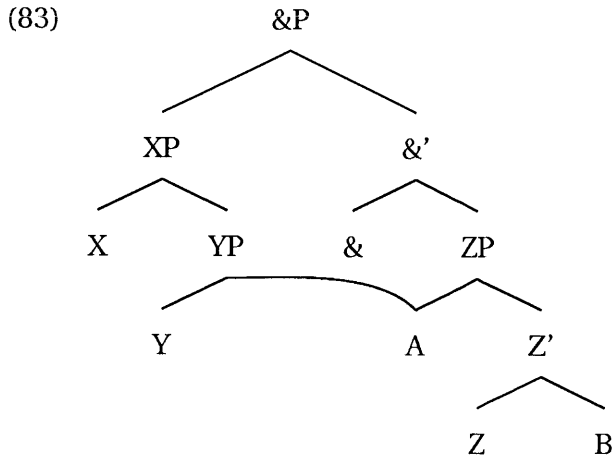
If we assume some LP statement 'A<O in OP', then A<O should be in the linearization of &P. Due to (75a), O<A should also be in the linearization of &P, since each is the right edge of the left and right conjuncts respectively. The result is a linearization contradiction.

Right node wrapping

Consider a grammatical and ungrammatical right-node wrapping structure:

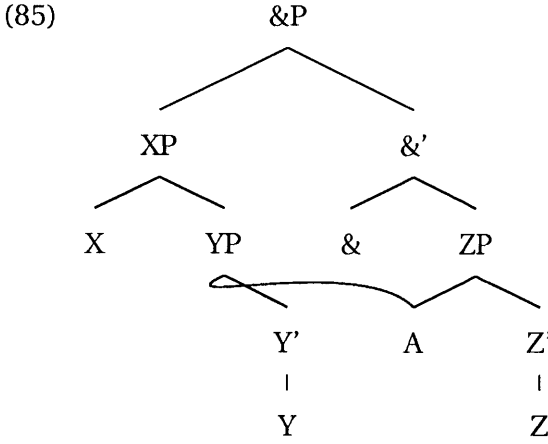
(82) **Grammatical right-node wrapping:**

John should congratulate , and then invite **Mary** to dinner.



(84) **Ungrammatical right-node wrapping:**

*John should give _ a present , and then invite **Mary** to dinner.



The grammaticality of (83) is correctly predicted by the LMC: crucially, the right edge of X, viz. A, is required to precede the right edge of ZP (B). Since this ordering is preserved in the right conjunct, a licit output is possible: X<Y<&<A<Z<B.

Things are slightly more complicated in the case of (85). Recall that in cases where the pivot is rightmost in the final conjunct, the right edge restriction follows from the fact that the the rightmost element in the left conjunct—O in (81)—had to precede A. Since A in turn preceded O in the left conjunct, a linearization contradiction was inevitable.

This does not follow for (85). Here, since the pivot is not rightmost in the final conjunct, the LMC does not require Y<A. Therefore, as far as the LMC is concerned, the following ordering should be licit:

(86) X<&<A<Y<Z

In words, as long as all of the material which A precedes in the left conjunct (Y above) is shoved into the right conjunct (between A and Z here) then a satisfactory mapping should obtain. In order to rule this out, the following assumption must hold:

(87) **Assumption regarding coordinate structures:**

For any right-node wrapping-compatible coordinate structure [A & B],
A is a Spellout domain, B is not

Applied to (85) this means that XP must be a Spellout domain, and ZP must not be. X will be adjacent to Y in Spellout(XP), since A is incompletely dominated here.¹⁵ Therefore ADJACENCY PRESERVATION will work to forbid an ordering like (86) in which Y is not immediately to the right of X. Crucially, though, the right conjunct may not be a Spellout domain in right-node wrapping structures, since Spellout of the right conjunct would have the effect of squeezing the pivot out of the conjunct internal order. I.e. if &' were a Spellout domain, then Spellout(&') would result in & being adjacent to Z, which the right-node wrapping ordering in which A intervenes between & and Z.

This results in a different version of the right-node wrapping generalization than either my own proposal or Wilder's predicts:

(88) **right-node wrapping generalization (B&K)**

For any coordinate structure [A & B] in which A is a Spellout domain and B is not, then the pivot may be non-rightmost in B.

Summary

Unlike my own proposal, this theory's account of the right-edge restriction in right node wrapping sentences depends crucially on a particular distribution of Spellout domains, stated in (87). It is unclear to me whether the necessary assumptions hold. Even a minimally more complex sentence should be ungrammatical, provided that it includes a

¹⁵A crucial assumption here is that the Spellout of XP will only include relations between nodes which XP completely dominates. So, while relations like 'X<A' are in the *linearization* of XP, they will not be included in the calculation of adjacency relations for the purposes of Spellout, since A is not completely dominated by XP. In the Spellout of XP, we will only have 'X<Y', meaning that X is adjacent to Y in Spellout(XP).

Spellout domain in the right conjunct. For instance, we might expect a sentence with TP coordination (90)—in which each conjunct contains a ν P/VP Spellout domain—to be significantly worse than its VP coordination counterpart (89), in which the right conjunct is not predicted to include a Spellout domain. I’m not sure this is the case, though my judgments are shaky:

- (89) a. The president gave an award to _ and then invited **the medalists** to dinner.
b. * The president gave _ an award and then invited **the medalists** to dinner.
- (90) a. <> The president gave an award to _ and the first lady invited **the medalists** to dinner.
b. * The president gave _ an award and the first lady invited **the medalists** to dinner.

In addition, because this proposal works by crucially weakening the theory of linearization, so that only edges are aligned with respect to each other, it must assume a theory of cyclic Spellout in which adjacency, rather than precedence, relations are preserved. It is therefore incompatible with Fox & Pesetsky’s (Fox & Pesetsky, 2005) theory of cyclic linearization, in which the conditions on Spellout are weaker. In this it again differs from my own proposal, which imposes no requirements on either the distribution of Spellout domains or the nature of cyclic Spellout.

However, it is worth pointing out that a crucial element of B&K’s theory, namely the availability of delayed Spellout in certain sharing structures, will play a significant role in the next chapter. Further discussion of how the theory in that chapter relates to B&K’s system and the analysis in this chapter appears in **chapter 3, section 3.7.3**.

2.6 Conclusion

This chapter has developed a strictly compositional theory of the linearization of multidominant structures. Following Wilder (1999), I have also assumed that linearization is sensitive to complete dominance relations. I have shown that—because complete

dominance is inherently sensitive to the domain of evaluation—the compositional linearization of RNR structures, when assigned a pivot internal multidominant analysis, can correctly account for the generalization that the pivot must be merged in a rightmost position in all non-final conjuncts (the right edge restriction) and that this requirement is obviated in the final conjunct (the right-node wrapping generalization).

Chapter 3

How to get off an island

3.1 Multi-gap dependencies and selective islands

3.1.1 Symbiotic and parasitic gaps

It is well known that some multi-gap dependencies ameliorate syntactic island effects. While the *parasitic gap* construction (Engdahl, 1983, a.o.) is certainly the best known of these (91), Bachrach & Katzir (2009; 2017) claim that across-the-board (ATB) movement also ameliorates island effects (92):

- (91) a. ?? That's the report that John fell asleep after reading _.
b. ✓ That's the report that John filed _ after reading _.
- (92) a. ?? Who did Gilligan talk to someone who admires _?
b. ✓ Who did Gilligan talk to someone who admires _ and Mary Ann go swimming with someone who can't stand _?

Example (92a) contains an instance of extraction which violates the complex NP constraint (Ross, 1967). The instance of ATB-movement in (92b) likewise violates the constraint, but is claimed to be acceptable. I will call this the *symbiotic gap* effect. Since the contrast in (92) is subtle for some speakers, the research reported here includes experimental confirmation of the effect.

I will follow Bachrach & Katzir in arguing that the symbiotic gap effect follows from a grammar which allows *delayed Spellout*. I will also extend delayed Spellout to the analysis of the parasitic gap effect. In other words: The parasitic gap effect is just a special instance of the symbiotic gap effect.

This analysis falls into the class of theories which seek to reduce parasitic gap formation to a variety of ATB-movement (Pesetsky, 1982; Williams, 1990; Nunes, 2004, a.o.). The primary goal of the chapter is to defend this analysis from classic arguments (Postal 1994) that parasitic gap formation is subject to restrictions, dubbed the *selective island* effects, not ordinarily present in ATB-movement. After showing that the full array of selective island effects also arises in symbiotic gap constructions, the main contribution of this chapter will be to argue that selective island effects arise from the mechanism of delayed Spellout itself.

Delayed Spellout works as follows. Spellout of a domain K only applies to constituents that K completely dominates, in the sense developed in the previous chapter. It follows, Bachrach & Katzir show, that Spellout of an incompletely dominated phrase will be *delayed* until the first Spellout domain that contains every position that phrase is Merged in. If syntactic islands arise when linearization of a moved element both inside of and outside of an island generates an ordering contradiction (as in Fox & Pesetsky's 2005 theory of *cyclic linearization*), delaying PF Spellout of that phrase until it has exited the island avoids that contradiction, correctly predicting island amelioration, as Bachrach & Katzir show.

In order to account for selective island effects, I will generalize Bachrach & Katzir's proposal, arguing that delayed Spellout has effects at both PF *and* LF Spellout. This generalization of Bachrach & Katzir predicts an array of semantic consequences for LF Spellout that correlate precisely with the consequences of delayed PF Spellout. In particular, when combined with Johnson's (2012, 2014) proposal that successive cyclic DP movement involves NP sharing, this extension predicts both the existence and distribution of the selective island effects. Since this analysis can account for selective island effects in both symbiotic gap and parasitic gap structures, it opens the way for a unification of ATB-movement and parasitic gap formation (contra Postal).

However, as will be demonstrated presently, this theory seems to overgenerate. Because, as we shall see, Johnson's proposal effectively invokes something like parallel Merge in every movement structure, whether it involves a single- or multi-gap dependency, delayed Spellout, and hence island obviation, should be predicted in every kind of movement structure. Put more bluntly, there should be no such thing as an island.

However, I believe that this problem is not fatal. Indeed, as Cinque (1990) and Postal (1998) have also observed, the same set of selective island effects arise in single-gap dependencies as arise in multi-gap dependencies. I will therefore propose a constraint which disfavors application of delayed Spellout in single-gap dependencies, thus predicting the existence of a contrast between single-gap and double-gap dependencies that involve extraction out of an island environment. Delayed Spellout in single-gap dependencies remain as a marginal possibility, explaining the selective island contrasts in these constructions.

This proposal does not itself offer a theory of islandhood. Instead, given the assumption that certain Spellout domains constitute islands for extraction, it predicts under what conditions islandhood is ameliorated, in particular the symbiotic and parasitic gap effects, and why the selective island restrictions should persist even under these conditions. An independent theory, not developed here, will be needed to specify the structural configurations that yield island effects in the first place.

3.1.2 Selective islands

Selective island effects are a class of environments which resist extraction out of islands even more than usual. While these effects appear in all manner of single-gap syntactic island environments — CNPC, adjunct, *wh*, etc. — they are notable for also appearing in multi-gap dependencies which otherwise filter out island effects, in particular parasitic and symbiotic gap sentences.

Anti-pronominality is one type of selective island effect. Certain syntactic environments forbid pronouns (93):

- (93) a. ✓ Gilligan speaks Hungarian, and Mary Ann speaks it, too.

- b. * Gilligan speaks Hungarian, and Mary Ann speaks in it, too.

Postal (1993, 1994) observes that those environments which forbid pronouns also block parasitic gap formation (94a), but not ordinary ATB-movement (94b):

- (94) a. That's the language that the professor wrote a grammar of _ after Gilligan spoke ✓ (*in) _.
b. That's the language that the professor wrote a grammar of _ and Gilligan speaks (in) _.

Based on this sort of observation, Postal argues that PG formation and ATB-movement involve distinct syntactic mechanisms which leave different sorts of material in the gap site.

However, anti-pronominal environments also block symbiotic gap formation (95):

- (95) That's the language that the professor wrote a grammar of _ and Gilligan knows someone who speaks ✓ (*in) _.

Indeed, as Postal also observes, the same contrast shows up even in single-gap dependencies:

- (96) That's the language that Gilligan knows someone who speaks ??(*in) _.

What (95) and (96) show is that the relevant distinction is not parasitic gap formation vs. ATB-movement, but instead whether movement occurs out of an island.

In addition to anti-pronominality, I will consider two other kinds of selective island effects: the observation (Cinque, 1990; Postal, 1998) that DPs are more amenable to extraction from islands than non-DPs, and scope reconstruction asymmetries which appear in island environments (Longobardi, 1987; Cinque, 1990; Ruys, 2011). The primary goal of this chapter is to explain the fact, documented more fully in **section 3.2.2**, that these restrictions arise in parasitic/symbiotic gap sentences as well. The secondary goal is to explain why the same contrast arises in single-gap dependency sentences.

3.1.3 The proposal

This dissertation will follow a tradition (Postal, 1966; Abney, 1987; Stanton, 2016) which holds that pronominals are syntactically bare determiners, lacking an NP complement:



Following Postal and Stanton, we will assume that an anti-pronominal context is a syntactic environment which forbids bare determiners like (97a).

As mentioned above, Bachrach & Katzir argue for a cyclic Spellout algorithm which allows Spellout of shared material to be delayed. This follows from the assumption that the Spellout domain of a phrase K , within a workspace W , is identical to the complete dominance domain of K in W . In addition, I will extend Bachrach & Katzir's model by arguing that cyclic Spellout returns both a PF and LF interpretation (Chomsky, 2000). Delayed Spellout therefore affects both LF and PF interpretation. That is, if a phrase avoids Spellout in a given syntactic position, it will not only fail to be pronounced in that position but crucially will also avoid semantic interpretation there.

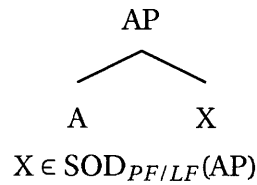
(98) **Spellout and complete dominance**

$$SOD_{LF}(K,W) = SOD_{PF}(K,W) = CDD(K,W)$$

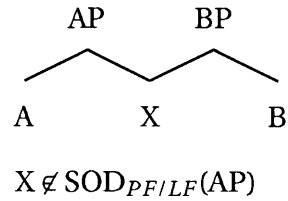
A node X is only visible for PF and LF Spellout, according to (98), if the root of the Spellout domain (SOD) dominates X in *every* position that X is merged. This is illustrated below:¹

¹Unless otherwise stated, all phrases in an example structure are assumed to occupy the same workspace.

(99) a.



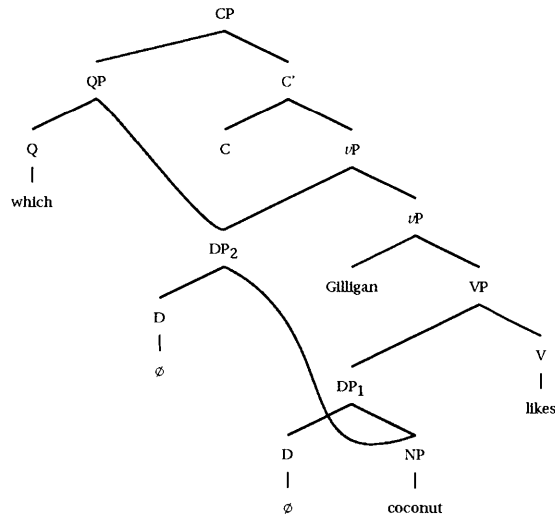
b.



If a phrase is shared both inside of and outside of a given SOD, as in (99b), that phrase will avoid Spellout within that SOD. When combined with the hypothesis that island-hood is related to cyclic Spellout (Fox & Pesetsky, 2005, 2007, a.o.), this predicts that movement out of an island, if achieved via delayed Spellout, should (i) permit island obviation and (ii) forbid semantic interpretation of the moved phrase within the island. In other words, delayed Spellout should correlate with the absence of what are sometimes termed ‘reconstruction’ effects within the island.

Johnson (2012, 2014) argues that movement structures can be built up in a manner which exploits syntactic sharing. For instance, according to Johnson, the phenomenon typically described as successive cyclic movement of a DP in reality involves sharing of a single NP between two null determiner heads. An example of this is provided below:

(100)

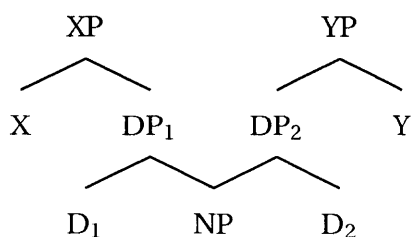


The advantage of this view of the grammar is that it allows a multi-dominant syntax to generate movement structures whose lower position is interpreted as a variable and higher position as an operator, ensuring interpretability at LF. What’s important for our purposes is that successive cyclic movement is taken to be NP sharing. That is, in the

above structure, movement from the base position to the edge of νP is accomplished not by standard internal Merge, but instead by sharing an NP between distinct determiners, forming two DPs. One DP is merged at the base position, the other at the edge.

Consider now what happens when Bachrach and Katzir's delayed Spellout mechanism is integrated into Johnson's theory of movement. In cases traditionally analyzed as DP movement, the moved element is actually the NP complement. For cases of ATB-movement, this will involve a configuration like the following:

(101)



A PF ordering paradox will be avoided so long as Spellout of the moved phrase, in this case the NP, is delayed. However, because LF Spellout is also delayed, this entails that the NP will not be interpreted within the island at LF or PF, leaving the determiner the only interpreted object in that position. Since a bare D is structurally identical to a pronoun, movement of a DP out of an island is both PF and LF interpretable — as long as the gap site is not a pronoun-rejecting context. On this view, both PF island amelioration and anti-pronominality effects crucially arise from the same mechanism: Delayed Spellout. I will also show how the other selective island effects under discussion, DP/non-DP asymmetries and scope reconstruction asymmetries, follow from the same array of factors.

My main goal is to explain the distribution of selective island effects in multi-gap dependencies. However, this leaves the question unanswered of why single-gap dependencies are subject to the same array of selective island effects. I believe that my analysis also sheds some light on this problem. As mentioned, Johnson's theory entails that every movement step will involve some sharing. Syntactic sharing entails the possibility of delayed Spellout, and delayed Spellout of course entails island amelioration. Thus, in this theoretical universe, it is always *possible* to get off an island. The differences in

acceptability between single-gap and multi-gap dependencies then come down to the assumption that the grammar countenances delayed Spellout in conjunction structures (symbiotic gaps) and adjunction structures (for parasitic gaps) but not for other kinds of phrases. This issue is discussed more fully in **Section 3.5.6**.

3.2 Data

3.2.1 Symbiotic gaps

Certain multi-gap dependencies mitigate island effects. Besides the parasitic gap construction, Bachrach & Katzir (2009) observe that their delayed Spellout mechanism predicts that ATB-movement should ameliorate island effects. These constructions are illustrated in (91) and (92), repeated below:

(102) a. ?? That's the report that John fell asleep after reading _.

b. ✓ That's the report that John filed _ after reading _.

(103) a. ?? Who did Gilligan talk to someone who admires _?

b. ✓ Who did Gilligan talk to someone who admires _ and Mary Ann go swimming with someone who can't stand _?

Again, I will refer to sentences like (103b) as symbiotic gap constructions.

While the parasitic gap phenomenon is well established in the literature, and has even been subject to experimental study (Phillips, 2006), the empirical status of the symbiotic gap effect remains more controversial. While it has been discussed sporadically (Bachrach & Katzir, 2009, 2017; Fox & Pesetsky, 2009; Larson & Parker, 2013), judgments contrary to (103b) are found elsewhere (Pesetsky, 1982; Munn, 1992).

In order to clarify the empirical status of this class of sentences across speakers, I verified the symbiotic gap prediction with a study on Mechanical Turk using Erlewine & Kotek's (2015) *Turktools*. There is a potential confound which makes direct comparison of the two sentences in (103) difficult: It is possible that a sentence like (103b) will be judged as less acceptable than its simpler counterpart (103a) simply because it is

more syntactically complex, thus obscuring any island amelioration effect.² To remedy this, I employed a 2×2 design with two factors. First, the presence vs. absence of ATB-movement; second, the presence vs. absence of an island. This paradigm is illustrated below:

- (104) a. That's the incident that the mayor read a report about _.
[-ATB, -island]
- b. That's the incident that the major read a report that was written about _.
[-ATB, +island]
- c. That's the incident that the TV journalist reported on _ and that the mayor received a report about _.
[+ATB, -island]
- d. That's the incident that the TV journalist reported on _ and that the mayor received a report that had been written about _.
[+ATB, +island]

The island effect in question is a contrast between full and reduced relative clauses, first discussed by Ross (1967).

If there is an island effect in both coordinate and non-coordinate sentences, then one expects a main effect due to the island, but no interaction. If the symbiotic gap effect is real, an interaction is predicted: the degradation associated with islands should be primarily driven by the non-coordinate sentences. The test involved a simple rating study, in which subjects were asked to read a sentence and judge its acceptability on a 7-point Likert scale. After testing 32 subjects, the z-score transformed ratings were analyzed using a multi-level mixed effects model.³ Crucially, the ATB×Island interaction was significant ($b=0.87$, $t(218) = 4.53$). Notably, while a degradation in acceptability was found to be associated with ATB-sentences in the [-Island] condition, in the [+Island]

²I know of one experimental study which examines symbiotic gap sentences: Larson & Parker (2013). While this study is interesting in other respects, its primary goal was not the confirmation of the symbiotic gap effect, and it does not take this confound into account.

³This model included random intercepts by subjects ($SD = 0.3298$) and items ($SD=0.0838$). The presence of an island significantly predicted a decrease in ratings, $b = -1.67$, $t(218) = -12.39$; the addition of ATB movement likewise significantly predicted a decrease in ratings, $b = -0.41$, $t(216) = -3.01$.

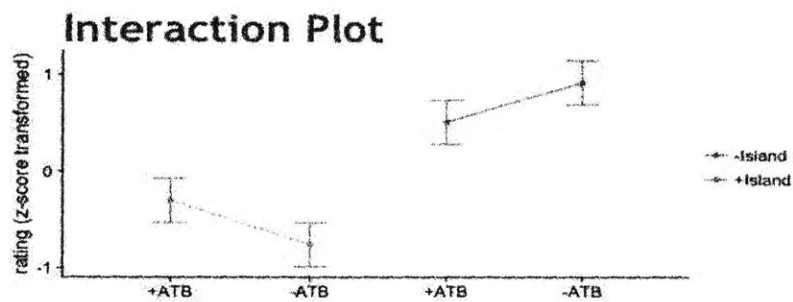


Figure 3-1: Interaction plot for CNPC island experiment

condition this pattern is reversed: [+ATB] is rated better than [-ATB] (see **figure 3-1**). This result supports the claim that the grammar incorporates Delayed Spellout.

A similar experiment, with 25 subjects, was conducted to test the symbiotic gap effect in sentences which violate the adjunct island condition. The items in this study involved the same 2×2 design. An example item set can be found below:

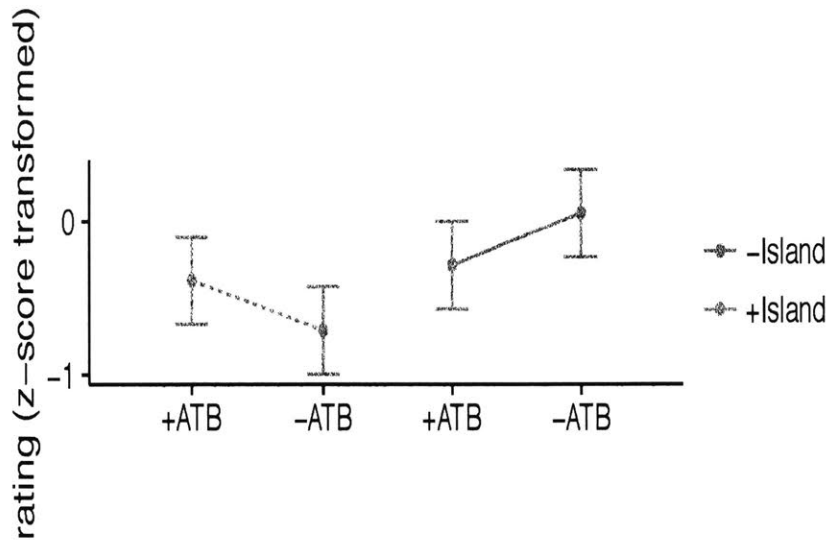


Figure 3-2: Interaction plot for Adjunct island experiment

- (105) a. That's the babysitter who the parent is angry that the toddler screamed at _.
[-ATB, -island]
- b. That's the babysitter who the parent is angry because the toddler screamed at _.
[-ATB, +island]
- c. That's the babysitter who the neighbor recommended _ and the parent is angry that the toddler screamed at _.
[+ATB, -island]
- d. That's the babysitter who the neighbor recommended _ and the parent is angry because the toddler screamed at _.
[+ATB, +island]

Again, the crucial interaction term was significant ($b = 0.6587$, $t(158) = 3.383$), and the interaction plot (**figure 3-2**) reveals a pattern much like the previous experiment.

In addition to CNPC and adjunct islands, informally gathered judgments from native English speakers lead me to believe that the symbiotic gap effect arises in other syntactic island types as well, such as subject islands⁴ and *wh*-islands (107):

⁴This observation differs from the prediction of Bachrach & Katzir's own model, which actually does not predict amelioration of subject islands via delayed Spellout. See Bachrach & Katzir (2009) and Fox & Pesetsky (2007) for further discussion of this matter.

(106) **Subject islands**

- a. ?? That's the book that people who love _ often turn into liberals.
- b. ✓ That's the book that people who love _ often turn into liberals and people who hate _ often wind up as conservatives.

(107) *wh*-island

- a. ?? That's the coconut that Ginger wonders who found _.
- b. ✓ That's the coconut that Ginger wonders who found _ and Mary Ann wonders who will eat _.

3.2.2 Selective islands

Although symbiotic and parasitic gap sentences generally ameliorate island effects, they are subject to a number of selective island restrictions. Although this is well established for parasitic gap sentences (Postal, 1993, 1994), I will show that the same restrictions hold of symbiotic gap sentences.

As discussed in the introduction, Postal (1993, 1994, 1998) observes that syntactic environments which forbid pronouns also forbids parasitic gap extraction. The relevant data is repeated below:

- (108) a. Gilligan speaks Hungarian, and Mary Ann speaks (*in) it, too.
b. That's the language that the professor studied _ after Gilligan spoke ✓ (*in) _.

Again, this restriction extends to symbiotic gap sentences. The relevant data is repeated below:

- (109) That's one language that Gilligan knows a man who speaks ✓ (*in) _ and Mary Ann also knows a woman who speaks ✓ (*in) _.

Stanton (2016) discusses several similar anti-pronominal contexts involving preposition-stranding. For instance, Stanton shows that certain locative prepositions forbid pronouns:

- (110) a. Gilligan swam in the lagoon and Ginger swam in it too.

- b. *Gilligan had lunch on the lagoon and Ginger had lunch on it too.

As we see in (110a), the preposition *in* is an instance of a locative preposition which accepts pronouns. As we see in (110b), *on* is not. Note that the preposition *on the lagoon* is actually ambiguous. Under one reading, it locates a position on the surface of the lagoon. On another, it locates a position along the lagoon's edge. It is this second reading which is unavailable in (110b). The first reading is pragmatically odd, though it is available if we imagine, for instance, that Gilligan and Ginger have some sort of raft which allows them to eat lunch on the lagoon's surface.

The examples in (112) illustrates how each environment behaves in cases of parasitic gap extraction, in these cases involving extraction from a subject island:

- (111) a. That's the lagoon that swimming in _ causes people to become frightened of _.
b. *That's the lagoon that having lunch on _ causes people to become frightened of _.

Again, this restriction also arises in symbiotic gap sentences, but not ordinary ATB-movement

- (112) a. *That's the lagoon that Gilligan knows someone who had lunch on _ and Mary Ann knows someone who swam in _.
b. That's the lagoon that Gilligan had lunch on _ and Mary Ann swam in _.

A similar pattern shows up when we construct examples using some of Stanton's other anti-pronominal contexts. Stanton discusses some pronominal restrictions involving temporal prepositions like *on Christmas eve*. First, we see that these prepositions generally forbid pronominal complements:

- (113) *Mary ordered Chinese takeout on Christmas eve and Fred ordered takeout on it too.

As with its locative counterparts, a temporal preposition cannot be stranded in an island context. (114a) illustrates this with extraction from an adjunct island:

- (114) a. *That's the holiday that Gilligan stopped celebrating _ after learning that he could just order Chinese takeout on _.
- b. That's the holiday that Gilligan stopped celebrating _ after learning about the origins of _.

Again, this restriction also holds of symbiotic gap sentences but not ordinary ATB-movement:

- (115) a. *That's the holiday that Gilligan knows someone who celebrates _ and Mary Ann knows someone who just orders Chinese takeout on _.
- b. That's the holiday that Gilligan celebrates _ and Mary Ann just orders Chinese takeout on _.

Additionally, Cinque (1990) and Postal (1998) observe that DPs (116a) are amenable to parasitic gap extraction, but non-DPs are not (116b):

- (116) a. Which country do people who who visit _ often write blog posts about _?
- b. *How sick do people who look _ often wind up _?

Again, this restriction applies to symbiotic gap sentences as well, but not to ordinary ATB movement:

- (117) a. *How sick did Gilligan talk to people who looked _ and Mary Ann speak with people who felt _.
- b. How sick did Gilligan look _ and Ginger feel _?

Finally, parasitic gap extraction blocks scope reconstruction:

- (118) a. How many people did you decide that I should talk to _?
- b. How many people did you call _ after deciding that Bill should talk to _?

Example (118a) allows two readings. Crucially, with of one of these readings it is not necessary that the person answering the question have specific people in mind. For instance, imagine that we are conducting a survey. In this case (118a) can have as an answer a certain number of people, without any associated inference that the answerer has specific people in mind. This reading disappears in (118b).

We again observe scope reconstruction effects in symbiotic gap constructions:

(119) How many people do you know a man who thinks I should talk to _ and Mary know a woman who thinks I should avoid _?

Example (119) only allows a wide scope reading for the *many* phrase with respect to the modal embedded within the island.

Finally, as was pointed out in the introduction, this range of selective island effects also arise in single gap dependencies (Postal 1998):

(120) **Anti-pronominality**

- a. ?That's the lagoon that swimming in _ causes strange rashes.
- b. *That's the lagoon that having lunch on _ causes strange rashes.

(121) **DP/non-DP asymmetry**

- a. ?Which country do people who visit _ often wind up in the hospital?
- b. *How sick do people who look _ often wind up in the hospital?

(122) **Scope reconstruction asymmetry**

- a. How many people do you think I should talk to _?
(✓ *how many* > *should*, ✓ *should* > *how many*)
- b. How many people do you know a man who should talk to _?
(?✓ *how many* > *should*, **should* > *how many*)

This dissertation will also seek to account for these data by arguing that delayed Spellout remains as a marginal possibility even in single-gap dependencies.

3.3 A theory of movement

3.3.1 Internal merge and its discontents

Johnson (2012, 2014) explores some of the consequences of a grammar with an operation like parallel Merge, which allows for phrases to be dominated by distinct, non-commanding mothers. This sort of grammar, Johnson shows, enables movement structures to be generated without some of the problems associated with internal Merge.

While internal Merge certainly offers an elegant approach to movement — a single syntactic operation is responsible for both movement and basic phrase structure building — IM also raises problems for semantic interpretation. How, given that the lower copy behaves semantically as some sort of variable and the higher as some sort of operator, can two sorts of semantic jobs be performed by one syntactic object?

Fox (1999, 2003) solves this problem with *trace conversion*, a post-derivational rule which converts the lower copy of a DP chain into a bound definite description at LF. Under Fox's account, trace conversion is a transformation which applies at LF to lower copies of DP-movement chains and has two effects: first, it deletes the lower determiner and replaces it with a definite determiner (123a); second, it adds a variable component to the DP which is bound by the higher copy (123b):

- (123) a. $[_{CP} \text{ [which coconut] }_7 [_{TP} \text{ Gilligan } [_{VP} \text{ eat [the coconut] }]]]$
 b. $[_{CP} \text{ [which coconut] }_7 [_{TP} \text{ Gilligan } [_{VP} \text{ eat [the ID-7 coconut] }]]]$

The lower copy will receive the following semantic interpretation:

- (124) a. $[[the]] = \lambda f: \exists!x[f(x)=1]. \iota x[f(x)=1]$
 b. $[[ID-7]]^g = \lambda x. x = g(7)$
 c. $[[[_{DP} \text{ the } [_{NP} \text{ [ID 7] coconut}]]]]^g$ is defined iff $\exists!x: \text{coconut}(x) = 1 \ \& \ x = g(7)$,
 Where defined, $= \iota x[\text{coconut}(x) = 1 \ \& \ x = g(7)] = g(7)$

The bound definite DP winds up denoting the entity introduced by $g(7)$, and introduces the presupposition that $g(7)$ is a coconut. The variable $g(7)$ can then be bound by the *wh*-phrase.

Fox assumes that trace conversion is a rule which applies at LF. Johnson instead argues that PM allows us to build the *effects* of trace conversion directly into the derivational syntax.⁵ Johnson's proposal simplifies the grammar, since trace conversion is no

⁵Johnson (2012, 2014) also raises the following conceptual objection to trace conversion. Fox's rule is designed to solve a problem which only exists for movement chains. However, there is no principled reason why one could not define a similar rule which effects the exact same structural change on any pair of *c*-commanding quantifiers. For instance, Johnson (2014: pg. 28-29) notes that one could imagine a grammar which applied a variant of the trace conversion rule to a sentence like (1):

(1) Every boy criticized no boy's mother.

longer needed as an independent rule.⁶ To see how this theory works, we first must be more explicit about how PM works. This is done in **section 3.3.2**. Then a version of Johnson's own proposal is presented. First we consider *wh*-DP movement (**section 3.2**). Next we discuss successive cyclic DP movement (**section 3.3.3**). Of particular importance for my own theory is Johnson's claim that DP movement involves sharing of an NP between *distinct* determiners that are also phonetically null. Finally, non-DP movement in Johnson's system is considered (**section 3.3.4**). This is important for my account of the DP/non-DP asymmetry discussed in **section 3.2.2**.

3.3.2 Parallel Merge: A review of how it works

As was already discussed in **section 1.2** of the introduction, a system which enables Johnson's structures to be built follows naturally from a grammar with the following properties. A derivation begins with a set of numerations (i.e. sets of lexical item tokens). Besides Merge, the grammar possesses a function Collect which combines the contents of two workspaces (Fox & Pesetsky, 2007). Finally, a standard version of external Merge (Chomsky, 1995) may apply between two syntactic objects within a workspace.

A key feature of the system is that only one kind of Merge exists. Movement structures are built from a sequence of parallel Merge operations. As was discussed in the introduction, parallel Merge is simply the application of external merge in two workspaces which happen to share the same constituent. Suppose we have a set of workspaces like

In this hypothetical grammar, Fox's rule is allowed to apply to the lower quantificational phrase *no boy*, converting it into a definite description bound by the higher quantifier *every boy*. The sentence would then have the interpretation *Every boy x criticized the boy x's mother*. This sort of thing is clearly unattested. Instead, the mechanism must be specific to movement. While for Fox this restriction to chains derived by movement is written into the rule's structural description, Johnson argues that it would be better to eliminate TC as a distinct transformation entirely, and to capture its effects in the theory of movement itself.

⁶Another important conceptual advantage of Johnson's system is that it allows us to maintain an extremely strong version of Chomsky's (1995) *extension condition* :

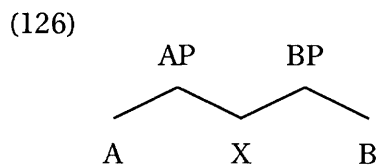
- (1) EXTENSION CONDITION:
Both arguments of Merge must be roots in some workspace

This version of the extension condition results in a highly restrictive grammar: while external Merge remains possible, internal Merge- which involves merger of a root syntactic object and a non-root which it dominates, is forbidden. However, parallel Merge is still possible, if it is understood as simultaneous external Merge.

the following:

- (125) a. $W_1 = \{X, A\}$
b. $W_2 = \{X, B\}$

Here, the same syntactic object X is shared between the two workspaces W_1 and W_2 . If $\text{Merge}(X,A)$ and $\text{Merge}(X,B)$ both apply, then X will be part of two syntactic objects $[_{AP} X A]$ and $[_{BP} X B]$. If Collect then combines the two workspaces, the result is a workspace in which X is shared. The resulting structure is graphically illustrated below:



Prior to application of Collect, when the two phrases occupy distinct workspaces, X will not count as shared between AP and BP. After application of Collect, X will be shared. In other words, a phrase only counts as shared between two or more positions if *each* position is within the same workspace. The mechanics of Collect will be important to the analysis in **section 3.5**; in particular, it allows us to derive the assumption that delayed Spellout is optional.

3.3.3 *Wh*-movement

With this view of syntax in place, we can build the effects of Fox's trace conversion directly in the derivational syntax. We can illustrate this with Johnson's approach to *wh*-movement.

For Johnson, the lower instance of the *wh*-chain is a bound definite description. I will assume this has a Foxian syntax: $[_{DP} D [ID-n N]]$. Note that this analysis of the DP differs somewhat from Johnson's implementation, who follows Elbourne (2005) in assuming that an index combines directly with the D, resulting in a structure like $[[D i] NP]$. This difference in the structure of the DP will have consequences for our own analysis in **section 3.5** (see footnote 14). The head of a chain is built by a Q morpheme (Hagstrom, 1998; Cable, 2007; Beck, 2006; Kotek, 2014, a.o.), which Merges with the DP.

Let us illustrate this system with a simple derivation of the sentence *Which coconut does Gilligan like?*. We begin with two workspaces:

- (127) a. $W_1 = \{ \textit{the}, \text{ID-7}, \textit{coconut}, \textit{Gilligan}, \text{T}, \textit{like}, \text{C} \}$
 b. $W_2 = \{ \text{Q}, \textit{the}, \text{ID-7}, \textit{coconut} \}$

Successive applications of Merge then form a bound DP in each workspace:

- (128) a. $W_1 = \{ [_{DP} \textit{the} [_{NP} \text{ID-7} \textit{coconut}]], \textit{Gilligan}, \text{T}, \textit{like}, \text{C} \}$
 b. $W_2 = \{ \text{Q}, [_{DP} \textit{the} [_{NP} \text{ID-7} \textit{coconut}]]$

Now the DP in W_1 Merges with the verb *like*, and the DP in W_2 Merges with the Q morpheme:

- (129) a. $W_1 = \{ [_{VP} \textit{like} [_{DP} \textit{the} [_{NP} \text{ID-7} \textit{coconut}]]], \textit{Gilligan}, \text{T}, \text{C} \}$
 b. $W_2 = \{ [_{QP} \text{Q} [_{DP} \textit{the} [_{NP} \text{ID-7} \textit{coconut}]]] \}$

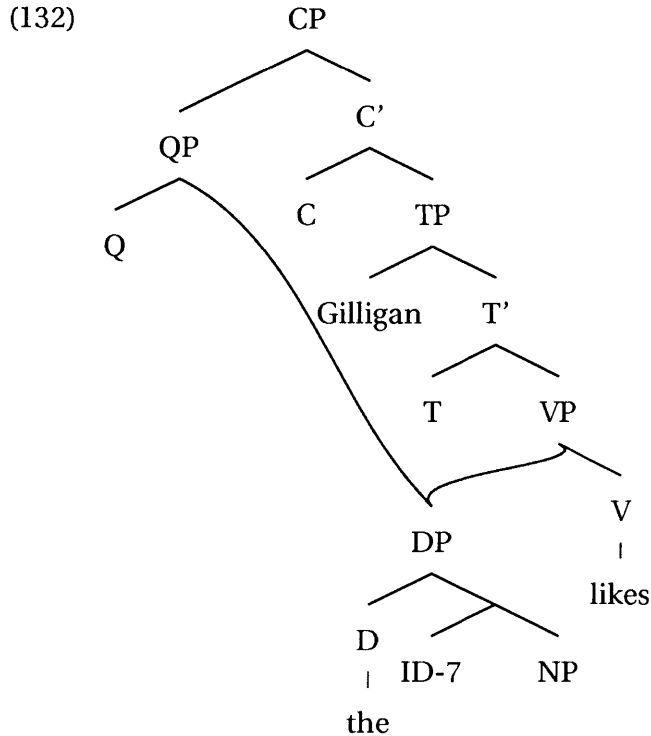
This is followed by a sequence of Merge operations within W_1 , building up a \bar{C} constituent.

- (130) a. $W_1 = \{ [_{C'} \text{C} [_{CP} \textit{Gilligan} [_{T'} \text{T} \text{T} [_{VP} \textit{like} [_{DP} \textit{the} [_{NP} \text{ID-7} \textit{coconut}]]]]]]] \}$
 b. $W_2 = \{ [_{QP} \text{Q} [_{DP} \textit{the} [_{NP} \text{ID-7} \textit{coconut}]]] \}$

Next Collect applies, forming a new workspace, W_3 , containing the elements of both W_1 and W_2 :

- (131) a. $W_3 = W_1 \cup W_2 =$
 $\{ [_{QP} \text{Q} [_{DP} \textit{the} [_{NP} \text{ID-7} \textit{coconut}]]], [_{C'} \text{C} [_{CP} \textit{Gilligan} [_{T'} \text{T} \text{T} [_{VP} \textit{like} [_{DP} \textit{the} [_{NP} \text{ID-7} \textit{coconut}]]]]]]] \}$

Finally, Merge of QP and C' can apply. The resulting structure can be graphically represented as follows:



The linearization of structures like (132) is determined by a set of interacting constraints. For instance, Johnson assumes a constraint CONTIGUITY, which has the effect of generally mapping phrases into contiguous strings. Competing is a constraint which requires the material dominated by QPs to be linearized to the left.⁷

The semantics of the QP will not be relevant to this dissertation. Johnson (2014) shows that this syntax is compatible with an analysis based upon Beck (2006) — in which the *wh*-DP introduces focus alternatives, which the Q morpheme converts into the ordinary semantic value — while Johnson (2012) assumes an analysis in which Q denotes an unrestricted existential quantifier (Heim, 2013).

⁷ From Johnson (2014, pg. 16):

(1) Contiguity

Let A be the set of vocabulary items dominated by some node A , and b a vocabulary item not in A . For every A in a phrase marker, if a linearization includes $b < a$ then it cannot include $A' < b$, for $a, b \in A$, and if it includes $a < b$, then it cannot also include $b < A'$, for $a, A' \in A$.

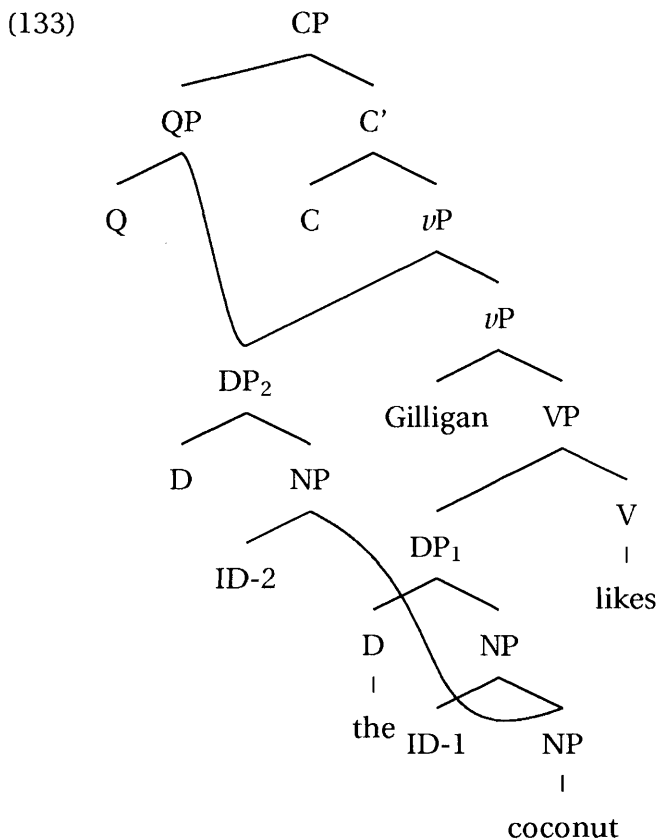
(2) The Wh-criterion

If a *wh*-QP is merged to a CP, then a grammatical linearization must position that QP so that it precedes everything in the QP.

3.3.4 Successive cyclicity

Successive cyclic movement follows a similar logic, except that rather than a QP parallel Merging with a DP, the movement step is accomplished by sharing a single NP between distinct (phonetically null) determiners.⁸ This assumption that it is the *NP*, rather than the full DP, which is actually involved with movement is especially important for my own proposal, since it entails that delayed Spellout will only target the nominal portion of the DP movement structure.

Consider *which coconut does Gilligan like* again. Supposing now that movement must pass through a *vP* edge, we assume that a D merges with NP, and then forms the complement of V. A second D also merges with the NP, before that DP remerges in a c-commanding position at the left edge of the *vP*. A Q morpheme will be merged with the highest DP:



The result is a structure which is both LF and PF interpretable. After my assumptions

⁸Johnson's version of quantifier raising works in the same way, except that the higher determiner is a quantifier. Successive cyclic movement is thus a case of QR feeding *wh*-movement.

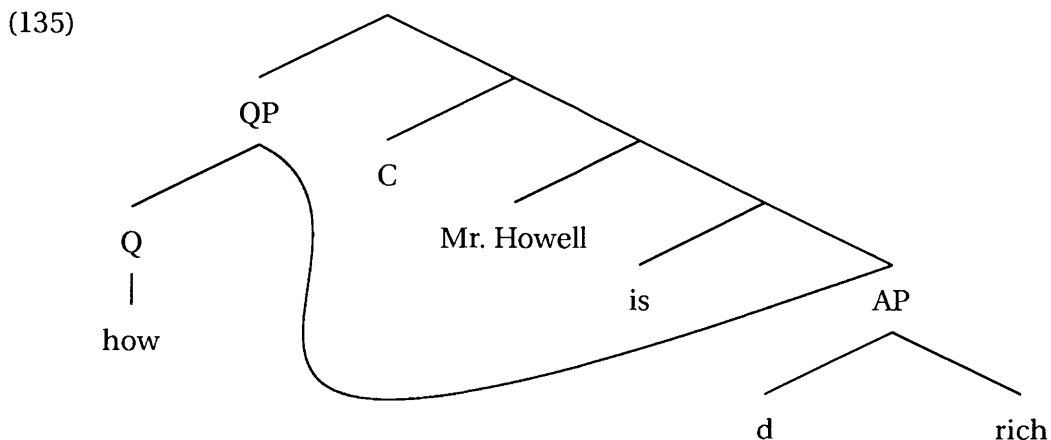
about cyclic Spellout have been introduced, **section 3.4.2** will discuss the LF and PF interpretation of this structure in more detail.

3.3.5 non-DP movement

Finally, we will consider how non-DP movement works in the system. Understanding this will be crucial to the account for the contrast between DP and non-DP movement out of island environments. The movement of a non-DP is illustrated below (134):

(134) How rich is Mr. Howell?

Johnson suggests a syntactic structure like the following for (134):



In this structure, Q merges with the AP, forcing the entire AP to effectively pied-pipe to the edge of the sentence at PF. Notice that, unlike successive cyclic DP movement, movement a non-DP must target the entire constituent. The syntactic material which made the lower DP interpretable as a variable, a phonetically null D and an variable function ID-*n*, would not be able to produce a semantically interpretable structure.⁹¹⁰

⁹If the definite determiner merged with an AP, it would produce a definite description, not a phrase with a predicative interpretation.

¹⁰A question remains of how successive cyclic movement works for non-DPs. It turns out that Johnson's system does provide a certain flexibility in which nodes are semantically interpreted. While it is required that a phrase be interpreted at some position, it does not need to be interpreted at *every* position it occupies. This allows for a derivation in which the moved non-DP merges into a higher position, but is interpreted only in its lower position. Each step of SC movement for a non-DP could involve merger of a semantically vacuous Q morpheme to the DP. This vacuous QP then would merge at the edge of a Spell-out domain.

3.4 Delayed Spellout

3.4.1 Cyclic Spellout and islandhood

A cyclic Spellout model proposes that aspects of phonological and semantic interpretation are computed cyclically at designated Spellout domains. For PF Spellout, we will assume Fox and Pesetsky's theory of *cyclic linearization*. According to Fox and Pesetsky, when Spellout applies to a given SOD, a set of linear precedence (LP) relations between nodes in that SOD is determined. Each subsequent round of Spellout will simply add to this set of LP relations. Fox and Pesetsky furthermore assume that structures whose linearization leads to ordering contradictions have an ill-formed PF interpretation. That is, for any two distinct nodes α and β , the set of LP statements may not contain both $\alpha < \beta$ and $\beta < \alpha$. This is illustrated below with a simple well-formed derivation:

- (136) What did Gilligan read =
[_{CP} **what** did Gilligan [_{VP} ~~what~~ eat ~~what~~]]
- Spellout(VP) = what < read
 - Spellout(CP) = what < did < Gilligan < read

Since the *wh*-phrase is moved to the left edge of each SOD, no ordering contradictions arise, and the structure has a well-formed PF interpretation. In general, for a phrase to move leftward it must be linearized at the left edge of every SOD it is Spelled out in.

An island, according to Fox and Pesetsky, is simply an SOD that disallows linearization of a moved element to its left edge:

(137) **Island for leftward movement of α**

An island for leftward movement is a Spellout domain that does not allow α to be pronounced at its left edge

(Fox and Pesetsky 2007, pg. 2)

I will assume that this definition can be extended to subject, adjunct, CNPC, and *wh* islands.

To see how this account works, consider a simplified case of extraction across an adjunct island SOD:

- (138) What did Gilligan sleep after reading? =
 [_{CP} what did Gilligan sleep [_{PP} after reading ~~what~~]]
- a. Spellout(PP) = after < reading < what
 - b. Spellout(CP) =
 what < did < Gilligan < sleep < after < reading
 (linearization contradiction 🧠)

Because the adjunct PP SOD forbids movement to its left edge, evaluation of the CP SOD will result in a contradictory ordering, and hence a malformed PF interpretation.

3.4.2 LF and PF Spellout

We will assume Fox and Pesetsky's theory, with one important addition: the syntax-to-semantics mapping is determined cyclically, along with the syntax-to-phonology mapping. While this is assumed in standard phase theory (Chomsky, 2000), it is not a direct concern of Fox and Pesetsky, or any works that build upon their theory (Ko, 2005; Sabagh, 2007, a.o.).

We will assume that cyclic Spellout must satisfy the following principles:

(139) **Spellout (PF):**

- a. For every $x, y \in \text{SOD}(X)$, the linear relation between x and y is determined at Spellout(X)
- b. The total set of linear precedence statements for a given derivation must not produce an ordering contradiction at Spellout(X)
- c. Once an LP relation is established, it cannot be altered by later iterations of Spellout.

(140) **Spellout (LF):**

- a. For every $x \in \text{SOD}(X)$, $\llbracket x \rrbracket$ is determined at Spellout(X)
- b. $\llbracket X \rrbracket$ must be semantically interpretable at Spellout(X)
- c. Once $\llbracket X \rrbracket$ has been determined, it cannot be altered by later iterations of Spellout.

While (139) simply encodes the principles of cyclic linearization which we discussed above, (140) adds the assumption that the semantic interpretation of any node is fixed cyclically, and that once the interpretation is fixed, it cannot be altered by subsequent iterations of Spellout.

3.4.3 Complete dominance and delayed Spellout

Following Chomsky (2000), phase theorists generally assume that Spellout applies to the c-command domain of designated *phase heads*, such as *v* or C. Bachrach and Katzir instead argue that dominance rather than c-command determines Spellout, and in particular that the Spellout domain of a phrase is restricted to the set of all nodes which are completely dominated by the root of that phrase. The definition of complete dominance already discussed in **section 2.4** is repeated below:

(141) **Reflexive dominance**

1. If $X = Y$, then X reflexively dominates Y,
2. If X reflexively dominates a mother of Y, then X reflexively dominates Y

(142) **Complete Dominance**

A node X completely dominates a node Y within some workspace W iff X reflexively dominates every mother of Y in W. The set of nodes completely dominated by X in W is called the **Complete Dominance Domain** of X in W, abbreviated CDD(X, W).

Now Spellout domains can be defined in terms of complete dominance domains:

(143) **Spellout Domain (SOD)**

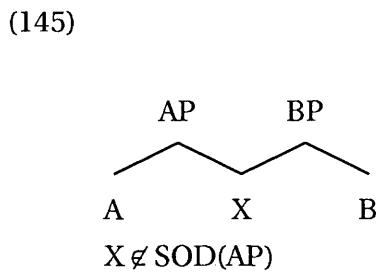
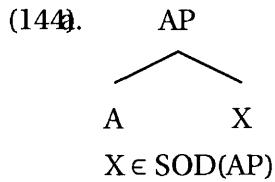
$SOD(X, W) = CDD(X, W)$, if X is a **Spellout node**.

Spellout Node

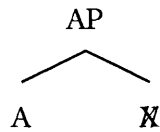
A designated syntactic object that triggers spellout of its spellout domain (= VP, CP, roots of specifiers and adjuncts).

Note the relationship between this proposal and the use of complete dominance in the analysis of right-node raising structures(**chapter 2**) will be discussed more fully in **section 3.7.3**.

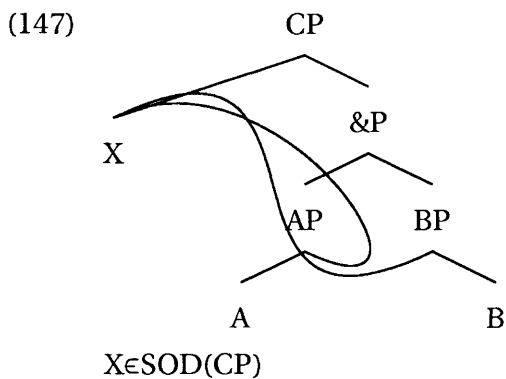
To illustrate how complete dominance and Spellout relate, consider the two syntactic structures below:



(146) **What Spellout sees:**



In (145) Spellout of X does not take place within either AP or BP, even though both are SODs, but is delayed until a point at which some SOD which does completely dominate X is evaluated. Consider a potential continuation of this derivation:



Now X is completely dominated by the CP SOD, allowing for it to be interpreted in its higher position at LF and PF. Notice that, since X avoided Spellout within AP and BP, it

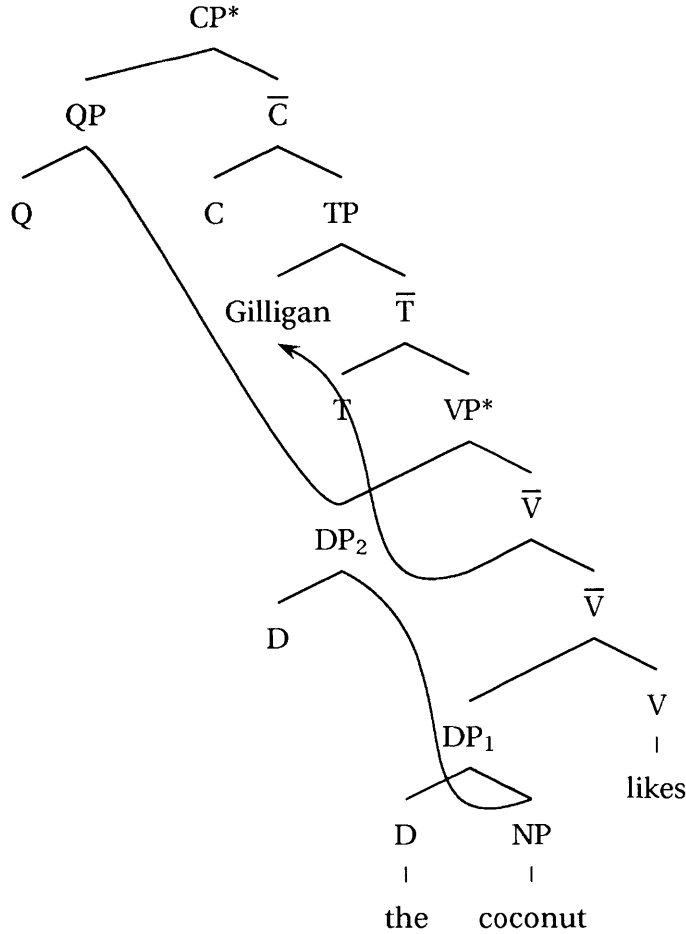
will only be linearized with respect to elements within these SODs once it has arrived at the edge of the CP. Consequently, if either of these SODs happen to be islands, the ordering contradiction which normally results from movement out of an island will be avoided — because Spellout of the moved phrase was evaded within the island SODs. This of course is just the symbiotic gap effect.

Crucially, complete dominance is defined *relative to a given workspace*. Consequently, if a phrase is shared between two positions in different workspaces, then that phrase will count as completely dominated within each workspace. If we assume that Spellout can apply either before or *after* Collect, it follows that there is optionality as to whether Spellout of shared material will be delayed. This is important to the explanation for why selective island effects arise in symbiotic gap sentences, but not ordinary ATB-movement.

3.4.4 A sample derivation: DP movement

Again consider the derivation for *Which coconut does Gilligan like*, now taking into account the theory of cyclic Spellout. The structure for the entire sentence will be as follows, with * marking the Spellout nodes:

(148)



Spellout will initially apply to the maximal VP projection. In order for the linearization algorithm to work properly, two assumptions about the linearization algorithm will be necessary:

1. a linearization constraint, similar to Johnson's *wh*-criterion (see footnote 7), which forces the material dominated by DP₂ to be pronounced at the left edge of the SOD
2. phonetically null determiners are not considered by linearization.¹¹

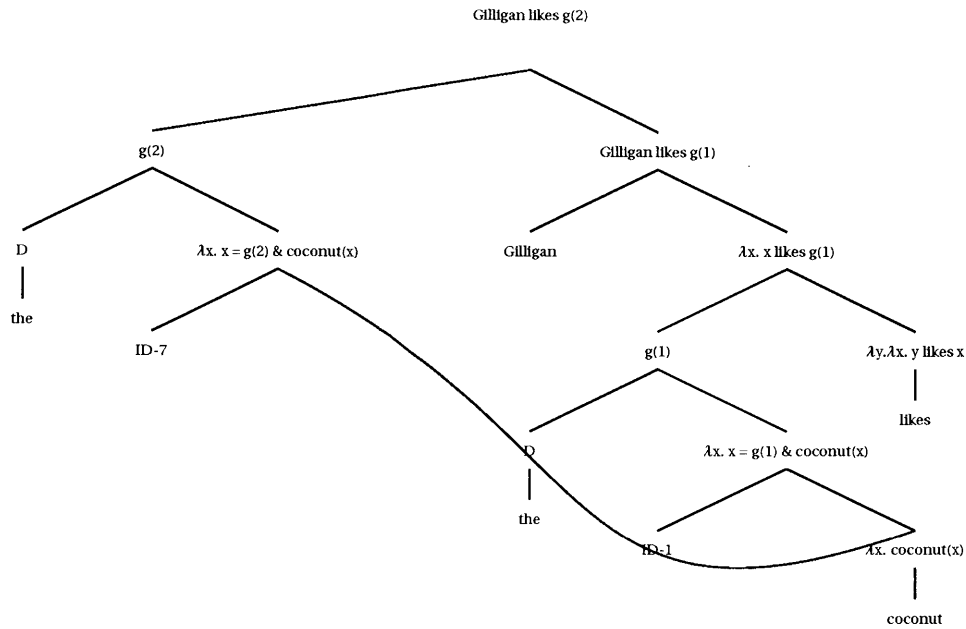
With these assumptions in place, we will arrive the following set of linearization statements:

(149) $Lin(VP^*) = \{\text{coconut} < \text{Gilligan}, \text{coconut} < \text{likes}, \text{Gilligan} < \text{likes}\}$

¹¹This assumption raises questions for cases where apparently null material is island-sensitive and therefore subject to linearization, e.g. null operators and null topics (Chomsky, 1977; Huang, 1984, a.o.). Discussion of this issue is beyond the scope of this dissertation.

This SOD will have the following semantic interpretation:

(150)



The next Spellout domain will be the matrix CP. It should produce the following linearization statements:

(151) $Lin(CP^*) = \{ \text{which} < \text{coconut}, QP < \text{Gilligan}, QP < VP, \text{coconut} < \text{Gilligan}, \text{coconut} < \text{likes}, \text{Gilligan} < \text{likes} \}$

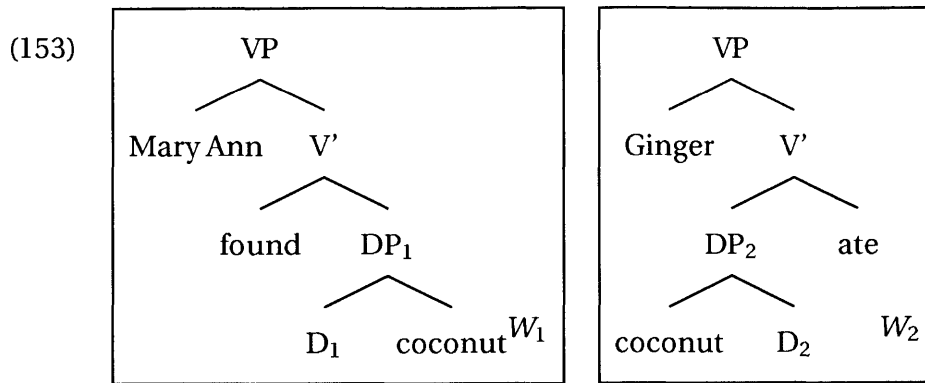
This set of LP statements is non-contradictory, and so causes no problems at PE. The semantic interpretation will proceed as follows. First the semantic value computed for VP* in the earlier round of Spellout will be retrieved, then this value will be used to compute the semantic value for the entire CP.

3.4.5 ATB-movement

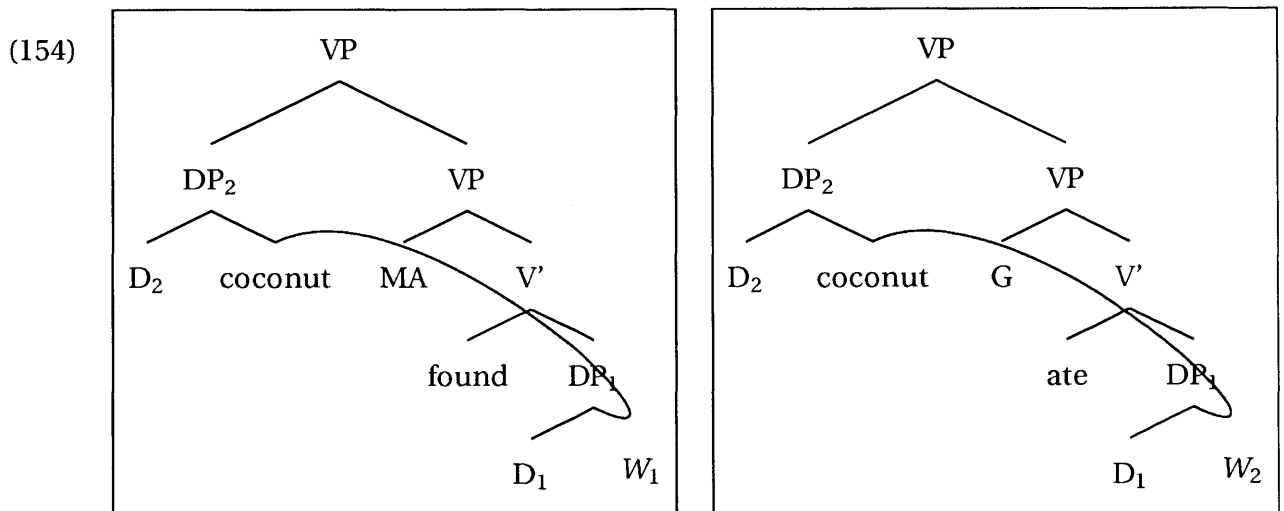
Ordinary ATB-movement will work as follows. Consider a sentence like (152):

(152) Which coconut did Gilligan said that Mary Ann found _ and Ginger ate _.

The ATB-structure is built up by parallel Merge of the NP *coconut* with distinct determiners. Each DP then combines with a V:



This is followed by successive cyclic movement of the NP to the VP edge:



Spellout of each VP delivers an ordering in which *coconut* precedes other material in the Spellout domain:

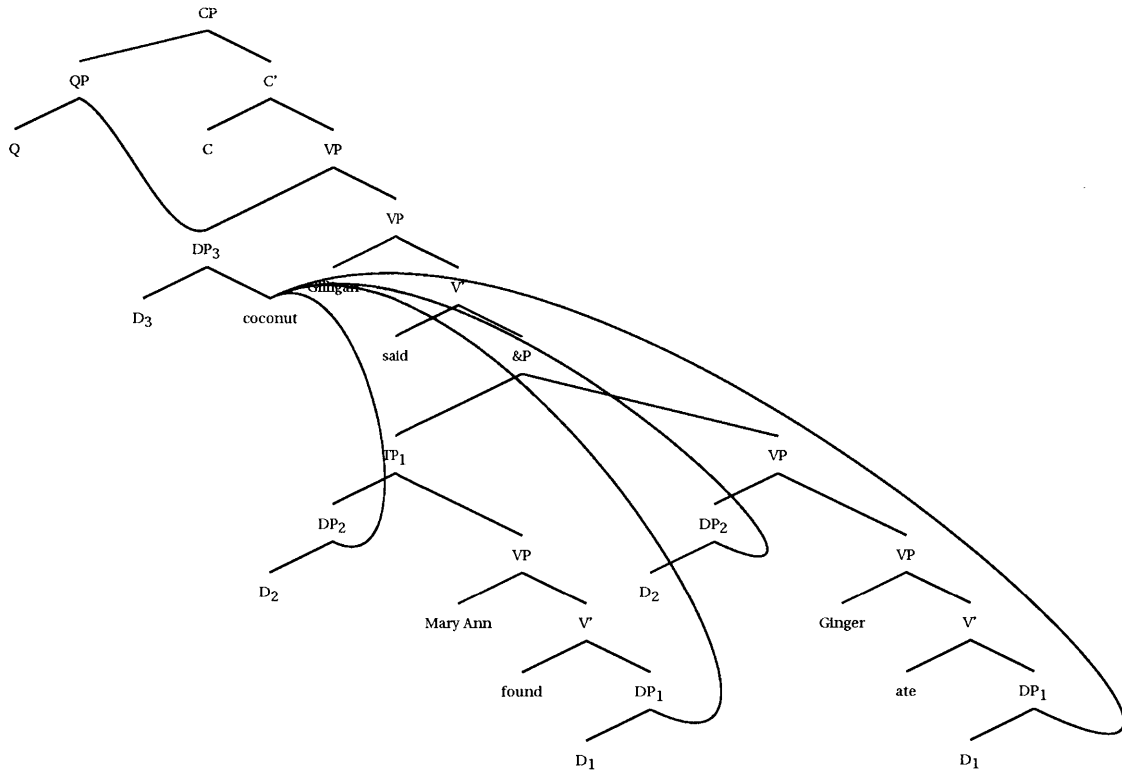
(155) $Lin(VP, W_1) = \text{coconut} < \text{Mary Ann} < \text{found}$

(156) $Lin(VP, W_2) = \text{coconut} < \text{Ginger} < \text{ate}$

Next, the entire coordinate structure is built up. An important assumption here is that, in cases of ordinary ATB movement, delayed Spellout is not necessary for movement out of a coordinate structure to result in a well-formed PF interpretation. That is, neither the coordinate structure or either of its conjuncts is inherently an island.¹²

¹²This analysis appears to be incompatible with a theory of coordinate structures which, like Munn (1993), takes the left conjunct to be the specifier of an &P, since we are also assuming that specifiers are island Spellout domains. However, the fact that, unlike other specifiers, left conjuncts are not generally subject to selective island effects might be taken as an argument against an such an analysis, and for the assumption made here that neither conjunct is inherently an island.

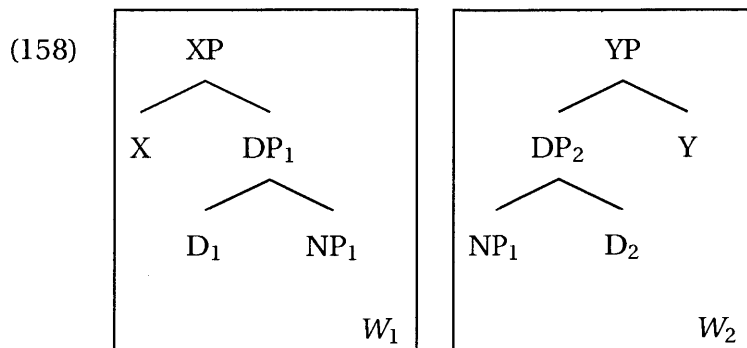
(157)



Within the higher VP and CP Spellout domains, the NP coconut will be linearized to the left of all material left inside of the coordinate structure. The result will be a well-formed linearization. The sentence is also semantically interpretable.

3.5 How to get off an island

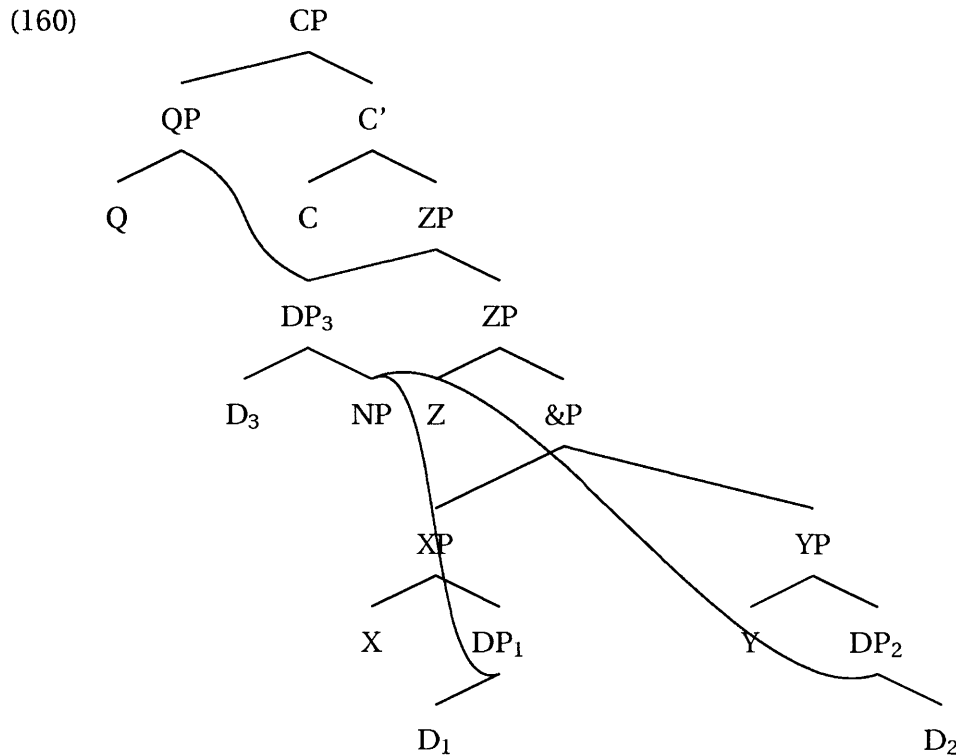
We are now ready to consider the derivation of symbiotic gap effects. As with the case of ordinary ATB-movement discussed in **section 3.4.5**, we first assume that a single NP is Merged with distinct, phonetically null determiners. Each DP is then Merged inside of some phrase XP and YP respectively:



Suppose that XP and YP are islands, and hence forbid movement to their left edge. At this point we have two options. First, Spellout may apply to each phrase at a point where it is occupying a different workspace. This will result in the following linearization:

(159) $Lin = X < NP_1, Y < NP_1,$

Now suppose we arrive at a point where the movement structure is complete:

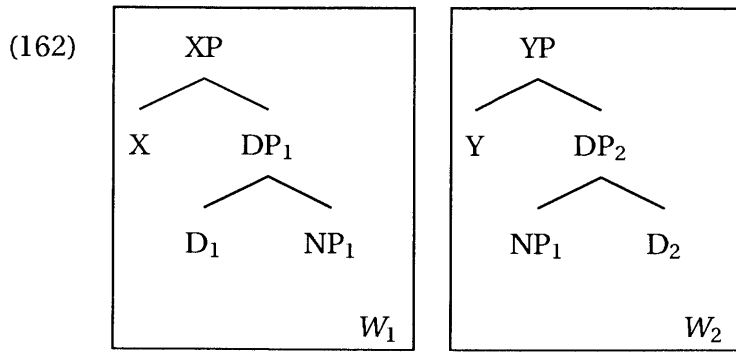


The linearization of SOD(CP) will order everything dominated by the QP, including the NP, to the left of X and Y. The result will contradict the ordering established in XP and YP:

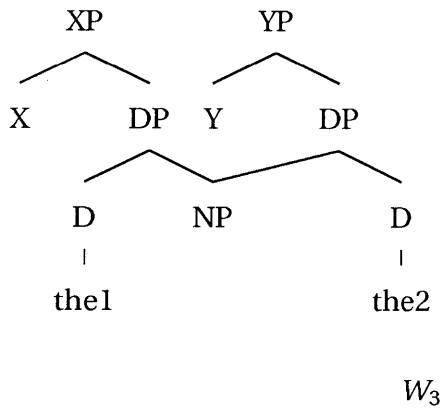
- (161) a. $Lin(XP) = X < NP$
 b. $Lin(XP) = Y < NP$
 c. $Lin(CP) = NP < Z < X < Y$

(linearization contradiction ☠)

However, an alternative derivation is possible. Imagine that, unlike the ordinary ATB-movement cases, Collect applies *prior* to Spellout of the XP and YP, so that the two DPs now occupy the same workspace:



COLLECT →



If Spellout applies to the phrases in this sort of workspace, Spellout of the NP will be delayed, since it is incompletely dominated within each SOD. Even if there is an island inside one or the other conjunct, movement can still take place, since Spellout of the NP is delayed. However, something is Spelled out: the two determiners.

With this we arrive at the major theoretical conclusion of this chapter. Delayed Spellout of the NP has two consequences. First, an island violation is avoided, since Spellout of the NP within the island is delayed. Second, delayed Spellout will result in an interpretable structure. Although the NP is not interpreted semantically, the island CP still receives a well-formed semantic interpretation, with the base-generated position of the movement structure interpreted as an unrestricted variable. Recall that the determiner denotes a function which applies to a type $\langle e, t \rangle$ predicate and returns the unique entity in the extension of that predicate. ID- n simply denotes the set of individuals identical to $g(n)$. In other words, it is a set consisting of a single individual, the one corresponding

to $g(n)$. $\llbracket [the\ ID-n] \rrbracket$ in turn will simply return that individual. We assume that this is identical to the semantics of an ordinary pronoun, except that pronouns typically also carry ϕ -features.

Since this semantic interpretation constitutes the LF output for cyclic Spellout of the island, this will be the only interpretation available for the final structure. Future rounds of Spellout will be incapable of changing it. More generally, we can conclude the following:

(163) A phrase which is moved out of an island may not be interpreted within that island

As we are about to see, this result will allow us to derive the full array of selective island effects discussed in **section 3.2.2**.

3.5.1 Anti-pronominality

The conclusion in (163) predicts anti-pronominality. The anti-pronominality effects observed in **section 3.2.2** show that the material left behind in extraction from island contexts differs from that left behind in non-island extraction: the syntactic object left behind in island contexts appears to have something in common with a pronominal, whereas the object left behind non-island movement behaves like a fully reconstructed DP.

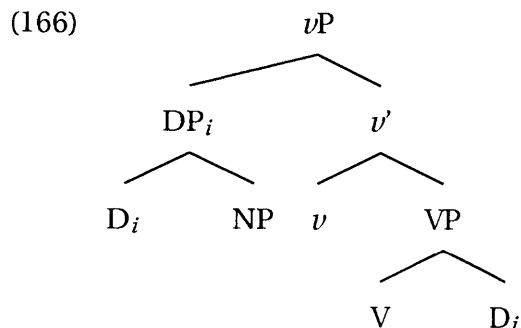
A similar contrast shows up elsewhere. Postal (1998) and Stanton (2016) notice that different types of \bar{A} -movement differ in their sensitivity to pronoun rejecting contexts. Some types of extraction, like topicalization, disallow movement out of pronoun rejecting environments:

(164) Hungarian, Mary refuses to speak (*in) _.

Others, like *wh*-movement and relativization, allow this sort of extraction (provided the movement is not out of an island, of course):

(165) That's one language that Mary refuses to speak (in) _.

Stanton (2016) provides an explanation for this contrast. Stanton argues that topicalization (and other types of movement with similar properties) forces *wholesale late merger* (Takahashi & Hulsey, 2009) of the NP complement of the moved DP at a site structurally higher than its base-generated position. As a consequence, the copy left behind in the base-generated position in this sort of configuration will be a bare D, with the full DP copy only appearing at a higher site:



For reasons not relevant to the present discussion, Stanton argues that in *wh*-movement and relativization merger of the NP necessarily occurs in the base-generated position of the chain. The copy in the base-generated position will be a full DP copy, NP complement included.

Following Postal (1966), Stanton argues that a bare determiner is structurally identical to a pronoun. Hence the lower copy in a case of topicalization might be expected to be excluded in anti-pronominal contexts.¹³ Conversely, the lower copy in *wh*-movement will be a fully reconstructed DP copy. Anti-pronominal sensitivity is not expected.

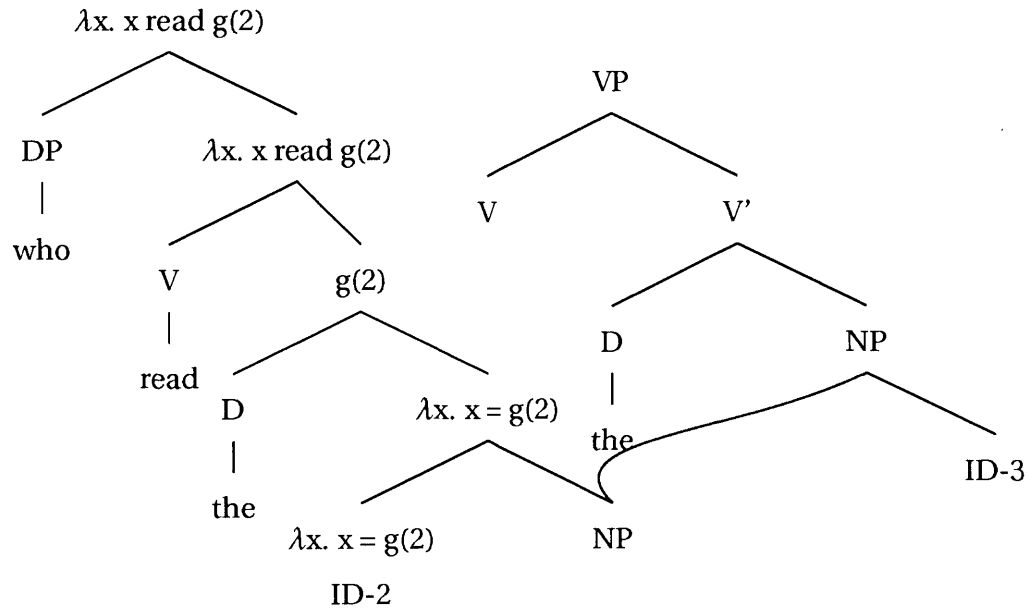
To put it in a slightly different way, Stanton argues that the grammar imposes certain restrictions on the reconstruction of the NP complement in a DP movement structure. This restriction in turn predicts anti-pronominality. We will adopt this basic insight of Stanton's paper:

(167) Anti-pronominal effects are related to the ability of the NP to be interpreted in the pronoun-rejecting site.

¹³While Stanton does not analyze the full range of Postal's anti-pronominality effects, she argues that, for pronoun-rejecting temporal and locative PPs, specific material in the NP must be present in order for the PP to receive the appropriate interpretation. Note that Stanton's examples of pronoun-rejecting prepositions, like *on the lagoon*, are generally ambiguous between a literal reading (where one is actually located on the lagoon), and an idiomatic one (where one is along the banks of the lagoon). It is this idiomatic reading which is rendered unavailable in cases like extraction from an island.

With that in mind, let us turn to the semantic interpretation of an island SOD when the NP is incompletely dominated:

(168)



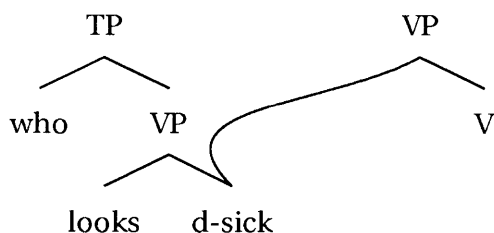
As we have seen, a cyclic Spellout algorithm which incorporates complete dominance into its definition of SODs will predict that the NP in this structure can avoid Spellout, and therefore avoid becoming trapped in an island. Moreover, the NP's avoiding Spellout will also prevent its reconstruction into the island. However, the phrase [_{DP} THE ID-2] will be Spelled out, and interpreted as a variable, in other words a D with no NP complement. This will have an effect identical to Stanton's use of WLM, except that the restriction on NP reconstruction now arises out of the interaction between a Johnsonian syntax and the Spellout algorithm, rather than the mechanics of Merge itself.¹⁴

¹⁴Note that this analysis is not easily made compatible with the specific analysis of the bound definite DP assumed by Johnson (2014), which has a structure inspired by Elbourne (2005): [[THE n] NP]. If the NP is not spelled out in this sort of structure, then the definite determiner will simply lack a type $\langle e, t \rangle$ complement to combine with, resulting in semantic uninterpretability. However, if we assume that referential pronouns combine with a phonetically null ONE predicate, as Elbourne (2005: pg. 95) suggests, then the grammar could generate a structure comparable to (168) in which the lower DP is replaced with [[THE n] ONE NP]. If the NP, but not ONE, is shared, then the Spelled out structure [[THE n] ONE] will be semantically interpretable. However, this variant of the analysis seems less attractive than the current proposal.

3.5.2 DP/non-DP asymmetry

We saw in **section 3.3.5** that Johnson's theory posits syntactic and semantic differences between DP and non-DP movement. Under my analysis, this is related to the DP/non-DP asymmetry in extraction out of islands. Suppose that we attempt to delay Spellout of a non-DP. Again, this would be accomplished by creating a workspace which the wh-AP is incompletely dominated within the island Spellout domain? This sort of structure would look like (169):

(169)



While this structure will avoid a linearization contradiction, it will also lack a well-formed semantic interpretation. In order for the delayed Spellout mechanism to produce a structure with well-formed semantic interpretation, there must be something Spelled out in the moved phrase's base-generated position which is semantically interpretable in that position. This requirement is satisfied in the case of DP movement, in which a DP [THE ID-n], which has the interpretation of a variable, will occupy that position. This is not possible given Johnson's syntax for non-DP movement. Within its base-position, something must be available with a predicative interpretation. Since the AP itself is the only thing merged in this position, delayed Spellout will result in an uninterpretable structure. In short, semantic interpretation of moved material in the base-generated position is obligatory for non-DPs, rendering the delayed Spellout mechanism inapplicable to avoid island violations.

3.5.3 Scope reconstruction

It's been observed that islands block scope reconstruction (Longobardi 1987, Cinque 1990, Ruys 2011). Consider the sentence below:

(170) How many people do you think I should talk to?

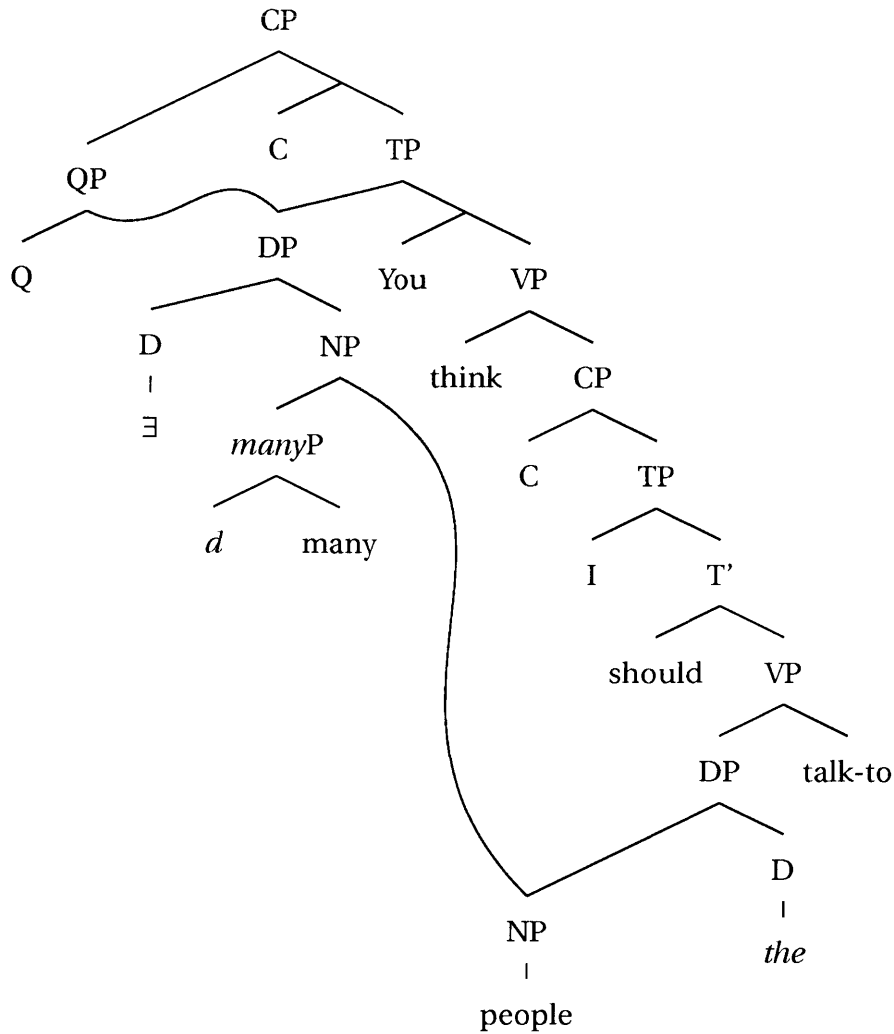
Example (170) is ambiguous. Under one reading, it means something like: *for what number n are there n people x such that you think I should talk to x ?* Under this reading, the *many* phrase takes scope over the modal *should*. To see what sort of LF generates this reading, we will assume that this sort of *wh*-phrase involves two sorts of quantification. First, a *wh*-operator which quantifies over degrees; second, existential quantification. Assume a semantics for the *many*-phrase similar to the analysis of Hackl (2000):

(171) $\llbracket \mathbf{many} \rrbracket = \lambda d. \lambda x. \text{measure}(x) = d$

After *many* combines with a degree d , it will denote a property of individuals which can combine with an NP via predicate modification. This NP will then combine with an existential quantifier.

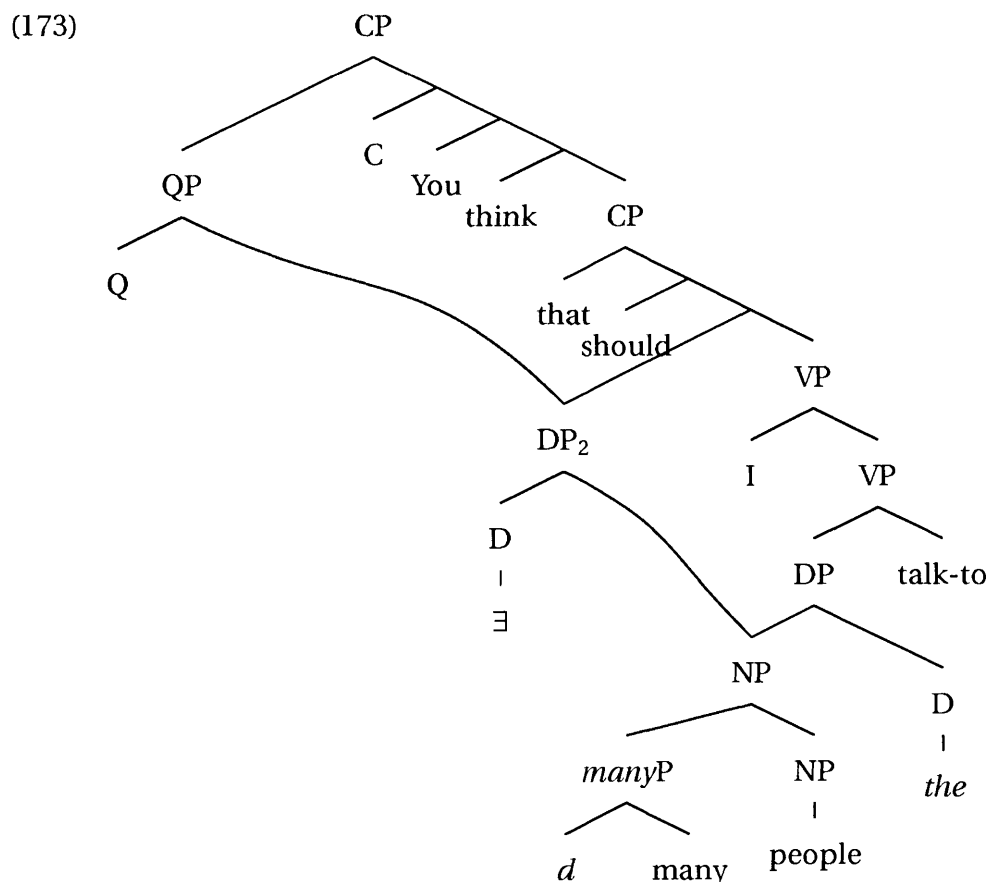
There are two positions in which the DP containing *many* can take scope. First, it may take scope above the modal:

(172)



The interpretation of this structure can be paraphrased as *what is the maximal degree n such that there are n-many people that you think I should talk to?*. For instance, suppose that you ask me for advice about who to talk to about a particular problem in phonology, and I response that you should speak with Bill and Tom. In this case, the answer to the question corresponding to the LF in (172) would be: two people (Bill and Tom).

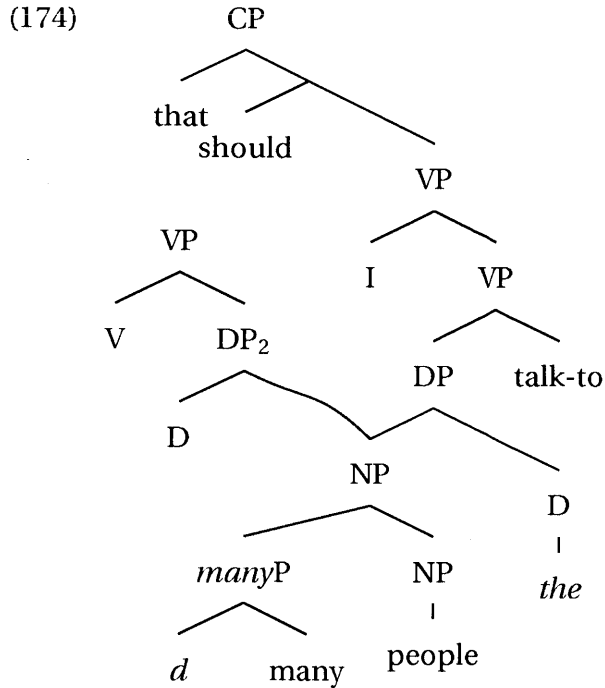
A second LF is available, in which the *many* phrase scopes below the modal:



The second LF corresponds to an interpretation with the following paraphrase: *what is the number n such that you think that there should be n -many people that I talk to?* This could correspond to a situation where I want you to talk to some number of people, but do not have any particular individual in mind.

Now let us consider these sorts of structures in island sentence. As discussed in **section 3.2.2**, the ambiguity disappears in these sentences. Instead, only a wide scope reading paraphrased is available.

This follows from our assumptions about how islandhood works. The point is that, even if the *many* phrase is merged in a position within the island, the mechanics of delayed Spellout will prevent its being interpreted inside of the island. At a relative clause island SOD, for instance, our syntax would posit a structure like (174):



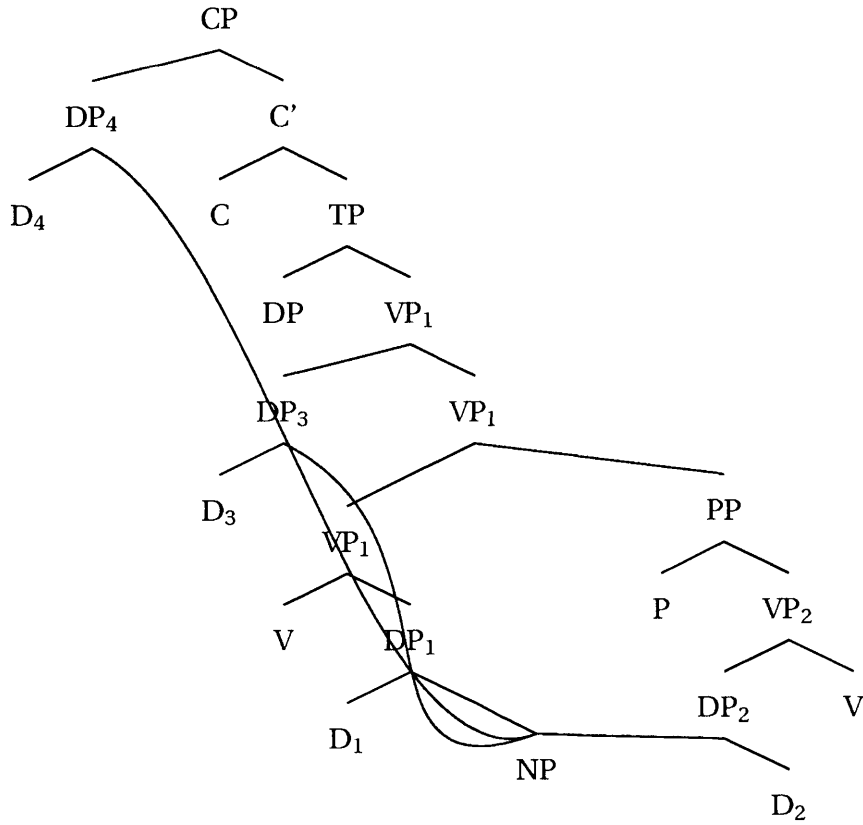
In this structure, the DP will be interpreted within the island as a bare D, and hence as a variable. Interpretation of the existential quantifier and *many*-phrase will be delayed until some point above the island, and hence above the modal. We correctly predict, then, that only the wide scope reading will be available in symbiotic and parasitic gap sentences. In single-island dependencies that cross island boundaries, the wide-scope interpretation — but not the narrow scope interpretation — should be marginally available.

3.5.4 Parasitic gaps

Next consider the derivation of a parasitic gap sentence like (175), with the structural description (176).

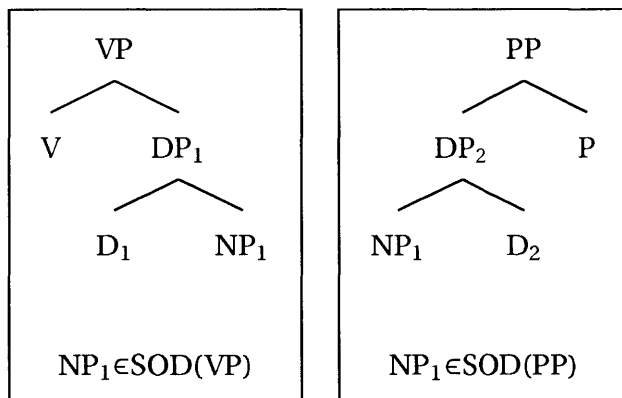
(175) Which coconut did Gilligan eat _ after picking _ ?

(176)

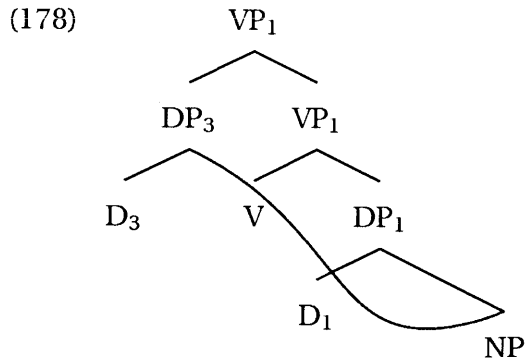


The derivation will begin with the VP and PP occupying distinct workspaces:

(177)

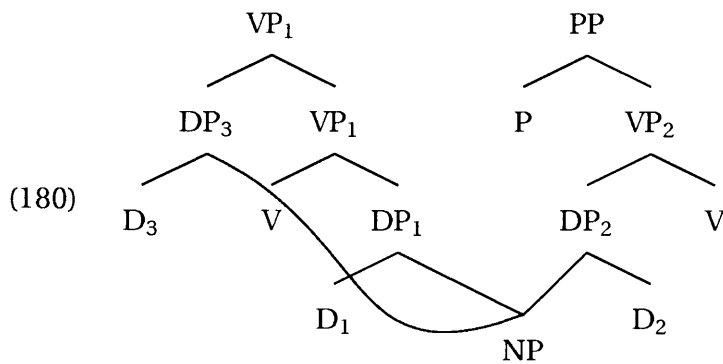


Within the workspace containing the main clause VP, a successive cyclic movement step may take place, and Spellout can apply:



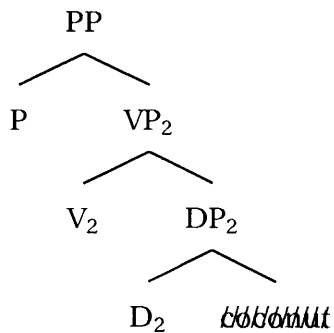
(179) $Lin = NP < V$

At this point, Collect Applies to the PP, adding it to the workspace containing the main clause VP:



$NP \notin SOD(PP)$

(181) **What Spellout sees:**



If Spellout applies at this point, the NP will fail to be interpreted within the adjunct, permitting further movement out of the island, but imposing selective island restrictions.

3.5.5 Delayed Spellout and single-gap sentences

Now consider extraction from an island structure, as in (182):

(182) ??Which skipper does Mary Ann know a professor who admires _?

In (182), the *wh*-phrase *which man* has been extracted from a relative clause, in violation of the complex NP constraint (Ross, 1967). According to the theory of islandhood adopted in this dissertation, the unacceptability is ultimately due to cyclic Spellout. Extraction out of the complex NP yields a linearization contradiction. However, the goal of this section will be to show that, in our Johnsonian framework, derivational possibilities exist which should allow delayed Spellout of the NP, and therefore avoid a linearization contradiction.

To see how this will work, let us focus on the crucial step in the derivation at which the island SOD is formed. We will assume that the SOD which determines islandhood in (182) is the relative clause itself.¹⁵ Following the conception of islandhood borrowed from Fox and Pesetsky, we will assume that linearization of the *wh*-phrase at the left edge of the relative clause is blocked.¹⁶

Each DP involved in the intermediate movement step initially occupies a separate workspace. Thus we might have a derivational stage like the following:

(183) $W_1 = [_{CP} \text{ who } [\text{ admires } [_{DP} \text{ the}_1 \text{ skipper }]]]$
 $W_2 = [_{DP} \text{ the}_2 \text{ skipper }]$
etc.

At this point, a number of possibilities are available. First, the island CP in W_1 could be Spelled out without first Collecting W_1 and W_2 . However, because the NP *skipper* would be completely dominated within W_1 , it would be linearized to the right of the

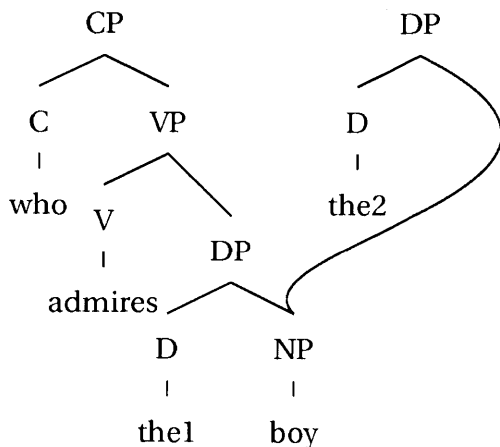
¹⁵The question of exactly what SOD is responsible for RC islands depends upon what is responsible for the RC's islandhood. Since RCs are arguably adjuncts, and since their creation involves RC-internal \bar{A} -movement, it seems plausible to assume that the RC itself constitutes either an adjunct or a subjacency island.

¹⁶This might be because movement of the *wh*-phrase is simply forbidden to the left edge of the island SOD, or because the *wh*-phrase is required to 'tuck in' at an inner specifier, CP position underneath the relativization operator. In a system like Fox and Pesetsky's, inner specifiers in languages like English are linearized in their base-generated position. See Fox & Pesetsky (2009).

V *admires*, which would ultimately lead to a linearization contradiction after further leftward movement.

However, as with the case of parasitic and symbiotic gaps, another possibility exists. Suppose that Collect applies between W_2 and W_1 , but that the second DP does not Merge with the island CP. At this point the W_1 would look like the following:

(184)



In this structure, the CP root of the relative clause completely dominates the determiner *the1*. The NP, which is also merged into a second DP, is not completely dominated by the CP. Now suppose that Spellout applies to the CP at this point. At PF Spellout, the NP will avoid linearization with respect to the other non-null terms in the SOD. Consequently, a linearization contradiction will not occur later on in the derivation. Again, though, this will lead to selective island effects.

3.5.6 Just-in-time Delivery

We are still left wondering why extraction of a DP from a pronoun-accepting context across an island boundary in a single-gap sentence should induce a degree of unacceptability which apparently is not found in ATB-extraction. That is, why is there a contrast between (185a) and (185b)?:

- (185) a. ??Who did Gilligan talk to someone who admires _?
 b. Who did Gilligan talk to someone who admires _ and Mary Ann go swimming with someone who can't stand _?

Or to put things more bluntly: why are there island effects in the first place?

At first glance, this might look like a lethal problem for the analysis. After all, Ross's (1967) discovery of islands stands as one of the great achievements of syntactic theory, and a theory which predicts islands to be non-existent should, all other things being equal, be dispreferred to one which derives them.

However, a theory which predicts extraction from islands to result in ungrammaticality is also inadequate, as it fails to derive the three-way contrast described in **section 3.2**. What the data shows is the following cline in acceptability:

(186) **Cline of acceptability across types of extraction:**

multi-gap non-selective island, non-island > single-gap non-selective island > single-gap multi-gap selective island

Any instance of extraction from selective island environments, whether or not it involves parasitic- or symbiotic-gap formation, results in unacceptability. The present theory predicts this, since selective islands are just those environments which require reconstruction into the gap site. All sentences involving extraction from a non-selective island environment are more acceptable than their selective island counterparts. The task now is to explain the contrast between single-gap island violating dependencies and multi-gap dependencies.

To explain this, we will have to look more closely at the derivational sequence which builds these structures. First, let us assume that there is a distinction between obligatory merge operations and optional operations, which by assumption will include adjunction and conjunction. In many cases, merger of a phrase in a particular location is necessary for a derivation to converge. For instance, arguments must be merged in a phrase where they will be assigned the appropriate thematic role. This also seems to hold of successive cyclic movement operations: McCloskey (2002), for instance, argues that *wh*-phrases are required to merge at each phase edge position. One could imagine implementing this idea with a distinction between feature-driven Merge and a non-feature-driven adjunction/conjunction operation, but the details of such a proposal will

be beyond the scope of this dissertation.¹⁷

I will assume a constraint which applies only to obligatory Merge operations, and which has the effect of forcing the operation to be immediately preceded by Collect of the relevant syntactic objects. I'll call this constraint **Just-in-time**:

(187) **Just-in-time**

An *obligatory* operation Merge(X,Y) must be immediately preceded by Collect(W_x, W_y), where $X \in W_x$ and $Y \in W_y$.

In general, this constraint has the effect of requiring Collect to apply immediately prior to Merge, and not earlier, and hence will minimize the amount of derivational ambiguity for any given phrase structure.

Returning to the model outlined in **chapter 1**, this constraint can be thought of as a restriction on possible workspaces. For instance, suppose that we have a well-formed workspace W_1

(188) $W_1 = \{A, \text{Merge}(B,C), D_1, \dots, D_n\}$

where Merge(A, Merge(B,C)) is well-defined and *obligatory*. Now consider a minimally different workspace, where Merge(B,C) is absent and where B and C are both elements:

(189) $W_2 = \{A, B, C, D_1, \dots, D_n\}$

The constraint **Just-in-time** doesn't allow W_2 to derive W_1 , even though Merge(B,C) is assumed to be well-defined. This is because in such a derivational sequence Merge(A,B) is not preceded by an operation that Collects A and B into a single workspace. Instead, the constraint is only satisfied via the following sequence of operations. First, we start off with two workspaces:

(190) $W_3 = \{A\}$

(191) $W_4 = \{B, C, D_1, \dots, D_n\}$

¹⁷There is an interesting literature in the minimalist grammar tradition (Stabler, 1997, a.o.) which attempts to formalize just this sort of distinction between normal Merge, which is taken to be feature-driven, and adjunction, which is not (Fowlie, 2015; Graf, 2014). This could form the basis for a more rigorous formalization of this idea.

This can be followed by Merge(B,C):

$$(192) \quad W_5 = \{\text{Merge}(B,C), D_1, \dots, D_n\}$$

This can of course be followed by Collect of W_3 and W_5 . The result is W_1 , but now **Just-in-time** is satisfied.

Let's now consider how **Just-in-time** operates in a sentence involving extraction from an island:

(193) What book did John leave after reading _?

(194) **LF for (193):**

$$[{}_{CP} [{}_{QP} Q [\text{THE book}_i]] [{}_{C'} \text{ did } [\text{John } [\text{leave }] [{}_{adjunct} \text{ after reading THE book}_i]]]]$$

The structural description of this sentence involves a single token of *book* merged within a QP, which is in turn merged in the specifier, CP position, as well as in a base position within the adjunct (internal structure of the adjunct is elided). Assume there are two SODs, CP and the adjunct itself.

Now consider a derivational stage where the QP occupies one workspace and the island SOD another:

$$(195) \quad \text{a. } W_1 = \{[{}_{QP} Q [\text{THE book}_i]]\}$$

$$\text{b. } W_2 = \{[{}_{adjunct} \text{ after reading THE book}_i]\}$$

Assume that the LI tokens *did*, *John*, and *leave* will be added by later iterations of Collect.¹⁸ The adjunct is of course an island, so if *book* is Spelled out inside of it, later Spellout at the higher position will result in an ordering contradiction. However, in order for delayed Spellout of *book* to apply, Collect(W_1 , W_2) must apply before Spellout of the adjunct, creating the following workspace:

$$(196) \quad W_3 = \{[{}_{QP} Q [\text{THE book}_i]], [{}_{adjunct} \text{ after reading THE book}_i]\}$$

¹⁸In fact, it's an entailment of **Just-in time** that every lexical item token which is the argument of an obligatory Merge operation must start off life in a workspace where it is the unique element. Hence, each LI token must start out in its own workspace.

(203) $W_4 = \{ \{ [_{VP} \text{ admire THE book}_i] , [_{VP} \text{ fell asleep } [_{adjunct} \text{ after reading THE book}_i]] \}$

This of course can be followed by Merge of the two VPs, and the other operations needed to complete the derivation. Crucially, because the VPs are merged in an, by assumption, optional coordination operation, **Just-in-time** is satisfied by this derivation. Adjunct parasitic gap structures will likewise not be constrained by **Just-in-time**, since adjunction is likewise an optional operation.¹⁹ In short, the contrast between single- and multi-gap dependencies comes down to whether **Just-in-time** is violated.

3.6 More issues

3.6.1 Some reconstruction puzzles

We have seen that this analysis predicts that semantic interpretation of the NP is banned in island contexts. Along with everything else, this predicts that reconstruction for the purposes of binding should also be banned. It turns out, however, that things are not as simple as one would hope.

One of the advantages of copy and remerge theory is that they make possible a syntactic theory of reconstruction into binding domains (Chomsky, 1995; Fox, 1999). This allows us to explain why condition C effects emerge in data like (204a) and why binding of a reflexive is possible in (204b):

- (204) a. *Which aspect of John_i does he_i despise _?
b. Which aspect of himself does John despise _?

Since our theory predicts that NP complements of moved DPs are not interpreted within islands, we should predict (i) a bleeding of condition C and (ii) the impossibility of reflexive binding into islands. This first prediction appears to be borne out:

¹⁹In order to account for subject parasitic gap constructions, we would have to assume that specifiers, or at least the class of specifiers that license parasitic gaps, fall into the same formal class as adjuncts. This might have something to do with the fact that subject parasitic gap licensing is notoriously restrictive in somewhat ill-understood ways. Indeed, Arad (2014) has argued that in languages like Hebrew they don't exist at all.

(205) ?That's the aspect of John_{i//j} that Mary was surprised because he_i despises _.

This sentence is no less acceptable on the reading in which *John* co-refers with the pronoun than one in which it does not.

Curiously, however, reconstruction of reflexives into islands does not appear any worse than extraction from an island is normally:²⁰

- (206) a. ?That's the aspect of himself that Mary was surprised because John despises _.
b. ?That's the aspect of Fred that Mary was surprised because John despises _.

While it remains unclear why this should be so, I should note that a similar puzzle arises for Stanton's analysis of anti-pronominality. Stanton's proposal is in line with the predictions of *wholesale late merger* that condition C is bled by A-movement. However, in certain cases Stanton's analysis appears to incorrectly predict bleeding of Condition A, for instance in A-movement:

(207) That aspect of himself seems to John _ to be unattractive.

If, following Takahashi & Hulse, we assume that late merger of the NP is possible in A-movement, then Stanton's theory should require the copy left behind in the position c-commanded by *John* to be a bare D, without the NP *aspect of himself*. This predicts reflexive binding to be impossible. Instead, Stanton suggests that the preference for late merger can be overridden by semantic requirements (such as a constraint requiring a bound variable or reflexive to be c-commanded by its antecedent at LF). This forces reconstruction where it is otherwise not permitted. The violability of Stanton's late merger constraint is perfectly compatible with the optimality theoretic logic of that system. However, it is unclear how a similar proposal could be made compatible with the architecture assumed in this dissertation.

²⁰David Pesetsky (p.c.) points out that a similar oddity arises with left dislocation structures in English:

- (1) a. ?That's the aspect of John_i that Mary was surprised because he_i despises it.
b. That's the aspect of himself that Mary was surprised because John despises it.

3.6.2 A syntactic question

There is at least one important difference between Johnson's (2012, 2014) original implementation of his multi-dominant syntax for movement and my own version of this theory. I assume an analysis of the DP borrowed from Fox (1999, 2002), in which the variable component is introduced by a *ID-n* function which adjoins to the NP:

(208) [_{DP} the [_{NP} *ID-n* N]]

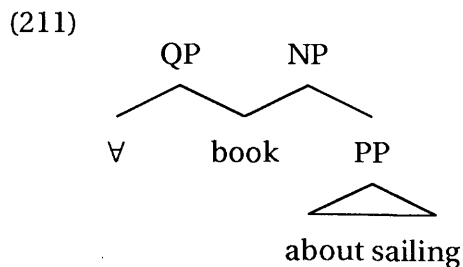
Johnson instead borrows Elbourne's (2005) analysis of the DP, where the variable component combines directly with the determiner:

(209) [_{DP} [the *n*] [_{NP} N]]

Kyle Johnson (p.c.) points out a potential problem with my implementation. A general issue with Johnson's syntax is that it seems to allow for a sort of anti-late merger effect. To see what this means, consider a possible derivation for a QR structure in of the sentence in (210):

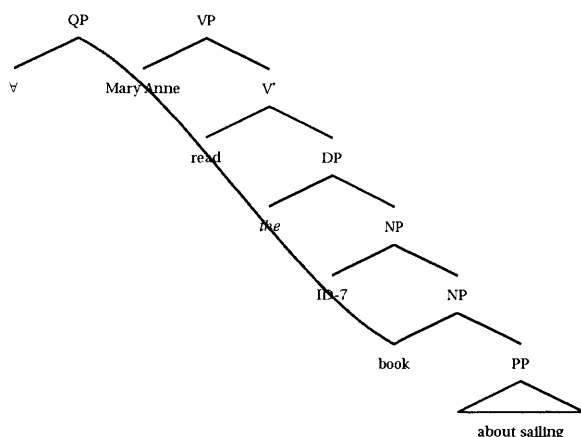
(210) Mary Anne read every book about sailing.

Suppose we first parallel Merge the object NP *book* with the PP *about sailing* and the universal quantifier \forall :

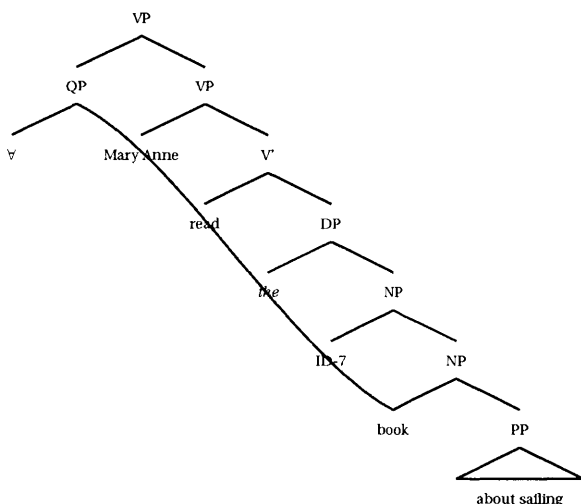


Next the the VP structure is built up (212), before merging the QP to the VP (213):

(212)



(213)



The point is that an insufficiently constrained version of Johnson's system makes the apparently pathological prediction that adjuncts can be merged in a lower copy of a moved DP but not a higher one. Johnson's (p.c.) concern was that whatever constraint rules out structures like (213) will also rule out structures necessary for my analysis in which ID-7 is not shared between every position the DP is merged in.

I believe that the appropriate generalization is that this sort of anti-late merger is banned for *semantically contentful* adjuncts. More explicitly, the problem is with the presupposition that this sort of adjunct will introduce. Notice that the DP in (213) introduces the presupposition that $g(7)$ is a book about sailing. If we assume *universal presupposition projection* (i.e. the universal quantifier presupposes that every member of the restrictor set satisfies the presuppositions of the scope) then the LF in (213) as a whole introduces the presupposition that *every contextually salient book is a book*

about sailing. But of course any context which satisfies this presupposition will be one in which all of the books are about sailing. So utterance of (210) is always contextually equivalent to its structurally simpler alternative *Mary Anne read every book*. If we assume that rational language users never use a more complicated sentence when a simpler one will do (Grice's maxim of manner) then sentence (210) with LF (213) is simply never utterable. Notice that the presupposition that winds up being introduced for a DP in which ID-7 is contained only in the lower DP is the tautology that *every contextually salient book is identical to itself*. Since this presupposition is satisfied in every context, the problem does not arise. While I believe that this represents a promising solution to this problem, determining whether it can be generalized to all cases of DP movement would take us into the complex and empirically murky question of which determiners introduce which sorts of presuppositions. Because this issue is complex and orthogonal to the main concerns of the theory, I refrain from developing these ideas further in this dissertation.

3.6.3 *wh*-in-situ

The analysis developed in this paper predicts a DP/non-DP asymmetry in extraction out of islands for languages with overt *wh*-movement, like English. What appears to be a similar contrast shows up in in-situ *wh*-constructions in languages like Chinese. Huang (1982) illustrates this for Chinese:

- (214) a. ni xiang-zhidao [shei mai-le sheme]?
 you wonder who buy-ASP what
 'What do you wonder who bought?'
 b. ni xiang-zhidao [shei weisheme mai-le shu]?
 you wonder who why buy-ASP book
 'Why do you wonder who beat Zhangsan?'
 c. ni xiang-zhidao [shei zeme mai-le shu]?
 you wonder who how buy-ASP book
 'How do you wonder who beat Zhangsan?'

Example (214a) shows that a *wh*-DP (*sheme*) is acceptable within a *wh*-island in Mandarin. However, in (214b) a non-DP *wh*-phrase *weisheme* is trapped inside of an island.

According to Huang, *wh* can only be interpreted within the embedded clause. A main clause interpretation similar to *why did you wonder who beat Zhangsan* _? is unavailable. Similarly, (214c) only has an interpretation where ‘who’ takes scope over ‘how’.

This appears to be a problem for the analysis presented in this paper. We have argued that a PF output constraint, by assumption Fox and Pesetsky’s ban on ordering contradictions, forbids movement out of islands from preceding as normal in cases of overt movement. To the extent that overt movement out of islands is possible at all, it is because of delayed Spellout of the *wh*-phrase. Among other things, this derives a DP/non-DP asymmetry for overt movement, since delayed Spellout is only possible for DPs. But in a language where *wh*-phrases remain in-situ, like Chinese, it is unclear why movement out of an island should be problematic at all, since no ordering contradiction should arise in the first place.

I think it is likely that restrictions like those in (214) are ultimately ‘weak island’ effects,²¹ like negative or factive islands. Abrusán (2014) has argued that ‘weak islands’ arise from interactions between the semantics of certain sorts of interrogatives and certain presuppositions. As an illustration, consider the islands that factive predicates form for manner adverbials:

(215) #How does Mary regret that John fixed the car?

Abrusán’s crucial assumption is that the domain of manners always contains contraries. That is, for any manner X, then some manner denoting the opposite of X also exists (e.g. wisely/unwisely, fast/slow, etc.). Weak island effects arise in the case of manner adverbials because of particular presuppositions. In the case of sentences like (215), the familiar factive presupposition is the culprit. A factive like *Gilligan regrets that he ate*

²¹Note that weak islands and selective islands are not the same thing. As Abrusán convincingly argues, weak islands are fundamentally semantic, and include such environments as negative islands and factive islands. These environments do not show standard SI effects, such as sensitivity to anti-pronominal contexts:

- (1) a. That’s the river that Mary didn’t have lunch on _
b. That’s the river that Mary knows that Bill had lunch on _.

Both of these sentences allow the idiomatic reading which is unavailable in other island contexts.

the coconut presupposes that Gilligan actually ate the coconut. Now take a factive in a *wh*-interrogative:

(216) What does Mary regret that John fixed?

Following a long tradition (Hamblin 1973, Karttunen 1977, etc.), Abrusán assumes that questions denote sets of propositions. In order for a question Q to be felicitous, the presupposition of every proposition in Q must be satisfied. This predicts that for a question like (216), the presupposition introduced will be:

(217) For every contextually salient non-human entity x , John fixed x

Returning to (215) this question will presuppose that, for every Manner in the domain, John fixed his car in that manner. Or more precisely, since Abrusán assumes manners to be properties of events:

(218) ‘for every manner X in D_M , the event of John’s fixing his car was in X ’

This is where the assumption about contrariety becomes crucial. Because whenever *fast* is in D_M *slow* is to, this presupposition will be satisfied only if the event of John’s fixing his car was simultaneously fast and slow. Since these two manners are incompatible, this presupposition will be unsatisfiable.

This theory makes a prediction which is relevant to our discussion of manner *wh*-adverbials in Mandarin. *How* questions are actually ambiguous. Consider:

(219) How did John climb Mt. Washington?

This sentence allows two kinds of answers. First, we might respond with a manner (well, fast, etc.). But one can also respond ‘He took the Tuckerman Ravine Trail.’ Following Tsai (1994), let’s call this the ‘means’ reading. Abrusán’s analysis should predict this response to be grammatical inside of a weak island, since the presupposition introduced by the ‘means’ reading aren’t predicted to be contradictory.

For Chinese, Tsai (1994) observes a syntactic correlate to the manner/means ambiguity:

- (220) a. Nimen zenmeyang da zhe chang lanqiu
 you how play this CL basketball
 ‘By what means do you play this basketball game?’
- b. Lanqiu, nimen da- de zenmeyang
 basketball you play DE how
 ‘In what manner do you play basketball?’

Tsai reports that while with (6a) speakers prefer the means answer, the post-verbal *zenmeyang* in (6b) renders a manner response obligatory.

Now consider the following pair, in which *zenmeyang* is embedded within a complex NP island:

- (221) a. Ni bijiao xihuan [[ta zenmeyang zhu] de cai]?
 you more like he how cook PNM dish
 ‘What is the means x s.t. you like better the dishes which he cooks by x ?’
- b. *Ni bijiao xihuan [[ta zhu- de zenmeyang] de cai]?
 you more like he cook DE how PNM dish
 ‘What is the manner x s.t. you like better the dishes which he cooks in x ?’

Post-verbal *zenmeyang* is unacceptable, while the preverbal equivalent is grammatical on the means reading. This contrast would make sense if we assume that the unacceptability of certain adverbials is an Abrusánian weak island effect and not fundamentally syntactic. While a full analysis of weak island effects in Mandarin is beyond the scope of this paper, Tsai’s observation suggests that a semantic analysis along the lines of Abrusán (2014) is compatible with Huang’s data.²²

At this point, one might wonder whether the DP/non-DP asymmetry discussed extensively in this paper is also simply a weak island effect. I think not, for the following reason. Chomsky (1982) points out that pied-piping of a PP out of an island is absolutely forbidden:

- (222) a. ?Which girl does John know a guy who talked to _.
- b. *To which girl does John know a guy who talked _.

²²Note, however, that Abrusán explicitly argues that *why* questions are different than manner questions, based on different behavior within semantic environments that ameliorate weak islands for manner adverbials. It would be interesting to see if the same differences in behavior carry over to the Mandarin data.

This is an instance of the familiar DP/non-DP asymmetry. It could be explained within our system, assuming that pied-piping is simply merger of Q with a constituent larger than the *wh*-DP (Cable, 2007). It is not explicable in semantic terms, however.

3.7 Conclusion

This proposal is a theory of how delayed Spellout of multi-dominant structures constrains both PF and LF interpretation. The central idea is that extraction from an island is possible so long as Spellout of the moved material is delayed within the island Spellout domain. Because Spellout simultaneously returns an LF and PF interpretation, however, this also has the effect of preventing semantic interpretation of moved material within the island.

Under most Merge based theories of movement, a proposal like this one would result in uninterpretability at LF. However, with Johnson's theory of DP movement, movement only targets the NP complement. As a consequence, only this NP fails to be interpreted in cases of movement out of islands. This, we have argued, allows a semantically interpretable LF to be formed in select cases, but imposes a variety of restrictions on what can move out of islands. These semantic restrictions perfectly correlate with the selective island effects discussed by Cinque and Postal. We have discovered that this analysis predicts such effects to appear in symbiotic and parasitic gap constructions, and can also be extended to single-gap dependencies.

Next, I will discuss three remaining issues. First, I will briefly compare this analysis with the resumptive pronoun analysis of selective islands. Second, I will reiterate the major conclusions of the chapter regarding symbiotic and parasitic gaps. Finally, I discuss how this proposal relates to the discussion of right node raising in the last chapter.

3.7.1 The resumptive pronoun analysis revisited

Both Cinque and Postal provided a *resumptive pronoun* analysis for the selective island data. In one version of this analysis, a derivation is available in which *wh*-phrases are base-generated in a position above an island, and anaphorically bind a potentially null

pronominal inside of the island. This is illustrated below for *which book does John know the man who wrote _?*, a sentence which violates the complex NP island constraint (Ross, 1967):

(223) *which book_i does John know ~~which book₁~~ [_{island} the man who wrote pro_i]*?

The island effect is ameliorated on this analysis because there has actually been no extraction out of the island. This derivation is assumed to be virtually identical to one with an overt resumptive pronoun, which also ameliorates island effects:

(224) *?Which book_i does John know a man who wrote it_i?*

Our own analysis should not be viewed as a competitor to this analysis so much as an attempt to derive explain it. Instead of assuming that null resumptive pronouns are independently inserted into the structure, we assume that they are an automatic consequence of delayed Spellout . In those circumstances where interpretation of the NP is blocked, viz. island environments, the DP occupying the base position should be treated by the interfaces as a phonetically null pronoun.

3.7.2 Parasitic gap formation and ATB-movement: A unified analysis?

A central conclusion of this chapter is that delayed Spellout is the mechanism responsible for both symbiotic and the more familiar parasitic gap construction. Thus, mine falls into the class of theories which attempt to reduce parasitic gap formation to ATB-movement (Williams, 1990; Nunes, 2001, a.o.). A classic argument against this approach (Postal, 1994) comes from the observation that parasitic gaps are, but not most ATB-gaps are not, subject to selective island effects. I demonstrated that an ATB-analysis of parasitic gaps can be saved from this objection, once the anti-reconstruction effects of delayed Spellout are taken into account. Selective island effects in parasitic gap sentences come about not because of some mechanism particular to parasitic gap formation, but because of the general mechanics of island amelioration itself.

This result is desirable both conceptually, since it simplifies the grammar, and empirically, since by retaining a common mechanism between parasitic gap and ATB-sentences

we are able to explain data otherwise troublesome for the alternative null operator analysis, such as parallel case matching effects in parasitic gap- and ATB-sentences sentences in languages like Russian²³ (Franks, 1993).

The problem of parasitic gaps is rather intricate, and the model outlined in this chapter cannot account for all of the phenomenon's peculiarities. I have, for instance, no account of the well known connectedness effects. As Kayne (1983) originally showed, parasitic gaps are banned when embedded inside of a subject island which is already contained within an island:

- (225) a. ✓ That's the book that we should destroy _ before someone steals a copy of _.
b. * That's the book that we should destroy _ before a copy of _ gets stolen by someone.

This sort of restriction was latter extended to adjunct islands by Longobardi (1985) and other syntactic islands by Chomsky (1986).

While I have no account for this restriction, it is worth noting that, following the pattern established by selective island effects, it can be extended to symbiotic gap and even single-gap sentences:

- (226) a. ✓ That's the book that John read a review about _ and Mary had a friend who stole a copy of _.
b. * That's the book that John read a review about _ and Mary had a friend who a copy of _ got stolen by.
- (227) a. ?? That's the book that Mary had a friend who stole a copy of _.
b. * That's the book that Mary had a friend who a copy of _ got stolen by.

It is a virtue of the present analysis that it ties all three construction types together. I hope that a fuller exploration of this model will allow us to solve this and other puzzles.

²³Franks (1993) demonstrates that, like its better known counterpart in ATB-movement, Russian parasitic gap formation is subject to case matching requirements. Cf. Oraviita & Taraldsen (1983), who provide evidence that Finnish parasitic gap constructions lack such a constraint. That is, Frank's data suggests the Russian is, but Oraviita & Taraldsen's suggests that Finnish is not, subject to a case-matching restriction on parasitic gap formation. The origin of such cross-linguistic variation remains a topic for further research.

3.7.3 Selective islands and RNR: A comparison with the results of the previous chapter

The leading idea of the previous chapter was that the right-edge restriction, and its obviation in right-node wrapping sentences, could be explained by a linearization algorithm that was sensitive to complete dominance in local domains. Within a right-node raising structure, material within each conjunct completely dominates the pivot with respect to the conjunct itself but not with respect to the entire coordinate structure. It follows that linearization of the pivot will apply internally to each conjunct, but not at the point where the coordinate structure itself is evaluated. As noted in the previous chapter, this linearization algorithm is actually agnostic with respect to which theory of cyclic Spellout one chooses, or even whether one assumes Spellout is cyclic at all. Unlike Bachrach & Katzir's system, it doesn't need to invoke a delayed Spellout mechanism to explain the right-node raising data. However, it is worth asking: To what extent is the algorithm compatible with the theory presented in this chapter? It turns out that the two ideas are in some tension with one another.

The major insight of the previous chapter was that, since the linearization algorithm *requires* the pivot to be linearized within each conjunct, the shared material will be linearized to the left of any material following it within a non-final conjunct, and so right-edge effects will emerge. The major insight of the present chapter is that, particularly in across-the-board and parasitic gap sentences, Spellout, and hence linearization, of the shared material is *optional*, permitting island obviation behavior. But if delayed Spellout can apply to material in ATB-movement, permitting otherwise contradictory linearizations from arising, why can it not likewise apply in right-node raising sentences, thereby allowing for obviation of the right edge restriction? In other words, if islands and the right edge effect both arise from the same constraint on linearization contradiction, and this constraint is obviated in coordinate structures for islands, why is it not obviated for the right edge effect?

It's interesting to compare the analysis of this section with that of Bachrach & Katzir. Recall that, for B&K, the right-edge restriction is ultimately a consequence of a lineariza-

tion algorithm which is distinct from Spellout, and which is insensitive to complete dominance. Hence their theory's predictions with respect to the right edge restriction are not as closely tied to the distribution of delayed Spellout environments as my analysis. Indeed, they predict a different distribution of delayed Spellout. Recall from the discussion in the previous chapter that within their system PF Spellout determines adjacency, rather than precedence, relations between terminal nodes. If one term is next to another in a Spellout domain, then it must be next to that term in all subsequent iterations of Spellout. This is encoded in the following constraint, repeated from the previous chapter:

(228) ADJACENCY PRESERVATION

If an adjacency relation is established within a Spellout domain K , then it is preserved in every Spellout domain which contains K

This difference with respect to my thesis plays a significant role in Bachrach & Katzir's analysis of right-node raising. Since adjacency relations are preserved, then without delayed Spellout of the pivot within all non-final conjuncts, right node raising structures are simply unlinearizable. If combined with the claim of this chapter that delayed Spellout entails anti-reconstruction effects, B&K's theory predicts that gap sites in all right-node raising structures should be subject to selective island effects. This appears not to be correct, as noted by Postal (1994):

(229) My friend speaks (in) $_$, but cannot read, **several Amazonian languages**.

This RNR sentence is well-formed, even when the gap site constitutes an anti-pronominal context. Therefore, the aspect of Bachrach & Katzir's theory which enables it to predict both delayed Spellout and a right-edge restriction makes it incompatible with the claim made in this chapter that selective islands correlate with delayed Spellout.

I have one final comment. It seems that selective island effects may actually appear within right-node raising sentences, in contexts that violates the right edge restriction. To see this, first recall the anti-pronominal cases of extraction from a PP discovered by Stanton:

- (230) a. *That's the river that John has a friend who had lunch on _.
b. *That's the holiday that John has a friend who eats turkey on _.

In both cases, to the extent that the sentence is acceptable at all, only a literal reading (in which John is somehow having lunch or eating turkey on a physical object) is available. Recall also that RNR bans preposition stranding. Hence the right-edge restriction of the last chapter predicts a contrast depending upon whether the gap site is in a rightmost position:

- (231) a. John last week had lunch on _ and this week went swimming in **the Charles river**.
b. *John had lunch on _ last week and this week went swimming in **the Charles river**.
(232) a. John last year ate turkey on _, and this year refused to celebrate **Thanksgiving**.
b. *John ate turkey on _ last year, and this year refused to celebrate **Thanksgiving**.

Interestingly, however, an effect reminiscent of the selective island phenomenon appears in the right-edge restriction violating sentences (231b) and (232b), at least according to my own judgments. That is, to the extent these sentences are acceptable at all, they only have a very odd literal reading and not the most natural and obvious reading. I do not offer a real analysis here, but to the extent that this observation is correct it is consistent with the pattern seen elsewhere in this chapter. Where the grammar can get around linearization contradictions at all, whether in extraction out of islands or the right-edge restriction, it must invoke delayed Spellout of the displaced phrase. This in turn predicts that selective island effects to arise.

Chapter 4

Islands at PF and LF

4.1 Introduction

The previous chapter incorporated an account of syntactic islandhood suggested by Fox & Pesetsky's (2005) theory of cyclic linearization. According to this proposal, the linear order of syntactic structures is computed cyclically, at particular Spellout domains. If a phrase is not ordered at the edge of a Spellout domain, then movement of that phrase out of it will result in an ordering contradiction, and therefore uninterpretability at PF. Islands are just those Spellout domains which prevent movement to an edge position.

This is a variant of the PF theory of islands (Merchant, 2001, a.o.). Abstracting away from the details of cyclic linearization, this sort of theory asserts that structures which involve extraction across an island boundary will lack a well-formed PF interpretation, and hence will be ungrammatical. This can be stated, very schematically, with a constraint like the following:

(233) **PF constraint on islands**

A phrase originating in an island may not be phonologically interpreted at any position outside of that island

In addition to the results of the previous chapter, a few other predictions follow from this proposal. First, it's reasonable to assume that when the phonology is not computed, as in cases of ellipsis, PF constraints will be irrelevant. Hence extraction out of an island

should be well-formed in ellipsis constructions. A second corollary is that sentences in which movement is invisible to PF, such as when a moved phrase is semantically interpreted outside of an island but pronounced inside of it, should satisfy the PF constraint in (233). In other words, covert movement, unlike its overt counterpart, ought to be immune from island effects.

As originally noticed by Ross (1969), sluicing, which is typically analyzed as an ellipsis construction, ameliorates island effects. This sort of *salvation by deletion* phenomenon is consistent with the predictions of the PF theory, as noted by Merchant (2001) and others. However, the second prediction does not seem to hold. Both quantifier raising and covert *wh*-movement can be shown to be island sensitive. In **section 4.2**, I review a both sets of data.

The result is an apparent contradiction. Accepting the description of salvation by deletion data that motivates the PF theory of islands results in a grammar that cannot predict island effects for covert movement. Developing a theory in which covert movement is subject to island effects, which would entail that the origins of island effects are located at LF, seems to rule out a PF based explanation of salvation by deletion. The goal of this chapter will be to resolve this apparent contradiction. The central claim of this chapter is that island effects stem from a filter which applies (in more or less the same way) to both LF and PF. That is, an LF constraint on islands (234) applies in parallel to the PF constraint (233):

(234) **LF constraint on islands**

A phrase originating in an island may not be semantically interpreted in a position outside of that island

For island violating movement structures, the PF constraint on linearization can be satisfied in two cases: covert movement and ellipsis constructions (in particular sluicing). However these mechanisms will only work if the constraint at LF is also satisfied. This means that island amelioration will only be possible as long as the moved phrase totally reconstructs to a position within the island boundary at LF. This predicts that semantic evidence should still show island sensitivity, even in environments that otherwise ame-

liorate island effects. A variety of rather intricate data testing this prediction is reviewed in **section 4.4**. Although the judgments are delicate, I believe I can tentatively claim that the results point in the right direction.

There are a number of ways to realize this idea. First, the version of delayed Spellout explored in Bachrach & Katzir's work can actually be extended to LF islands, in a way which predicts these data. I will discuss how below. I will also discuss another possibility: that the input to semantic interpretation requires, in addition to information about structure, reference to some ordering of nodes.¹ That is, the interfaces demand an ordering at LF as well as PF. Suppose also that the LF ordering is determined in the same way as the PF ordering, with the same conception of islands and the same notion of ORDER PRESERVATION. If it is further assumed that the two orderings are not necessarily isomorphic, that is that PF may linearize two terms one way and LF another, then this theory will be able to derive both the full array of data.

4.2 Islands at PF and LF

In this section, I review two arguments which seem to point in opposite directions. First, in **section 4.2.1**, I examine evidence from sluicing which supports a PF theory of islands. Second, in **section 4.2.2**, I examine evidence which shows that covert movement is island sensitive. As discussed in the introduction to this chapter, this is inconsistent with the PF theory of islands.

4.2.1 An argument for the PF theory of islands

Here I'll focus on two sorts of island constraints: relative clause (235a) and adjunct islands (235b):

(235) a. **Relative clause island**

*That's the kind of guy_i that Mary has a friend who likes _i.

b. **Adjunct island**

¹This sort of idea has been mooted in the literature on dynamic semantics to account for ordering effects on presupposition projection (Heim, 1990; Rothschild, 2011).

*Which guy did Mary go home after she talked to _?

As first noticed by Ross (1969), sluicing constructions seem to ameliorate this effect. This is illustrated below for both island types :²

(236) a. **Salvation by deletion: Relative clause islands**

Mary has a friend who likes a particular guy. I wonder which one Δ .

b. **Salvation by deletion: Adjunct islands**

Mary went home after she talked to some particular guy. I wonder which one Δ .

There are a number of ways for accounting for salvation by deletion. The first, which shall be adopted here, assumes that movement out of the island really happens, and that PF deletion has the effect of saving the sentence from ungrammaticality. Alternatively, one might assume that the true antecedent for the sluice is not a structure which requires an extraction out of an island, and therefore that the appearance of island obviation is illusory. For instance, one might assume that the true antecedent for the sluice is some sub-constituent of the island itself, which might be termed a ‘short source’. Additionally one might also argue, as Barros (2014, a.o.) does, that sluicing can be licensed by structures which do not in fact appear as explicit antecedents in the discourse. While a full discussion of the arguments for and against the PF deletion theory is beyond the scope of this chapter, I can at least defend the PF deletion theory against a short source alternative.

²There is an inconsistency between earlier chapters and the description of the set of islands that are subject to salvation by deletion as PF islands. As Bachrach & Katzir note, violations of the coordinate structure constraint are not subject to amelioration in what I am calling symbiotic gap contexts:

- (1) a. * That’s the guy_i that I talked to Mary and a friend of _.
b. * That’s the guy_i that Fred insulted _ and I talked to Mary and a friend of _.

This leads Bachrach & Katzir to conclude that the coordinate structure constraint is something other than a Spellout island. Within the terms of **chapter 3**, this means that it is not a PF island. However, as Ross (1969) and others observe, salvation by deletion *can* ameliorate the CSC:

- (2) ✓ Bill talked to Mary and a friend of some student. Now tell me which one.

Accepting the explanation of salvation by deletion in terms of a PF theory of islands obviously entails that the CSC is a PF island. This contradiction cannot, unfortunately, be resolved in this thesis.

Merchant (2001) defends a sort of hybrid account in which only some islands are saved by PF deletion. He argues that the adjunct and relative clause islands illustrated in (236) are best explained by a short source, rather than a PF deletion, account. Merchant's argument relies on the putative unacceptability of data like the following (Merchant 2001: pg. 211; Merchant's judgments):

- (237) a. * They hired no people who spoke a lot of languages. Guess how many!
b. * They didn't hire anyone who speaks a Balkan language. Guess which one!

However, I believe that the ungrammaticality of both can be explained independently. Sentence (237a) can be plausibly analyzed as an 'inner island' effect that can be explained by an anomalous semantic interpretation (Rullman, 1995; Beck & Rullman, 1999; Fox & Hackl, 2006). According to these theories, the question would be paraphrasable as 'what is the maximal number n such that they hired no people who speak n languages'. Since there is no upper-bound on the number of languages that they could have hired people who don't speak, the sentence lacks an exhaustive true answer, and hence is infelicitous.³

The second sentence (237b) is not ungrammatical at all, according to my judgments, on a reading where the indefinite correlate *a Balkan language* takes wide scope. As discussed in **section 4.5** below, this is the preferred interpretation for indefinite correlates anyway.

Returning to the island sentences in (236), one might note that a short source analysis is available for both. Since each island has an embedded clause, which can also act as an antecedent, a resolution of the ellipsis is available which does not violate the relevant island condition. An example of each potential antecedent is provided for (236a) below:

(238) **Relative clause islands: Short vs. long source**

- a. ✓ Δ = Mary has a friend who likes ~~which one~~
b. ✓ Δ = Mary's friend likes ~~which one~~

³According to David Pesetsky (p.c.), sentence (237a) does have a marginal reading available in which there is a specified set of languages and no one was found who speaks those particular languages. This is predicted by the analyses discussed here.

However, one can also construct antecedents in which the short source is an implausible candidate for antecedent. In these sentences (contra the judgements reported in Merchant) sluicing appears to be successful.

First, look at the relative clause sentences again. In (239), selection of the short source as the antecedent require the speaker to ask a question about the identity of some individual who is liked by Mary's friends. However, the most plausible reading is the one where the speaker is asking for the identity of the individual who none of Mary's friends like. This is only consistent with selection of the long source, which in turn requires extraction out of an island:⁴

(239) **Relative clause islands: Short source blocked**

- a. Mary has no friends who like some particular guy in our department.
Can you tell me which guy Δ ?
- b. ✓ Δ = Mary has no friends who like ~~which one~~.
- c. # Δ = Mary's friends like ~~which one~~.

A similar sort of example can be constructed for adjunct islands. In the sentence below, selection of the short source (240c) would directly contradict the assertion of antecedent sentence. This leaves the long source (240b), which involves extraction from an island, as the only possible antecedent:⁵

(240) **Adjunct islands: Short source blocked**

- a. ✓ Mary went home without talking to some particular guy.
Can you tell me which guy Δ ?
- b. ✓ Δ = Mary left without talking to ~~which guy~~.
- c. # Δ = Mary talked to ~~which guy~~.

⁴An alternative analysis, which cannot be ruled out, is that this environment licenses the subjunctive mood in the relative clause, making available a short source sluice which is paraphrasable as 'which guy do those friends not like'. I'm not sure how to rule this alternative out.

⁵An alternative analysis is available in which without is decomposed at LF in such a fashion that the negation operator appears within the adjunct. In such a case a short source would be available.

4.2.2 An argument for the LF theory of islands

A PF theory of islands seems to predict that, all other things being equal, covert movement should be insensitive to PF constraints that affect overt movement. If island effects are determined at PF, then it follows that covert movement should be island insensitive. Here, I review two arguments that covert phrasal movement, both of quantificational DPs and *wh*-phrases, is in fact sensitive to the same island constraints as overt movement.

Quantifier raising

The exact nature of the locality restrictions on quantifier raising is controversial. Conventional wisdom says that QR is strictly clause-bounded, resembling in this respect scrambling processes in languages like German and Dutch (Johnson & Tomioka, 1997, a.o.).⁶

However, at least a few authors (Rodman, 1976; Farkas, 1981; Reinhart, 1997) have argued that QR is sensitive to the same island constraints as ‘ordinary’ \bar{A} -extraction.

For instance, Reinhart argues that the clause boundedness of QR can be circumvented by a sufficiently supportive context. Below is an example adapted from Reinhart, with the context provided by me:

(241) *Context:* Every day, five new patients enter the hospital, and each is assigned to a different doctor. That doctor takes responsibility for giving that patient a tranquilizer.

a. A doctor usually makes sure that we give every new patient a tranquilizer⁷.

$\forall > \exists$

b. A doctor usually examines the possibility that we give every new patient a tranquilizer.

$\# \forall > \exists$

⁶Cechetto (2004) offers an analysis where long-distance QR is permitted only in contexts, like antecedent contained deletion, that require it to ensure semantic interpretability. This fundamental prediction, that inverse scope out of embedded non-island clauses is banned in non-ACD environments, contradicts the judgments reported below.

⁷Examples from Reinhart (1997: pg. 336)

In the first example (241a), the universal quantifier is embedded within a non-island clause. Given the context, only the inverse scope reading, in which the universal quantifier scopes over the existential quantifier in the matrix clause, should be felicitous. In (241b), the context again rules out the surface scope reading. However, in this sentence, the universal is embedded within an complex NP island. The observation is that an appropriate reading can be constructed for the first sentence, but not for the second, indicating that the island constraint is active for QR.

Some more examples are provided below⁸

(242) a. An Amtrak official typically announces that every train from New York is stopping at the station.

($\forall > \exists$)

b. An Amtrak official typically announces if every train from New York is stopping at the station.

(* $\forall > \exists$)

(243) a. At every party convention, a politician announces that every state delegation supports the nominee.

($\forall > \exists$)

b. At every party convention, a politician reads an announcement that every state delegation supports the nominee.

(* $\forall > \exists$)

The first pair of examples (242) involves an *if*-clause adjunct. The sentence (242a), in which the universal quantifier is embedded within a non-island clause, permits a reading where the Amtrak official making the announcement can vary by the train. In the second sentence, this inverse scope reading is blocked. The second example again involves a complex NP island. As with the previous examples, the inverse scope reading is

⁸Fox & Sauerland (1995) observe that wide scope is more likely to be judged acceptable in generic sentences, as all of the examples here are. They develop an analysis where this apparent wide scope reading for the universal is in fact an illusion fostered by the presence of a generic operator in the matrix clause, with the universal quantifier itself remaining in situ. While I have no account for the genericity contrast, it's worth noting that this sort of analysis doesn't explain the island sensitivity witnessed in (241)-(243).

available only in the non-island sentence (243a).

All three pairs of examples reveal a contrast between quantifiers embedded in tensed non-island clauses and island clauses. The existence of this contrast suggests, in line with the analysis of Reinhart (1997), that long-distance QR is at least possible for many English speakers, and that this long-distance QR process is sensitive to syntactic island constraints.

Covert *wh*-movement

The next argument hinges crucially on the assumption, originating with Pesetsky (1987, 2000), that an in-situ *wh*-phrase in a language like English has two options for attaining the appropriate scope. The first option involves covert phrasal \bar{A} -movement. The second involves what Pesetsky (2000) terms feature movement, in which the *wh*-phrase remains in-situ (not undergoing phrasal movement) and attains scope via a mechanism distinct from \bar{A} -movement.

Importantly, Pesetsky argues that only this second operation is sensitive to ‘Beck’ intervention effects (Beck, 1996). If an intervener, like negation or *only*, appears between the in-situ *wh*-phrase and its scope site, and if the first option of covert phrasal movement is blocked, then the sentence is anomalous. More specifically, it is claimed that the *pair-list reading* of the multiple question is blocked. Pesetsky uses this theory to derive the difference between intervention sensitivity in superiority obeying and superiority violating multiple questions in English.

Kotek (2014) argues that the first operation, covert phrasal movement, is sensitive to islands, and the second operation, feature movement, is not. Her evidence for this claim uses Beck intervention effects as a diagnostic for covert phrasal movement. If a given intervener renders a sentence anomalous on its pair list reading, then phrasal movement has been blocked from applying. If covert phrasal movement is island sensitive, then an intervener which appears above the island should render the pair-list reading unavailable:

(244) **Kotek’s claim:**

Beck Intervention effects can diagnose island sensitivity of covert movement.

- a. An intervener can block the pair-list reading (though not a single pair reading)

First, consider a case where the in-situ *wh*-phrase is clause bounded, and an intervener appears in the matrix clause. This is schematically illustrated in (245a). An example is provided in (245b). The pair-list answer is judged to be acceptable.

(245) **Intervener over non-island clause boundary**

- a. *wh1* ...INT.. [*clause* ... *wh2*]
- b. Q: Which linguist doesn't it that we invited which philosopher?
A: Chomsky doesn't like it that we invited Nozick; Norvin that we invited Derrida; Donca that we invited Aristotle.

However, when the in-situ *wh*-phrase appears inside of an island, as in (246), then an intervener above the island should block the pair-list reading. This is judged to be true, with a pair-list reading reported as anomalous.

(246) **intervener over island boundary**

- a. *wh1* ... INT ... [*island* ... *wh2*]
- b. Q: Which linguist won't come if we invite which philosopher?
A: #Chomsky won't come if we invite Nozick; Norvin if we invite Derrida; Donca if we invite Aristotle.

Finally, when an intervener above the island clause is absent, phrasal movement above the island is unnecessary. This means that a pair-list reading should be available when there is no intervener present (247) or when it is contained within the island (248):

(247) **Non-island clause boundary over intervener**

- a. [*wh1* ... [*island* ... *wh2*]]
- b. Q: Which linguist will come if we invite which philosopher?
A: Chomsky will come if we invite Russell; Norvin if we invite Lewis; Donca if we invite Plato.

(248) **Island clause boundary over intervener**

a. [*wh*1 ... [*island* ... INT ... *wh*2]]

b. Q: Which linguist will come if we don't invite which philosopher?

A: Chomsky will come if we don't invite Nozick; Norvin if we don't invite Derida; Donca if we don't invite Aristotle.

In these cases, both sentences are judged to have a pair list reading, in line with the predictions of Kotek's analysis.

4.2.3 Summary

This section has reviewed evidence both for and against a PF theory of islands. The evidence for comes from salvation by deletion phenomena in sluicing: Island constraints which otherwise render an extraction ungrammatical seem to be lifted in cases, like sluicing, where the island environment is elided. I demonstrated this with data from sentences that contain both relative clause and adjunct islands. Salvation by deletion has been shown to be active in a variety of other island types as well (Ross 1969, Merchant, 1999, a.o.).

Second, I reviewed two arguments that covert movement is sensitive to islands. The first came from quantifier raising, where complex NP and adjunct island environments were shown to block quantifier raising, even when the only true reading of the sentence (given the context) required it. This limitation was shown not to apply to quantifiers embedded within non-island tensed clauses, suggesting that island constraints are indeed active for QR. The other argument came from covert *wh*-movement, where I reviewed Kotek's (2014) arguments that Beck intervention effects can be used to diagnose the availability of covert phrasal movement, and that this sort of movement is blocked when an in-situ *wh*-phrase is stuck inside of an island. Taken together, these two arguments suggest that covert movement is generally island sensitive in much the same way as its overt counterpart.

In the next section, I attempt to reconcile these two arguments by proposing that movement is subject to parallel LF and PF island constraints, both of which must be

satisfied if a sentence is to be well formed.

4.3 Analysis

4.3.1 PF and LF Constraints

The claim of this chapter is that a pair of constraints applies, to LF and PF respectively, that cause derivations which involve movement out of an island to crash at the interfaces. These constraints are repeated below:

(249) Constraints on extraction from islands

a. PF

A phrase originating in an island may not be phonologically interpreted at any position outside of that island (modulo salvation by deletion)

b. LF

A phrase originating in an island may not be semantically interpreted in a position outside of that island

In words, each constraint prevents a phrase which is merged inside of an island environment from receiving an LF or PF interpretation at some position outside of that island. Below, I state what the predictions of the theory will be.

4.3.2 *Wh*-in-situ and multiple questions

In order to allow for the correct interpretation of movement structures, I will assume a *PF neglect rule*, which states that, when a phrase occupies multiple c-commanding positions within a given syntactic structure, one of those positions may be ignored at PF Spellout. When the lower position is ignored, the result is overt movement, while neglect of the higher position triggers covert movement. In addition, I will assume an *LF neglect rule* which like wise allows for a phrase to be ignored in a given position. When this rule applies to the higher position in a movement structure, the result is LF reconstruction.

As discussed in **section 4.2.2**, I will assume, with Pesetsky (2000) and Kotek (2014), that there are in principle two ways of interpreting *wh*-in-situ (250a):

(250) a. ***wh*-in-situ**

Which boy read *which* book?

b. **Covert phrasal movement**

$[_{CP} \text{Which boy } [\lambda x \text{ [which book } [\lambda y \text{ [C } [_{TP} x \text{ read } y]]]]]]]$

c. **Feature movement**

$[_{CP} \text{ which boy } [\lambda x \text{ [} C_y \text{ [_{TP} x \text{ read which}_y \text{ book }]]]]$

First, covert phrasal movement may apply to the in-situ *wh*-phrase. As with quantifier raising, the phrase is then interpreted in its derived position, binding a variable in the lower position (250b). Second, what Pesetsky (2000) calls feature movement may apply. I will represent this with an index that appears both on the *wh*-phrase (*y* in (250c)) and on the complementizer head where the *wh*-phrase takes scope. In subsequent work, Beck (2006) and Kotek (2014) argue that *wh*-phrases introduce alternatives (Hamblin, 1973; Rooth, 1992). Pesetsky's feature movement is reinterpreted as the effect this set of alternatives has on the focus value of the entire question constituent. Crucially, no matter how the idea is implemented, the *wh*-phrase itself remains in-situ at LF in this sort of process. The important thing is that the relationship between the phrase and its scope site is decided via a mechanism other than phrasal movement.

As discussed in **section 4.2.2**, covert phrasal movement can freely apply over a Beck intervener, whereas feature movement cannot. For instance, the LF in (251b) will be well formed, that in (251c), which violates this constraint, will be ill-formed:

(251) a. Which boy didn't read which book?

b. **Phrasal movement: no intervention effect**

$[_{CP} \text{Which boy } [\lambda x \text{ [which book } [\lambda y \text{ [C } [_{NEG} [_{TP} x \text{ read } y]]]]]]]]]$

c. **Feature movement: Intervention effect**

$[_{CP} \text{ which boy } [\lambda x \text{ [} C_y \text{ [_{NEG} [_{TP} x \text{ read which}_y \text{ book }]]]]]]$

Finally, I will assume that a pair-list LF is distinguished from a single-list LF by the presence of multiple interrogative complementizers (Fox, 2012; Kotek, 2014). For a sentence

Now consider the case of sluicing. A key assumption (adopted from Merchant) is that ellipsis obviates the PF constraint. However, the LF constraint should remain active. Since LF reconstruction is possible, this will require the following sort of configuration, with both PF and LF being well-formed:

(253) a. John left without taking some book. I wonder which one Δ .

b. **Antecedent LF:**

[C [John left [without taking some book]]]

c. **Sluice PF:**

which one ~~John left without taking~~ ~~which one~~.

d. **Sluice LF:**

[C_x [John left [without taking which_x one]]]

In short, the PF island constraint is irrelevant (due to salvation by deletion), but the LF constraint remains in force. The *wh*-phrase may be pronounced in a position outside of the island. It may not be semantically interpreted there.

At this point, the reader might notice that the hypothesized LFs in (253) seem to contradict a widely reported fact about sluicing: That the correlate to the *wh*-phrase remnant must typically take the widest possible scope. In particular, indefinite correlates will tend to take wide scope with respect to other quantifiers. This follows from a common assumption, not made in this dissertation, that *wh*-phrases must generally take scope at a CP position, and that the correlate must show parallel scope. However, in the LFs above, neither the *wh*-phrase nor its correlate, the indefinite in the antecedent sentence, must take scope outside of the island position. In fact, such an LF would violate the island constraint which holds at LF. But how then do we explain the widely observed scope facts? The credibility of my analysis of sluicing depends on resolving this puzzle. My solution will hinge on another well known fact about questions in general: They typically presuppose a unique true exhaustive answer. I will show that this presupposition forces the contexts to be such that wide scope for the correlate will typically be the only plausible one. This is discussed more in **section 4.5**.

But first, I'd like to show how the system discussed so far makes some additional

predictions about sluicing.

4.4 Predictions II: When deletion provides no salvation

As we have seen, the pair of generalizations in (249) predicts that ellipsis should ameliorate island effects, but with certain caveats. In sluicing constructions, although overt movement at PF is permitted, semantic interpretation at the derived position is not permitted. That is, the *wh*-phrase must reconstruct at LF, and any constraints which apply at LF should remain in effect. This is stated below:

(254) Prediction

When PF constraints are lifted (as in sluicing), LF constraints should remain active

An entailment of (254) is that sluicing should be acceptable only in environments where the site at which the *wh*-phrase is interpreted is not outscoped by a Beck intervener. The first goal of this section is devoted to testing this prediction, using examples modeled on Kotek's data discussed above. The second is to point out an additional prediction made by the system for binding theory judgments.

4.4.1 Intervention effects and sluicing

Parallelism condition

It's typical to assume that the deleted constituent in sluicing must correspond in some fashion to a constituent in the antecedent sentence. It's also been commonly observed that focused constituents in the ellipsis domain are not relevant to this calculation. Here, I adopt a kind of LF parallelism condition, which states that the sluiced constituent must be LF-identical to some antecedent, *modulo focus*. For concreteness (and compatibility with the thesis that the two LFs must be equivalent), I assume a *structural* theory of focus alternatives (Fox & Katzir, 2011). According to this theory, focus alternatives consist of sets of LFs (rather than semantic values). A set of alternatives is constructed by replacing

each focused constituent of an LF with alternative constituents that are of equal or lesser complexity. The definition is provided below:

(255) **LF parallelism condition**⁹

A parallelism domain PD satisfies the parallelism condition if it is LF-identical to an antecedent constituent AC, modulo focus marked constituents

PD is LF-identical to AC modulo focus marked constituents, if there is a focus alternative to PD, PD_{alt} s.t. PD_{alt} is LF-equivalent to AC

(256) An LF ϕ is a focus-alternative of an LF ψ iff ϕ can be derived from ψ via successive replacement of focused constituents of ψ by constituents that are of equal or lesser structural complexity¹⁰

Some background

The data below will crucially rely on multiple sluicing, since the intervention effect diagnostics to be used are most clearly visible when testing pair-list readings. The desired sort of LF can be schematically constructed as follows:

(257) a. **Antecedent LF**

$[_{CP} C [\dots (INT) \dots [_{island} \dots QP_1 \dots QP_2]]]$

b. **Sluicing LF**

$[_{CP} C_7 C_8 [\dots (INT) \dots [_{island} \dots [_{which_7} NP] \dots [_{which_8} NP]]]]$

⁹This definition is adapted from Takahashi & Fox (2005), whose definition is in turn inspired by (Rooth, 1992; Romero, 1998; Heim, 1997).

¹⁰A more explicit definition is provided below (Fox & Katzir 2011: pg. 97):

- (1) S' is at most as complex as S , given some context C , iff S' can be derived from S by successive replacement of sub-constituents of S with elements of the substitution source for S in C
- (2) The substitution source for S in C is the union of the following sets:
 - a. The lexicon
 - b. The sub-constituents of S
 - c. The set of salient constituents in C
- (3) $F(S,C)$, the set of focus alternatives of S given $C = \{S': S' \text{ is derived from } S \text{ by replacing focus constituents } x_1, \dots, x_n \text{ with } y_1, \dots, y_n, \text{ where } y_1 \text{ is at most as complex as } x_1 \text{ given } C, \text{ etc.}\}$

If the prediction stated in (254) is correct, then a context containing the appropriate antecedent LF will license the sluicing LF, provided that no intervener separates the *wh*-phrases from the interrogative complementizers marking the scope of the question.

First, let's take a look at an example that has all the relevant attributes in place, except for the presence of an island: ¹¹

(258) I think that John recommended every book_{*i*} to a student who later cited it_{*i*}. Now tell me exactly which book to which student.

This sentence has a very particular set of properties needed for multiple sluicing in English to be well formed. First, multiple sluicing is subject to a *clausemate* condition, meaning that both *wh*-phrases must emanate from the same clause. Second, English multiple sluicing is often reported to be degraded when both *wh*-phrases are DPs. This sentence obviously satisfies both of these constraints. Additionally, satisfaction of the parallelism condition demands that *every book* in the antecedent clause must be a focus alternative of *which book* in the elided constituent. The relevant LFs are provided below:

(259) a. **Antecedent constituent**

[every book [λx [[a student who cited x] [λy [John recommended x to y]]]]]

b. **Elided constituent**

[[which book]_F [λx [[which student]_F [λy [John recommended x to y]]]]]

Interveners and islands

The stage is now set for the crucial examples. In the first one, the multiple sluice construction has both *wh*-phrases emerging from a complex NP island:

(260) **Context:** There are three boys and three girls , with three rumors floating around that each one of the boys has talked to exactly one of the girls (who happens to be the girl that that boy has a secret crush on). Mary tells you about each rumor...
 A: Mary repeated the rumor that every boy talked to the girl who he likes...
 ...Tell me exactly which boy to which girl.

¹¹See Romero (1998) for more discussion of non-definite correlates to sluicing sentences.

A few things are of note here. First, the context forces a pair-list reading, since it's clear that multiple boys have rumors floating around that they have talked to a girl. Second, the constraints on sluicing require that the two *wh*-phrases and their antecedent DPs are both base-generated inside of the island. Thus the LF of the antecedent must look like (261a) and that of the sluicing sentence (261b):

(261) a. **Antecedent LF:**

[Mary repeated [the rumor that [every boy λx [x talked to the girl that x likes]]]]

b. **Parallelism LF:**

[C_x [C_y [Mary repeated [the rumor that [which_x boy λx [x talked to [which_y girl]]]]]]]]

At this point, the reader might observe that the antecedent LF seems to entail something weaker than what is presupposed by the question. The antecedent merely entails that Mary repeated some rumor about every boy talking to a girl. The question, however, requires that for every boy there is a different rumor. It seems as though, contrary to the prediction of the theory, the universal quantifier should have to scope outside of the complex NP island in order to provide a suitable antecedent. However, it should be noted that the context already supports the stronger presupposition demanded by the question, since it is asserted that a rumor exists for each boy and that Mary has told the speaker about each rumor. In other words, although the logical entailment of the antecedent is too weak to satisfy the presupposition of the question, the overall contextual entailment is not. Notice also that a short source interpretation for the sluice is unavailable, since that reading would entail that each boy has in fact talked to the girl that he likes, not merely that there are rumors about this doing so.¹²

¹²This example also reveals an ontological fact about rumors: That they are cumulative. Hence, if I heard a rumor about John and a separate rumor about Mary, then it can truthfully be said that I heard a rumor about John and Mary. This cumulative property of rumors allows for the antecedent sentence in this example to be true even though the context only makes salient a number of separate rumors. This fact also plays a role in the following sluicing sentence, which involves only a single *wh*-phrase. Imagine that a number of rumors are floating around about who got jobs. In particular, there are rumors about each of the current fifth years getting a job. John has heard a majority of these rumors. Now consider the following discourse:

In any case, the sluicing construction seems to be well formed. Now consider a sentence identical to the first, but with an intervener (in this case the negative indefinite *no one*) in the matrix clause. The context has been minimally adjusted so that the antecedent makes sense:

- (262) a. **Context:** There are three boys and three girls, with various rumors floating around that each one of the boys has talked to one or another of the girls. Various people have repeated rumors about each boy, but for each boy, no one has repeated the rumor that he talked to the particular girl that he has a crush on.
- A: No one repeated the rumor that every boy talked to the girl who he likes...
 ...Tell me exactly which boy to which girl.

Although the sluicing sentence should have a well-formed PF, to satisfy the LF constraint on islands both *wh*-phrases must reconstruct and be interpreted via Pesetsky's feature movement:

- (263) a. **Antecedent LF:**
- [No one repeated [the rumor that [every boy λx [x talked to the girl that x likes]]]]
- b. **Sluicing LF:**
- [C_x [C_y [No one repeated [the rumor that [which_x boy λx [x talked to [which_y girl]]]]]]]]

The presence of the intervener should render this sentence ill-formed at LF, and the sentence should be judged as anomalous. The prediction seems to point in the right direction: The sentence is much harder to judge as true than its counterpart in (260).

(1) John heard a rumor that most fifth years got jobs. Can you tell me exactly which ones?

In the antecedent of this sentence, the correlate *most fifth years* is embedded inside of an island, hence must take narrow scope. However, the fact that John has heard rumors about the majority of fifth years entails, by the cumulative nature of rumors, that he has heard a rumor about most of them. This allows the first sentence to be true. As with the example in the main text, the presupposition on the sluiced question that the question has a unique exhaustive true answer is satisfied by the context, even though the antecedent itself would be too weak.

For another example, imagine that rumors are flying around the internet about compromising documents linking various members of the Trump family to different high ranking officials in the Russian government. Jane, an investigative reporter, is convinced of the accuracy of most of these reports. However, for each member of the Trump family, she doubts the claim that that particular Trump is linked to a particular Russian official. Now I say to you:

- (264) a. Jane is skeptical that each member of the Trump family is linked to a particular Russian official. Now tell me which Trump to which Russian.
 b. Jane isn't convinced that each member of the Trump family is linked to a particular Russian official. Now tell me which Trump to which Russian.

The sentences differ in that, in (264b), an intervener c-commands the correlates to the *wh*-phrases in the sluice. However, in both sentences the correlates occupy positions embedded inside of non-island clauses. Hence covert movement to positions above the negation operator is possible, in principle, allowing for well-formed LFs:

- (265) a. **Antecedent LF:**
 $[[\forall \text{ Trump}] \lambda_x [\text{a RO}] \lambda_y C [(\text{NEG}) [\text{Jane is skeptical} [\text{that } x \text{ is linked to } y]]]]$
 b. **Sluicing LF:**
 $[[wh \text{ Trump}] \lambda_x [wh \text{ RO}] \lambda_y C_Q [(\text{NEG}) [\text{Jane is skeptical} [\text{that } x \text{ is linked to } y]]]]$

Now consider, given the same scenario, sentences involving islands:

- (266) a. Jane will be proven wrong if each member of the Trump family is linked to a particular Russian official. Now tell me which Trump to which Russian.
 b. Jane won't be proven correct if each member of the Trump family is linked to a particular Russian official. Now tell me which Trump to which Russian.

Now we predict a contrast between (266a), which lacks a Beck intervener and so should be well-formed, and (266b), in which the presence of a negation operator is predicted to render the sentence anomalous.

A similar set of examples can be constructed with CNPC islands:

- (267) a. Jane (didn't) claim that each member of the Trump family is linked to a particular Russian official. Now tell me which Trump to which Russian.
- b. Jane (didn't) repeat the claim that each member of the Trump family is linked to a particular Russian official. Now tell me which Trump to which Russian.

Again the prediction is that sentence (267a), whether or not a negation operator is present in the main clause, should be acceptable on the pair-list interpretation. However, in (267b), where the *wh*-phrases and their correlates emanate from a complex NP island, should show a contrast. The sentence should only be acceptable when the negation operator is absent.

4.4.2 Extending the prediction: Binding theory

The prediction that the LF island constraint remains active when its PF counterpart is lifted should extend to other empirical domains. Here, I will focus on binding theory.

Wh-movement is known to feed reflexive binding:

- (268) Bill_i knows which picture of himself_i Mary Likes _.

In this sentence, a reflexive is bound by the matrix subject. Since the reflexive is inside of a *wh*-phrase, reconstruction of that phrase to its base position would result in a violation of condition A of the binding theory.

Since Condition A presumably applies at LF, this effect should be blocked in any environment which requires LF reconstruction. The LF constraint therefore should entail that sluicing environments that involve extraction out of an island will also prevent binding of a reflexive in the derived position.

This is tested in the examples below:

- (269) Mary_i knows we'll hire an applicant who praises some aspect of her_i in the interview. But she'd be shocked if she_i knew...
- a. <>which aspect of herself_i Δ.
- b. <> which aspect of her_i Δ.

- (270) Mary_i knows the winning applicant must praise some aspect of her_i in the interview. But she'd be shocked if she_i knew...
- a. <> which aspect of herself_i Δ.
 - b. <>which aspect of her_i Δ.

The antecedent sentence contains a CNPC island. Parallelism requires the *wh*-phrase to originate within that island. If LF reconstruction is mandatory, then binding of a reflexive at the derived position, as in (269a), should be impossible, whereas binding of a non-reflexive pronoun should be permitted (269b). By contrast, both the reflexive and non-reflexive pronoun should be well-formed in the example (270), where the indefinite correlate does not emanate from an island.

Below is another pair of examples, identical to the one above, except that it now involves an adjunct island, rather than a CNPC island:

- (271) John_i knows that that I can make fun of some aspect of him_i. But he_i'd be shocked to know...
- a. <> which aspect of himself_i Δ.
 - b. <>which aspect of her_i Δ.

John_i knows that I'll come to the party if I can make fun of some aspect of him_i.
But he_i'd be shocked to know...

- (272) a. <> which aspect of himself_i Δ.
b. <>which aspect of her_i Δ.

Unfortunately, these sentences are quite hard to judge, so all I can do here is point out what the predictions are.

4.5 Sluicing and scope parallelism

The story presented above relies on the assumption that scope reconstruction of the *wh*-remnant in sluicing is not only possible, but in some cases necessary. This assumption sits uneasily alongside a variety of observations that sluicing forces wide scope for

its correlate in the antecedent. For instance, Barker (2013) observes that contexts which render the wide scope reading for an indefinite implausible also forbid that indefinite from acting as the correlate for a sluicing sentence:

- (273) a. # John photographed a woman or the Empire state building yesterday. But I don't know who.
- b. # Every teacher called more than two students yesterday. But I don't know who.
- (*cf. John called more than two students yesterday...*)

In example (273a), the most plausible reading of the antecedent is one where the indefinite takes narrow scope with respect to disjunction. That is, it is either the case that John photographed some arbitrary woman or that he photographed the Empire State Building. This contrasts with a reading where there is some particular woman such that I either photographed her or the Empire State Building yesterday. Barker observes that the preference for the narrow scope of the indefinite inhibits it from acting as a correlate to the *wh*-phrase in the continuation sentence.

Similarly, it's been observed that modified numerals in direct object position resist taking wide scope with respect to universal subjects (Szabolcsi, 2011). Barker observes that such a modified numeral cannot act as the correlate to a *wh*-phrase remnant in a sluicing continuation (273b).

These facts do not follow from the LF parallelism condition stated above, at least when the possibility of LF reconstruction for the *wh*-phrase is factored in. The LFs for (273a) provided below should meet the parallelism condition stated above:

- (274) a. **Antecedent LF (273a)**
- [_{CP} C [John photographed [∃ woman]₇ or [the Empire State Building]]]]
- b. **Sluicing LF (273a)**
- [_{CP} C₇ [John photographed [[who₇]_F or [the Empire State Building]]]]]

I think that an alternative story for the facts observed by Barker comes, however, from the presuppositions introduced by questions. A question typically introduces a presup-

position that it has a unique true exhaustive answer (Dayal, 1996). To see what this predicts for the examples at hand, consider the denotation for the LF in (274b):

$$\begin{aligned}
 (275) \quad \llbracket (274b) \rrbracket &= \\
 &\{\lambda w. \text{John photographed } x \text{ in } w \vee \text{John photographed the EST in } w \mid x \text{ is a woman}\} \\
 &= \\
 &\{\lambda w. \text{John photographed Ishani in } w \vee \text{John photographed the EST in } w, \lambda w. \\
 &\text{John photographed Juliet in } w \vee \text{John photographed the EST in } w, \lambda w. \text{John pho-} \\
 &\text{tographed Lilla in } w \vee \text{John photographed the EST in } w, \dots\}
 \end{aligned}$$

The relevant presupposition that only one of the possible answers denoted by the question is true.¹³ Hence there must be some particular woman who John photographed (if he didn't photograph the Empire State Building).

However, according to Barker, the most natural reading of the antecedent is one where it takes narrow scope with respect to negation. This encourages an inference that the speaker has no particular woman in mind, and in fact doesn't care who the woman is. This is inconsistent with the requirements imposed by the question, leading to the question's oddness. A similar story will suffice to explain Barker's other example (273b).

4.5.1 More on indefinites and islands

An analysis which enables the apparent scope of the sluiced *wh*-phrase to mismatch its actual scope at LF might also give us a handle on some facts about indefinites and islands. The acceptability of sluicing correlates with a preference for a wide scope indefinite reading in the antecedent clause. Some examples from Reinhart (1997: pg. 340) are provided below:

- (276) a. If *a certain linguist* shows up, we're supposed to be particularly polite. But do you remember who?
- b. Max will believe everything that *someone* will tell him. And you can easily guess who.

¹³More exactly, the presupposition is that there is one answer which is true and entails all of the other true answers. But in this case no answer entails another. So what's presupposed is that only one of the answers can be true

From one point of view this corroborates the standard view that if the *wh*-phrase must take wide scope, then so must its correlate. However, the question arises of why indefinites may take wide scope with respect to an island in the first place. This is, of course, a well researched question (Fodor & Sag, 1982; Reinhart, 1997; Schwarzschild, 2002). The standard assumption only makes sense if the indefinite manages to take genuinely wide scope at some position outside of the island. This is the story told by Reinhart, who argues that indefinites and *wh*-phrase share a mechanism for scope taking (the binding of choice function variables) not available to other quantifiers, which allows them take scope at positions unavailable to other quantifiers.

However, according to other theories, such wide scope is only apparent. For instance, Schwarzschild (2002) assumes that indefinites are ordinary existential quantifiers which behave scope island constraints. The phenomenon of apparently wide scope is an illusion fostered by the fact that indefinites, but not other quantifiers, can be contextually delimited so that only a single element falls into their restriction. If I utter, *While walking down Comm Ave today, I saw a turkey*, the turkey I'm referring to might be the only contextually salient turkey around. This is referred to as a *singleton indefinite*. Schwarzschild observes that this results in scope neutralization with respect to other quantifiers:

(277) Every motorist on Commonwealth Avenue saw a turkey this morning.

If there's only one salient turkey, then this sentence will have the same truth conditions regardless of which scope the indefinite takes with respect to the universal quantifier. Schwarzschild goes on to show that this effect of scope neutralization suffices to explain the apparently wide scope effects associated with indefinites in island contexts, along with a variety of other peculiar facts associated with the phenomenon.

When it comes to sluicing, if the indefinite correlate actually takes scope within the island, as Schwarzschild's theory predicts, then so must its correlate *wh*-phrase within the sluicing sentence. Since the apparent wide scope of the indefinite is an illusion, the apparent scope of the *wh*-phrase must be as well. This is possible once we incorporate the insights of Pesetsky, Beck, and Kotek into our analysis. As we discussed earlier, an LF

is available in which the apparent scope of the *wh*-phrase in some sense is illusory: The *wh*-phrase is interpreted in its base-position, and only establishes a relationship with the interrogative C head via feature movement:

(278) a. **Antecedent clause**

[_{CP} [_{C'} C [if [[∃ linguist]shows up] [We should be polite]]]]

b. **Parallelism domain**

[_{CP} [_{C'} C₇ [if [[*which*_{F7} linguist]shows up] [We should be polite]]]]

These two LFs are equivalent, provided we assume that (i) feature movement is not relevant to the calculation of parallelism, and (ii) the existential is a focus alternative of the *wh*-phrase. Therefore, accepting Schwarzschild's analysis entails accepting something like the arguments being advanced here, and vice versa.

4.6 Two theories of LF islandhood

4.6.1 Bachrach & Katzir revisited, again

As Bachrach & Katzir (2007) shows, their system can actually be developed to make interesting predictions about islandhood at LF. While their specific proposal is complex in ways largely irrelevant to this thesis, it boils down to the following constraint:

(279) **Constraint on Spellout of variables**

If a phrase denotes a free variable at some Spellout domain K, it must denote a free variable for all SODs containing K

This predicts that a phrase semantically interpreted as a free variable within a particular Spellout domain cannot be reinterpreted as a bound variable within a higher SOD. In the case of movement structures, this means that movement must be successive cyclic for purposes of LF, as well as PF, interpretation.

To understand precisely how this works, it's important to understand that within Bachrach & Katzir's system, the distribution of delayed Spellout configurations differs

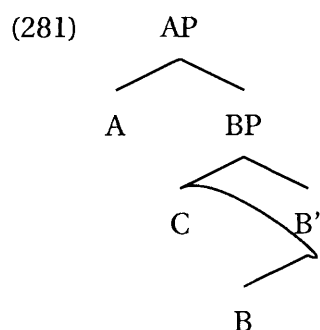
from the proposal made in this dissertation. This is because their definition of complete dominance differs from my own. Their own definition is provided below:

(280) **Complete dominance à la Bachrach & Katzir**

X completely dominates Y in C iff

- a. X is the only mother of Y in C, or
- b. Y has a mother in C, and X completely dominates every mother of Y in C.

To see how this definition works (and how it differs from my own system) consider a simple internal Merge configuration:



Here A is merged in two positions, once as the daughter of B', and again as the daughter of BP. Using the definition of complete dominance employed elsewhere in this dissertation, BP would completely dominate C, because BP reflexively dominates both mothers of C. However, according to Bachrach & Katzir's definition, BP *fails* to dominate CP: Because C has two mothers, (280b) applies; Since BP is the mother of B', it completely dominates it. However, BP is not the mother of itself, so it does not completely dominate itself. Because BP does not completely dominate one of the mothers of C (i.e. BP itself), it does not completely dominate C. Therefore, $C \notin \text{CDD}(\text{BPAP})$. Now consider AP. Because it completely dominates BP (it is the only mother of BP), it completely dominates every mother of C, meaning that it completely dominates C.

If we suppose that Spellout domains are co-extensive with complete dominance domains, then this proposal derives the escape hatch effect of movement. Suppose that BP is a Spellout domain. This entails that the adjacency relations and semantic interpretation of BP must be fixed in all higher Spellout domains. Since C is not completely

dominated by BP, it will avoid Spellout. Therefore, even if C denotes a variable at either position it is merged in, the constraint (279) will not apply, enabling the variable to be bound at a higher Spellout domain. Among other things, this permits movement. However, if AP, rather than BP, is a Spellout domain, then C will be Spelled out. This means that if it denotes a free variable within the SOD it must denote a free variable within all SODs that contain AP. This blocks an LFs that involve movement of C to higher positions and interpretation of the lower positions as bound variables.

This theory derives the constraint on LFs stipulated at the outset of this chapter. If a phrase is merged inside of an island (i.e. an SOD that prevents movement to its edge) then binding of this phrase at the position inside of the island by an occurrence of the phrase at a position outside of the island is blocked. This blocks LFs from being constructed that involve movement across an island boundary. Notice, though, that this crucially relies on Bachrach & Katzir's specific definition of complete dominance, which allows for delayed Spellout at edge positions. This assumption is incompatible with the proposal made in **chapter 3**, since the crucial result of that chapter is that delayed Spellout environments are selective island environments. If we were to combine the two proposals, the result would be the absurd prediction that all successive cyclic movement structures are selective islands. Therefore, in the next subsection, I propose an alternative account that is compatible with the results of **chapter 3**.

4.6.2 Cyclic linearization at LF

Another way to implement the LF constraint in terms parallel to its PF counterpart is to assume that LF, as well as PF, imposes an ordering on syntactic structures. I call this the LF ordering hypothesis:

(282) **LF Ordering Hypothesis**

LF Spellout, like its PF counterpart, imposes an ordering on phrases

- a. LF and PF orderings asymmetric and subject to ORDER PRESERVATION
- b. Covert phrasal movement receives the same LF ordering as overt movement

Since the LF ordering of phrasal movement is always ‘overt’, it will follow that all movement structures are subject to island effects at LF.

An internal merge structure can be interpreted in two ways at LF. First, the moved phrase can be interpreted in its derived position. For this sort of structure to be semantically interpretable, the lower position will have to be interpreted as a variable, via an operation like trace conversion (Fox, 1999). In this case, the LF ordering algorithm will have to ignore the lower position or treat it as a distinct phrase from the higher position, in order to avoid an ordering contradiction. Conversely, the higher position can be ignored, via something like an LF neglect rule. Application of this sort of rule results in LF reconstruction.

As we saw above, ordering restrictions at PF force overt movement to be successive cyclic. The same thing applies for covert phrasal movement. Interpreting a phrase in its derived position also forces movement to be successive cyclic. To see why, consider the following sentence, supposing that inverse scope is possible:

(283) A different doctor said that every patient recovered.

For this structure to be well-formed given the LF ordering hypothesis, an LF like the following will be necessary:

(284) LF:

[_{CP1} *every patient* [a doctor said [_{CP2} *every patient* [that [*every patient* [recovered]]]]]]]

In this structure, the universal quantifier *every patient* first raises to the specifier position of CP2, and then to a position c-commanding the existential in the matrix clause, allowing an inverse scope reading. When CP2 is Spelled out, PF will order this phrase in its base position, while LF will order at the left edge position:

(285) a. Spellout_{PF}(CP2) = that < every < patient < recovered

b. Spellout_{LF}(CP2) = \forall < that < recovered

Spellout of CP1 will again order the universal quantifier in its base position at PF and at a position to the left of the indefinite at LF. The result is a consistent ordering at both interfaces and an inverse scope reading at LF:

- (286) a. $\text{Spellout}_{PF}(\text{CP1}) = \text{a doctor} < \text{said} < \text{that} < \text{every} < \text{patient} < \text{recovered}$
 b. $\text{Spellout}_{LF}(\text{CP1}) = \forall < \exists < \text{said} < \text{that} < \text{recovered}$

An alternative derivation, in which the quantifier fails to covert movement of the universal quantifier fails to apply successive cyclically, would result in an LF ordering contradiction. \forall would be ordered after *that* in CP1, but before it in CP2. In other words, in order to interpret a phrase in some position outside of the Spellout domain where it originated, that phrase must undergo successive cyclic movement at LF.

This allows us to define a notion of LF island, comparable to the notion of PF island stated above:

(287) **Definition of a LF island**

An LF island for movement of α is a Spellout domain that does not allow α to be interpreted at its left edge

4.7 Conclusion

This chapter has generalized the PF theory of islands to both interfaces, arguing that parallel constraints apply to each interface, forbidding a phrase originating from an island from being interpreted at a position outside of the island. I showed how this proposal derives an apparently inconsistent set of data: salvation by deletion phenomena, which seems to point towards a theory of islands as arising from constraints at PF; and island effects in covert movement, which seems to point towards a theory of islands as arising from constraints at LF. I then showed that this theory makes other interesting predictions involving multiple sluicing constructions, and that the data seems to point in the right direction. Finally, I discussed an implementation of this proposal in terms of cyclic linearization, arguing that both LF and PF structures must receive an asymmetric linearization.

Bibliography

- Abney, S. (1987). The English Noun Phrase in its Sentential Aspect. PhD thesis, MIT.
- Abrusán, M. (2014). Weak Island Semantics. Oxford University Press.
- Arad, T. (2014). The nature of resumptive pronouns: Evidence from parasitic gaps. PhD thesis, Tel Aviv University.
- Bachrach, A. & Katzir, R. (2007). Spelling out qr. In Proceedings of Sinn und Bedeutung.
- Bachrach, A. & Katzir, R. (2009). Right-node raising and delayed spellout. In Grohmann, K., editor, InterPhases: Phase-Theoretic Investigations of Linguistic Interfaces. Oxford University Press.
- Bachrach, A. & Katzir, R. (2017). Linearizing structures. Syntax, 20(1):1–40. To appear in *Syntax*.
- Barker, C. (2013). Scopability and sluicing. Linguistics and Philosophy, 36:187–223.
- Barros, M. (2014). Sluicing and identity in ellipsis. PhD thesis, Rutgers University.
- Beck, S. (1996). Quantified structures as barriers for If movement. Natural Language Semantics, 4:1–56.
- Beck, S. (2006). Intervention effects follow from focus interpretation. Natural Language Semantics, 14:1–56.
- Beck, S. & Rullman, H. (1999). A flexible approach to exhaustivity in questions. Natural Language Semantics, 7(3):249–297.
- Booij, G. (1985). Coordination reduction in complex words: A case for prosodic phonology. In van der Hulst, H. & Smith, N., editors, Advances in nonlinear phonology, pages 143–160. Foris, Dordrecht.
- Brooks, G. (1990). Friends in low places. No Fences. Capitol Nashville.
- Cable, S. (2007). The grammar of Q: Q-particles and the nature of wh-fronting, as revealed by wh-questions in Tlingit. PhD thesis, MIT.
- Cechetto, C. (2004). Explaining the locality conditions of qr: Consequences for the theory of phases. Natural Language Semantics, 12(4):345–397.

- Chaves, R. (2007). Coordinate structures: Constraint-based syntax-semantics processing. PhD thesis, University of Lisbon.
- Chomsky, N. (1965). Aspects of the Theory of Syntax. MIT Press.
- Chomsky, N. (1977). On *wh*-movement. In Culicover, P., Wasow, T., & Akmajian, A., editors, Formal Syntax. Academic Press, New York.
- Chomsky, N. (1982). Some concepts and consequences of the theory of Government and Binding. MIT Press.
- Chomsky, N. (1986). Barriers. MIT Press.
- Chomsky, N. (1995). The minimalist program. MIT press.
- Chomsky, N. (2000). Derivation by phase. In Kenstowicz, M., editor, Ken Hale: A life in language. MIT Press.
- Cinque, G. (1990). Types of A'-dependencies. MIT Press, Cambridge, MA.
- Citko, B. (2005). On the nature of merge: external merge, internal merge, and parallel merge. Linguistic Inquiry, 36(4):475–496.
- Curry, H. (1961). Some logical aspects of grammatical structure. In Structure of Language and Its Mathematical Aspects, Proceedings of the 12th Symposium in Applied Mathematics. American Mathematical Society.
- Dayal, V. (1996). Locality in WH Quantification. Kluwer.
- Engdahl, E. (1983). Parasitic gaps. Linguistics and Philosophy, 6:5–34.
- Engdahl, E. (1986). The Syntax and Semantics of Questions with Special Reference to Swedish. Spring.
- Farkas, D. (1981). Quantifier scope and syntactic islands. In Proceedings of CLS 17. Chicago.
- Fodor, J. D. & Sag, I. (1982). Referential and quantificational indefinites. Linguistics and Philosophy, 5:355–398.
- Fowlie, M. (2015). Minimalist Grammars with Adjunction. PhD thesis, UCLA.
- Fox, D. (1999). Reconstruction, binding theory, and the interpretation of chains. Linguistic Inquiry, 30(2):157–196.
- Fox, D. (2003). On logical form. In Hendrick, R., editor, Minimalist Syntax. Blackwell.
- Fox, D. (2012). Multiple *wh*-questions: uniqueness, pair-list, and second order questions. Course notes, MIT.

- Fox, D. & Hackl, M. (2006). The universal density of measurement. Linguistics and Philosophy, 29:537–586.
- Fox, D. & Katzir, R. (2011). On the characterization of alternatives. Natural Language Semantics, 19:87–107.
- Fox, D. & Pesetsky, D. (2005). Cyclic linearization of syntactic structure. Theoretical Linguistics, 21(1-2):1–46.
- Fox, D. & Pesetsky, D. (2007). Cyclic linearization of shared material. Talk presented at Bogaziçi University, April 6 2007.
- Fox, D. & Pesetsky, D. (2009). Rightward movement, covert movement, and cyclic linearization. Lecture at Ben Gurion University.
- Fox, D. & Sauerland, U. (1995). Scope illusions. In Proceedings of NELS 26.
- Franks, S. (1993). On parallelism in across-the-board dependencies. Linguistic Inquiry, 24(3):509–529.
- Gazdar, G., Klein, E., Pullum, G., & Sag, I. (1985). Generalized Phrase Structure Grammar. Blackwell.
- Graf, T. (2014). Models of adjunction in minimalist grammars. In Formal Grammar 2014.
- Hagstrom, P. (1998). Decomposing questions. PhD thesis, MIT, Cambridge, MA.
- Hall, B. (1965). Aplikativnaja poroždajuščaja model' i isčislenie transformacij v ruskom jazyke by S. K. Šaumjan, P. A. Soboleva. Language, 40:397–410.
- Hamblin, C. L. (1973). Questions in montague english. Foundations of Language, 10(1):41–53.
- Heim, I. (1990). Presupposition projection. In van der Sandt, R., editor, Reader for Nijmegen workshop on presupposition, lexical meaning, and discourse processes. University of Nijmegen.
- Heim, I. (1997). Predicates or formulas? evidence from ellipsis. In Proceedings of SALT VII.
- Heim, I. (2013). Functional readings of questions. MIT.
- Heim, I. & Kratzer, A. (1998). Semantics In Generative Grammar. Blackwell, Malden, MA.
- Huang, C.-J. (1984). On the distribution and reference of empty pronouns. Linguistic Inquiry, 15:531–574.
- Huang, C.-T. J. (1982). Move wh in a language without wh-movement. The Linguistic Review, 1:369–416.

- Hudson, R. (1976). Conjunction reduction, gapping, and right node raising. Language, 52:535–562.
- Johnson, K. (2012). Towards deriving differences in how *wh*-movement and QR are pronounced. Lingua, 122(6):529–553.
- Johnson, K. (2014). Multidominance and movement. lectures presented at the Hebrew University of Jerusalem, December 2014.
- Johnson, K. & Tomioka, S. (1997). Lowering and mid-size clauses. In Katz, G., sook Kim, S., & Haike, W., editors, Proceedings of the 1997 Tübingen Workshop on Reconstruction, pages 185–206.
- Kayne, R. (1983). Connectedness. Linguistic Inquiry, 14(2):223–249.
- Kayne, R. (1994). The antisymmetry of syntax. MIT Press, Cambridge, MA.
- Ko, H. (2005). Syntactic edges and linearization. PhD thesis, MIT.
- Kotek, H. (2014). Composing Questions. PhD thesis, MIT, Cambridge, MA.
- Larson, B. & Parker, D. (2013). Across-the-board movement is actually asymmetrical. Presented at Annual LSA Meeting.
- Longobardi, G. (1985). Connectedness, scope, and c-command. Linguistic Inquiry, 16(2):163–192.
- Longobardi, G. (1987). Extraction from NP and the proper notion of head government. In Giorgi, A. & Longobardi, G., editors, The syntax of noun phrases, pages 57–112. Cambridge University Press.
- McCawley, J. (1982). Parentheticals and discontinuous syntactic structure. Linguistic Inquiry, 13:91–106.
- McCloskey, J. (2002). Resumption, successive cyclicity, and the locality of operations. In Epstein, S. D. & Seely, T. D., editors, Derivation and Explanation in the Minimalist Program. Blackwell.
- Merchant, J. (2001). The syntax of silence: Sluicing, islands, and the theory of ellipsis. Oxford University Press, Oxford.
- Munn, A. (1992). A null operator analysis of ATB gaps. The Linguistic Review, 9:1–26.
- Munn, A. (1993). Topics in the Syntax and Semantics of Coordinate Structures. PhD thesis, University of Maryland, College Park, MD.
- Nunes, J. (2001). Sideward movement. Linguistic Inquiry, 32(2):303–344.
- Nunes, J. (2004). Linearization of chains and sideward movement. MIT Press, Cambridge, MA.

- Oraviita, I. & Taraldsen, T. (1983). Case-marking and parasitic gaps in Finnish. University of Tromsø.
- Pesetsky, D. (1982). Paths and categories. PhD thesis, MIT.
- Pesetsky, D. (1987). *wh*-in-situ: movement and unselective binding. In The representation of (in)definiteness. MIT Press.
- Pesetsky, D. (2000). Phrasal movement and its kin. MIT Press, Cambridge, MA.
- Peters, P. & Ritchie, R. Phrase linking grammars. Manuscript.
- Phillips, C. (1996). Order and Structure. PhD thesis, MIT, Cambridge, MA.
- Phillips, C. (2006). The real-time status of island phenomena. Language, 82(4):795–823.
- Postal, P. (1966). On so-called ‘pronouns’ in English. Monograph series in language in linguistics, 19:177–206.
- Postal, P. (1993). Parasitic gaps and the across-the-board phenomenon. Linguistic Inquiry, 24(4):735–754.
- Postal, P. (1994). Parasitic and pseudoparasitic gaps. Linguistic Inquiry, 25(1):63–117.
- Postal, P. (1998). Three Investigations of Extraction. MIT Press, Cambridge, MA.
- Reinhart, T. (1997). Quantifier scope: How labor is divided between *qr* and choice functions. Linguistics and Philosophy, 20:335–397.
- Rodman, R. (1976). Scope phenomena, “movement transformations”, and relative clauses. In Partee, B. H., editor, Montague Grammar, pages 165–176. Academic Press, New York.
- Romero, M. (1998). Focus and reconstruction effects in *wh*-phrases. PhD thesis, University of Massachusetts, Amherst.
- Rooth, M. (1992). A theory of focus interpretation. Natural Language Semantics, 1:75–116.
- Ross, J. R. (1967). Constraints on variables in syntax. PhD thesis, MIT, Cambridge, MA.
- Ross, J. R. (1969). Guess who? In Binnick, R., Davidson, A., Green, G., & Morgan, J., editors, Proceedings of CLS, volume 5, pages 252–286.
- Rothschild, D. (2011). Explaining presupposition projection with dynamic semantics. Semantics & Pragmatics, 4(3):1–43.
- Rullman, H. (1995). Maximality in the semantics of *wh*-constructions. PhD thesis, University of Massachusetts, Amherst.

- Ruys, E. (2011). Semantic reconstruction and the interpretation of chains. In Reich, I. & et al., editors, Proceedings of Sinn und Bedeutung, volume 15, pages 515–529, Saarbrücken, Germany. Universaar-Saarland University Press.
- Sabbagh, J. (2007). Ordering and linearizing rightward movement. Natural Language & Linguistic Theory, 25(2):349–401.
- Sampson, G. (1975). The single mother condition. Journal of Linguistics, 11:1–11.
- Schwarzschild, R. (2002). Singleton indefinites. Journal of Semantics, 19(3):289–314.
- Stabler, E. (1997). Derivational minimalism. In Retoré, C., editor, Logical Aspects of computational linguistics. Springer.
- Stanton, J. (2016). Wholesale late merger in A'-movement: evidence from preposition stranding. Linguistic Inquiry, 47:89–126. <http://ling.auf.net/lingbuzz/002131>. To appear in *Linguistic Inquiry*.
- Szabolcsi, A. (2011). Quantification. Cambridge University Press.
- Takahashi, S. & Fox, D. (2005). Maxelide and the re-binding problem. In Georgala, E. & Howell, J., editors, Proceedings of SALT XV, pages 223–240, Ithaca, NY. Cornell University.
- Takahashi, S. & Hulsey, S. (2009). Wholesale Late Merger: Beyond the A/A' distinction. Linguistic Inquiry, 40:487–526.
- Toman, J. (1985). A discussion of coordination and word-syntax. In Toman, J., editor, Studies in German grammar, pages 407–432. Foris, Dordrecht.
- Tsai, W.-T. D. (1994). On nominal islands and LF extraction in Chinese. Natural Language & Linguistic Theory, 12(1):121–175.
- Šaumjan, S. & Soboleva, P. A. (1963). Applikativnaja Poroždajuščaja Model' i Isčislenie Transformacij v Russkom Jazyke. Izdatel'stvo Akademii Nauk SSSR, Moscow.
- Wexler, K. & Culicover, P. (1980). Formal Principles of Language Acquisition. MIT Press.
- Whitman, N. (2009). Right-node wrapping: Multimodal categorial grammar and the “friends in low places” coordination. In Hinrichs, E. & Nerbonne, J., editors, Theory and evidence in semantics, pages 235–256. CSLI.
- Wilder, C. (1997). Some properties of ellipsis in coordination. In Alexiadou, A. & Hall, T. A., editors, Studies on Universal Grammar and typological variation, pages 59–107. Benjamins, Amsterdam.
- Wilder, C. (1999). Right-node raising and the LCA. In Bird, S., Carnie, A., Haugen, J. D., & Norquest, P., editors, Proceedings of WCCFL 18.
- Williams, E. (1990). The ATB theory of parasitic gaps. The Linguistic Review, 6:265–279.

Yatabe, S. (2001). The syntax and semantics of left-node raising in japanese. In Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar, pages 325–344.