The Linear Limitations of Syntactic Derivations

by

Colin Pierce Bryon Davis

B.A., Linguistics, University of Minnesota, 2014

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

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2020

Author . . . .
Department of Linguistics and Philosophy
September 8, 2020

Certified by . . . .
Danny Fox
Anshen-Chomsky Professor of Language & Thought
Thesis Supervisor

Certified by . . . .
Sabine Iatridou
Professor of Linguistics
Thesis Supervisor

Certified by . . . .
David Pesetsky
Ferrari P. Ward Professor of Modern Languages and Linguistics
Thesis Supervisor

Accepted by . . . .
Kai von Fintel
Andrew W. Mellon Professor of Linguistics
Linguistics Section Head

The Linear Limitations of Syntactic Derivations

by

Colin Pierce Bryon Davis

Submitted to the Department of Linguistics and Philosophy
on September 8, 2020, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

In this dissertation, I identify and analyze several new generalizations about how phrasal displacement and discontinuity are constrained in natural language. These patterns reveal, I argue, that many limitations of syntactic derivations are attributable to the way in which syntactic operations are cyclically interleaved with the component of the grammar that establishes word order. This finding has consequences for many phenomena, and for the architecture of grammar in general.

Chapter 1 introduces the theory of cyclic linearization that serves as the foundation for this work, and several principles about the locality of movement that importantly interact with it. Chapter 2 shows that these concepts predict the cross-linguistic distribution of stranding in intermediate positions. Chapter 3 extends these considerations to an analysis of possessor extraction in colloquial English, a phenomenon subject to numerous intricate but systematic limitations. Chapter 4 provides further evidence for the theory advanced here from constraints on sub-extraction in Russian. Chapter 5 argues that certain facts about parasitic gaps in sentences with overlapping moved phrases reveal further evidence that linearization constrains syntactic derivations, with consequences for the nature of movement more generally. Chapter 6 argues that linear order constrains extraposition, and proposes an account of this phenomenon that addresses a number of puzzles about its distribution. Chapter 7 diverges from the theme of linearization to explore parasitic gaps in relative clauses, which connect to several results from chapter 5. Chapter 8 summarizes the findings of this dissertation, and discusses several more general implications of the framework advanced here.

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

Thesis Supervisor: Sabine Iatridou
Title: Professor of Linguistics

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

Acknowledgments

Looking back over these last 5 years, I realize that I took the first step on the path of research that would lead to this dissertation, and many related endeavors, on a certain unremarkable night in the fall of 2016. Since then I have often felt overwhelmed and under-slept, but I can see that pushing myself to my limits during this time has made me stronger and more capable than I would have otherwise been. Dedicating so many years to the intense study of this rather specific field of science has turned out to be a good decision, in no small part due to the intensity of the people in the linguistics department at MIT.

For positively affecting the work in this dissertation in one way or another, I thank Adam Albright, Andrei Antonenko, Itai Bassi, Jonathan Bobaljik, Tanya Bondarenko, Željko Bošković, Kenyon Branan, Noam Chomsky, Justin Colley, Stephen Crain, Patrick Elliott, Michael Yoshitaka Erlewine, Suzana Fong, Naomi Francis, Martin Hackl, Heidi Harley, Laura Kalin, Roni Katzir, Heejeong Ko, Filipe Kobayashi, Loes Koring, Anton Kukhto, Nick Longenbaugh, Travis Major, Martina Martinović, Dana McDaniel, Andrew McInnerney, Michelle Sheehan, Shigeru Miyagawa, Takashi Morita, Mitya Privoznov, Phillip Shushurin, Juliet Stanton, Donca Steriade, Abdul-Razak Sulemana, Shoichi Takahashi, Gary Thoms, Henrik Torgersen, Harold Torrence, Coppe van Urk, Bartosz Wiland, Michelle Yuan, Stanislao Zompì, and Erik Zyman. Chapters 2 and 3 of this dissertation are enriched versions of material respectively published in *Glossa* and *Linguistic Inquiry*, so it's only proper that I express thanks for the feedback I received from those journals. Similarly, reviewers for and audiences at the following conferences have all affected this research: the Linguistic Society of America's 92nd, 93rd and 94th annual meetings, West Coast Conference on Formal Linguistics 36 and 38, Formal Approaches to Slavic Linguistics 27 and 28, Workshop on Altaic Formal Linguistics 14, North East Linguistics Society 49 and 50 (and shortly 51 also), and Chicago Linguistic Society 55. Andrei Antonenko and Tanya Bondarenko deserve additional credit, since collaborative work with them on Russian serves as the basis for chapter 4.

For fun times, I thank Andrei Antonenko, Ömer Demirok, Itai Bassi, Tanya Bondarenko, Christopher Baron, Isa Bayırlı, Kenyon Branan, Justin Colley, Patrick Elliott, Enrico Flor, Boer Fu, Yadav Gowda, Peter Grishin, Filipe Hisao Kobayashi, Anton Kukhto, Katya Morgunova, Sophie Moracchini, Roger Paul, Mitya Privoznov, Erin Olson, Vincent Rouillard, Leo Rosenstein, Phillip Shushurin, Juliet Stanton, Yasu Sudo, Henrik Torgersen, Adina Williams, Gregor Williams, Michelle Yuan, Stanislao Zompì, and Hanzhi Zhu. There are many stories I could tell about the non-academic drama and mischief that has occurred during these 5 years. However, in order to maintain the privacy and dignity of those involved, I will leave that to the reader's imagination.

Now for some miscellaneous thanks. I thank the MIT linguistics department for their exceptional support during the bizarre circumstances of 2020, even though I fortunately did not have to take full advantage of it. I thank the department staff (Mary, Matt, Leo, Jen, Aggie, Chris, Christie, in no particular order) for being friendly and dependable. I thank my undergraduate advisor Claire Halpert for warning me that graduate school is a thorny path, and I thank myself for taking that path anyway. I thank Roger Paul for his superhuman effort to introduce a certain scruffy group of linguists to the Passamaquoddy community, though time constraints have prevented any Passamaquoddy from making it into this dissertation. I also thank whiskey, primarily Scotch, secondarily bourbon. Finally, I thank Tanya once more, for reasons she already knows.

I apologize to my cohort-mates (Rafael Abramovitz, Itai Bassi, Justin Colley, Suzana Fong, Verena Hehl, Maša Močnik, and Mitya Privoznov) for being the loudest one in the office and for having sticky hands, to my parents and various friends for often being a reclusive workaholic, and to anybody whose name I forgot to mention above but should have.

These acknowledgments would not be complete without thanking, in traditional alphabetical order, my main advisors at MIT: Danny Fox, Sabine Iatridou, David Pesetsky, and Norvin Richards. I have learned different things from each of them. Danny has a special ability to cut right to the heart of any issue, and from him I have gained a much deeper appreciation for the value and power of good theorizing. Sabine has helped me to remain grounded, methodical, and coherent, as well as to navigate the tumult of PhD student life and the process of turning into a grown-up linguist. David has taught me to have the courage to push the boundaries of what we think we know, to flip puzzles around until their weak point is exposed, as well as numerous nuggets of wisdom about writing, presenting, and being an academic in general. Norvin has helped me to develop as a fieldwork researcher, to be analytically agile, and to draw connections between disparate cross-linguistic facts. I am grateful to have been able to learn from and with these people over these last few years. I can tell that they've done something right, since despite the fact that I wish it weren't time to leave MIT, I can tell that I'm ready to move on.

Table of Contents

List of Abbreviations

1	1st person
2	2nd person
3	3rd person
a.o.	and others
ACC	accusative
AUX	auxiliary
CL	cyclic linearization
CLF	classifier
COP	copula
DAT	dative
GEN	genitive
LBE	left branch extraction
NEG	negation
NOM	nominative
PG	parasitic gap
PL	plural
POSS	possessive
PST	past
REL	relativizer / relativization
SG	singular

Chapter 1

The framework

1 Introduction

This dissertation explores a number of new generalizations about how phrasal displacement and discontinuity are limited in human language, which I argue lead to a deeper understanding of several fundamental principles of grammar. Empirically, this research deals with phenomena including (but not limited to) pied-piping, stranding, sub-extraction, extraposition, and the interpretation of parasitic gaps. By examining these topics, and the various aspects of syntactic theory with which they are intertwined, this dissertation amasses a body of evidence in support of one general idea with many implications—that the limitations on possible sentences are, to a significant degree, attributable to the interleaving of purely structural syntactic processes with the linear morpho-phonological component of the grammar.

Here I advance a theory in which at certain points in the syntactic derivation, the content of the structure so far built is evaluated by morpho-phonology and thus, among other processes, assigned linear ordering. This is termed *linearization*. A general framework along these lines is assumed by a great deal of recent work in syntactic theory. However, this dissertation departs from the majority of work in this domain by arguing that the process of establishing linear ordering is directly responsible for many pervasive and principled restrictions on what syntactic derivations can achieve. Fundamental to this perspective is the hypothesis that once a given cycle of the derivation undergoes linearization, the relative ordering established for its contents must be preserved by any later word-order changing syntactic processes.

In this dissertation I follow much preceding work in taking the relevant cycles of the derivation where linearization applies to be *phases*. In what follows I introduce the basics of phase theory, before proceeding to describe the framework advanced in this dissertation in greater detail.

1.1 Content of this chapter

Next, section 2 provides background on the basics of phase theory, and the hypothesis of linearization by cyclic spell-out. Section 3 describes the version of phase theory that I argue for in this dissertation, and summarizes its core predictions. Section 4 introduces several concepts about the locality of syntactic movement, which we will see throughout this dissertation interact, and conflict, with the requirements of linearization. Finally, section 5 concludes by briefly previewing the main results of each of the following chapters.

2 Cyclic spell-out and phase theory

Chomsky (2000, 2001, 2008, 2013) and much following work[1] argues that syntactic structures are derived bottom-up by successive applications of the operation *Merge*,[2] and that at certain points in this process, the structure built so far undergoes *Spell-Out*. The Spell-Out operation subjects the structure in question to semantic composition and assigns it morpho-phonological form, thus yielding its Logical Form (LF) and Phonological Form (PF).[3] By hypothesis, these representations must be established in order to allow what would otherwise be an abstract syntactic structure, consisting only of formal features, to interface with (at least) two independent cognitive systems. A structure's LF is interpreted by the conceptual-intentional system, which is concerned with meaning. A structure's PF is referenced by the sensori-motor system, which determines how the structure will be expressed by the body in the external world—primarily via speech in the case of spoken language. The way in which these systems and representational levels are interconnected is diagrammed in (1) below:

---

[1]For my purposes, summarizing the basics of Chomsky's proposals is all that is necessary, but see Citko (2014) for an overview of the large body of research in this vein.

[2]In particular, a given syntactic element is introduced into the structure in progress via *external* Merge, whereas relating an element already in the structure to an additional position within it is accomplished via *internal* Merge, generally termed "movement".

[3]In many works, the term "Spell-Out" is used to refer to the mapping to both LF and PF as I have just described, whereas in others the terminology "Transfer" is used in this sense, leaving the former term to refer strictly to the mapping to PF. Since this dissertation focuses on the relationship between syntax and the requirements of PF, I will maintain use of the term "Spell-Out" here.

(1)   *An architecture for natural language*

<div align="center">

Syntactic Structure

│

┌─────────────┐
│ Spell-Out │
└─────────────┘

Logical Form   Phonological Form

Conceptual-Intentional System   Sensori-Motor System

</div>
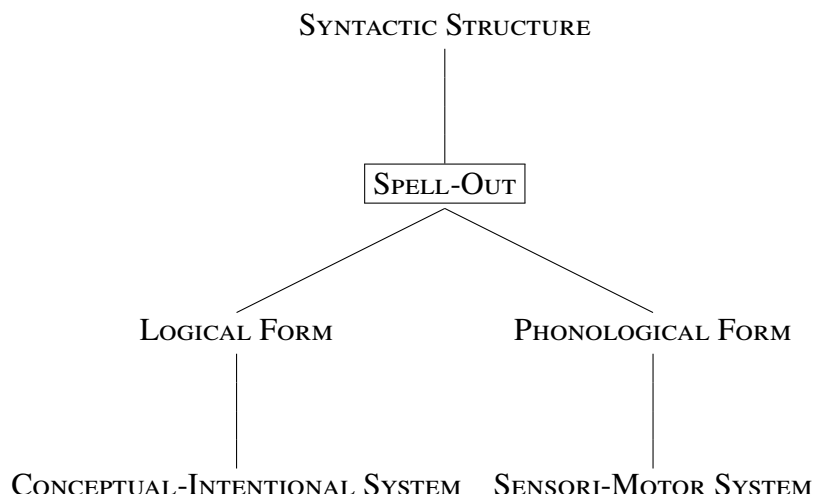
For a syntactic structure to be expressed as a spoken utterance, spell-out must determine at least two sets of information about it. Foremost it is necessary to establish how the structure is to be pronounced, by assigning morpho-phonological content to it.[4] Since it is only possible to pronounce a certain amount of material at once, the elements of a structure must, for the most part, be pronounced one at a time. In order for this to occur, it is necessary to also decide in what temporal order those elements will be uttered. This is precisely what the linearization process achieves. As previewed above, the interaction between the syntactic derivation and the demands of linearization is central to this dissertation. Before describing the approach to linearization that I defend here, however, it is necessary to first overview several important concepts about spell-out from Chomsky's work.

Chomsky's research initiated the hypothesis that spell-out is triggered by certain dedicated phrases termed *phases*, which, once built, submit the content of their complement to LF and PF. For most works that adopt phase theory, at least CP and vP are taken to be phases. Under this view, most clauses contain at least two phases, and therefore reach their final cumulative LF and PF incrementally in at least two steps, in cyclic phase-by-phase fashion. This is schematized in (2) below, where boxes demarcate the domains to which each instance of spell-out has applied:[5]

---

[4]It is often assumed that morpho-phonological form is only assigned to the terminal nodes of a syntactic structure. This concept is a standard axiom of the Distributed Morphology approach to the syntax-morphology mapping (Halle and Marantz 1993; Harley and Noyer 1999, and many others), and this is what I will assume here as well.

[5]For convenience I will generally only attach category labels to heads and maximal projections, not to intermediate (X') nodes.

(2)   *Phase-by-phase spell-out of the clause (Chomskyan phases)*

    a.   *#1: Spell-out of VP triggered by v*

vP
[PHASE]

v   VP

V O
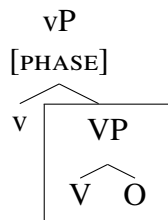
    b.   *#2: Spell-out of TP triggered by C*

CP
[PHASE]

C   TP

S

T vP

[PHASE]

v   VP

V O
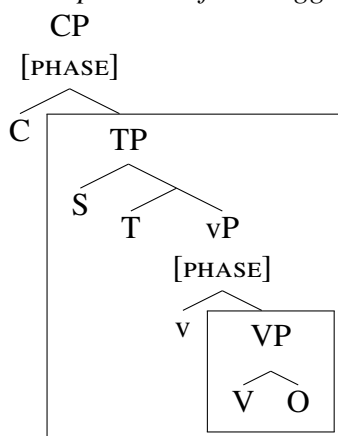
As argued by a great deal of work in this vein, the concept that spell-out applies to only the complement of a phase explains a variety of facts about the locality and cyclicity of syntactic operations. However, such results depend on one further hypothesis—that once a phase's complement undergoes spell-out, its contents cannot be referenced by later syntactic operations. This constraint is known as the *Phase Impenetrability Condition* (PIC).[6] One phenomenon often argued to be a consequence of the PIC is *successive-cyclic* movement, which has been identified and analyzed by a significant amount of research (Chomsky 1973, 1977, 1986; Du Plessis 1977; Henry 1995; Cole and Hermon 2000; McCloskey 2000, 2001, 2002; Nissenbaum 2000; Legate 2003; Sauerland 2003; Bruening 2001, 2006; Barbiers 2002; Torrence 2012; Abels 2003, 2012; Wiland 2010; Henry 2012; van Urk 2015; van Urk and Richards 2015; Korsah and Murphy 2019; Davis 2019, 2020, and many others). An instance of movement is successive-cyclic if it does not proceed directly to its termination point, but rather involves multiple shorter movement steps,

---

[6]This discussion is framed in terms of the "strong" PIC of Chomsky (2000), for which spell-out is triggered as soon as the head of the next projection above the relevant phase is merged. There are other proposals about the timing of spell-out, such as the "weak" PIC in Chomsky (2001). I will not dwell on these differences, since this dissertation will not pursue Chomsky's theories of spell-out. These concepts will be discussed in more detail in chapter 8, however.

as shown in (3). While the schema in (3) shows a three step movement path, a given movement path only needs one intermediate step to qualify as successive-cyclic:[7]
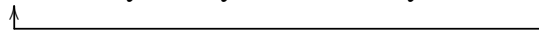
(3) *A successive-cyclic movement path*
$$[_{XP} \ \alpha \ \text{X} \ ... \ [_{YP} \ t \ \text{Y} \ ... \ [_{ZP} \ t \ \text{Z} \ ... \ t \ ]$$

Successive-cyclic movement is a central topic of this dissertation. To see how the theory under discussion predicts that movement will sometimes require intermediate steps in this way, let's consider what must occur to facilitate *wh*-movement from an embedded CP, for instance, assuming that CP is a phase (and leaving the phasehood of vP aside for convenience).

   If syntactic structures are constructed bottom-up by successive applications of merge, then in a sentence like (4) below, the embedded CP will have been completed before the construction of the matrix CP begins. In principle, there is no reason why the moved *wh*-phrase here could not have remained in situ in the lower VP right until the point at which the matrix C merges and triggers *wh*-movement to its specifier, as the arrow here shows:

(4) *Hypothetical non-successive-cyclic movement from an embedded CP*
$$[_{CP2} \ \text{What did you say} \ [_{CP1} \ \text{that you ate} \ t \ ]]?$$

While a string like that in (4) is indeed acceptable in English, under the theory of phases outlined above, the hypothetical derivation just described is not predicted to be available. This is because once the embedded CP is constructed, spell-out of its complement will be triggered, as represented by the box around TP in (5a) below. If the *wh*-phrase has not yet moved from TP at this point, then it will be trapped due to the PIC, and thus will be unable to later move into the matrix clause (5b):

(5) *PIC versus non-successive-cyclic movement from CP*

   a. *Spell-out of embedded TP upon completion of embedded CP*
       $[_{CP1_{[Phase]}}$ that $[\boxed{\text{TP}}$ **you ate what** $]]$
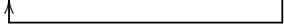
   b. *PIC bans movement into matrix CP from spelled-out TP*
       * **What** did you say $[_{CP1}$ that $[\boxed{\text{TP}}$ **you ate** t $]]$?

In contrast, if the *wh*-phrase moves to the specifier of the embedded CP (6a) before C triggers spell-out of TP (6b), then the *wh*-phrase successfully avoids being prematurely trapped by the PIC. This allows the *wh*-phrase to move beyond the clause

---

[7]I generally use the notation "*t*" to mark gaps left by movement, though I will also use "__" on occasion. When indices are required I typically use numbers rather than letters.

in which it originated and land in the matrix clause (6c):

(6)   *Successive-cyclic movement to CP edge feeds further movement*

    a.   *Movement to embedded CP edge*
        [$_{CP1}$ What that [$_{TP}$ you ate *t*]]

    b.   *Spell-out of embedded TP*
        [$_{CP1}$ What that [TP **you ate *t***]]

    c.   *Movement from embedded CP edge succeeds*
        ✓ [$_{CP2}$ What did you say [$_{CP1}$ *t* that [TP **you ate *t***]]]?

In general, these concepts predict that a phrase moving from the complement of any phase must form an intermediate landing site in the phase edge, in successive-cyclic fashion, before it can move beyond that phase. Indeed, many of the works cited above argue that movement from CP, in addition to certain other phrases, is successive-cyclic in the way expected if such constituents are phases.

    While the empirical basis for successive-cyclicity is strong (and will be added to by several results of the coming chapters), the explanation for it remains a topic of ongoing research. The version of phase theory just described attributes successive-cyclicity to a need for moving phrases to avoid the spell-out of each phase they exit. This dissertation will depart from this proposal, by advancing an approach under which moving phrases in fact never escape the spell-out of the phases they pass through, as described in the next section. As we are about to see, the approach defended here argues that effects like successive-cyclic movement are not about bypassing spell-out. Instead, this dissertation's approach argues that successive-cyclic movement occurs to ensure that the linearization information that cyclic phase-by-phase spell-out establishes is contradiction-free.

3   Cyclic linearization and order preservation

This dissertation advances a theory of phases generally referred to as *Cyclic Linearization* (CL; Fox and Pesetsky 2005a,b; Takahashi 2004; Ko 2007, 2011, 2014; Müller 2007; Sabbagh 2007; Podobryaev 2009; Takita 2010; Fanselow and Lenertová 2011; Jenks 2011, 2013a,b; Medeiros 2013; Overfelt 2015; Erlewine 2017; Davis 2019, 2020, a.o.). CL shares with the phase theory just discussed the idea that spell-out applies to syntactic derivations in phase-by-phase fashion, in addition to the general idea that the properties of spell-out are importantly intertwined with the motivation for successive-cyclic movement. However, CL differs in numerous other ways which, as we'll see, lead this theory to attribute a particularly great deal of

explanatory power to the way in which cyclic spell-out determines linearization. In this dissertation, I will argue that the same laws of linearization that for CL predict the distribution of successive-cyclic movement (and a variety of other phenomena) also explain several new generalizations about patterns of displacement and discontinuity in a variety of languages. While the analysis of these patterns will lead to consequences for many domains of syntactic theory, the influence of CL is a pervasive theme throughout this dissertation.

One important distinction of the CL approach concerns where spell-out applies. In particular, CL hypothesizes that a given instance of spell-out does not only apply to the complement of the phase that triggers it (7a), but rather applies to the entire phase (7b), including any constituents in its edge:[8]

(7) *Two approaches to spell-out*

    a. *Complement-only spell-out (Chomsky)*
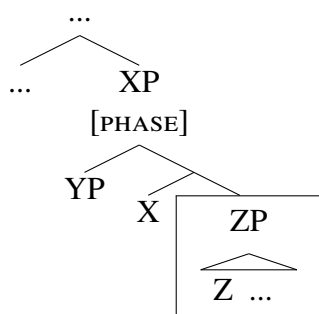
    b. *Spell-out of entire phase (CL)*
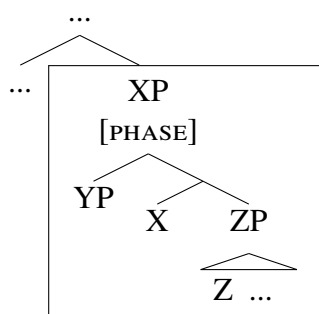
If no material in a given phase is exempt from spell-out of that phase, then it cannot

---

[8]Grohmann et al. (2017) observe that under the phase theory in Chomsky's work, the fact that material in the edge of a matrix CP is not dominated by a subsequent phase makes it unclear how such material is spelled-out. This puzzle does not arise under CL, for which spell-out within a given phase applies to all of its contents. While this is a welcome result, it raises an additional puzzle if Fitzpatrick (2006) is correct that the possibility of not spelling-out the edge of a matrix CP accounts for the availability of certain truncated questions like "*(Do) you have a moose?*". While intriguing, it is not clear that a phase-centric analysis of such sentences is the only possibility, however.

be the case that successive-cyclic movement serves to allow a moving phrase to avoid spell-out. Furthermore, if spelled-out constituents are inaccessible for later syntactic operations as the PIC describes, then the hypothesis that entire phases undergo spell-out incorrectly predicts the impossibility of ever moving from a phase, regardless of whether movement takes advantage of the phase edge or not. For this reason, CL abandons the PIC, and hypothesizes that previously spelled-out material remains accessible in principle for the entire derivation.[9] This set of assumptions is straightforwardly incompatible with the explanation for successive-cyclicity furnished by the phase theory in Chomsky's work. As expected, work adopting CL argues in favor of a different explanation—one that derives successive-cyclic movement, and a variety of related effects, from the principles of linearization.

For CL, successive-cyclic movement serves to bring a moving phrase to the linear edge of each phase it passes through, as a means of achieving a contradiction-free final ordering for the derivation in question. This understanding depends on the following hypothesis about the nature of linearization information:

(8)   *Order Preservation*
        Information about linearization, once established at the end of a given Spell-Out domain (=phase), is never deleted in the course of a derivation.
        (Fox and Pesetsky 2005a, p. 6)

Assuming (8), once a phase spells-out and establishes a relative linear ordering for the elements it contains, that ordering information must be retained, and hence cannot be overridden if the order of those elements is eventually reversed by a later movement operation. Rather, if such later order reversal occurs, the previously established order and the order derived later on will contradict each other, resulting in a crash at PF. In other words, (8) entails that the ordering information established upon the spell-out of a given phase must be preserved by any subsequent movement that occurs in the derivation in question.

To see concretely what this line of reasoning predicts, next I demonstrate in detail why successive-cyclic movement via phase edges satisfies order preservation, and thus yields a contradiction-free PF.

3.1   Successive-cyclicity and Cyclic Linearization

As an initial demonstration of how the CL approach functions, let's examine a derivation with non-successive-cyclic *wh*-movement from an embedded CP once

---

[9]See Martinović (2019) for a recent argument that spelled-out constituents indeed remain accessible for later syntactic operations, based on an examination of the interaction of syntactic movement and post-syntactic affixation in Wolof.

again, but this time with CL in mind. As before, I will speak as if only CPs are phases for convenience. The relevant hypothetical derivation is shown once more in (9) below, where the *wh*-moving object *what* moves to the matrix spec-CP without forming an intermediate landing site in the specifier of the embedded CP:

(9)   *Non-successive-cyclic movement from an embedded CP*
        [$_{CP2}$ What did you say [$_{CP1}$ that you ate *t*]]?

Here *what* had not moved to the edge of the embedded CP at the time when it was completed and then spelled-out, as shown in (10) below. Therefore in this situation, spell-out of the embedded CP orders *what* to the right of the rest of the content of the embedded clause, as (10b) shows. Here and throughout this dissertation, I use the notation "<" to encode linear precedence relations. An ordering statement [$\alpha <$ β] means "α linearly precedes β". If α and β are internally complex, [$\alpha < \beta$] means "the nodes dominated by α linearly precede the nodes dominated by β" (though this level of detail will usually be unnecessary).[10]

(10)   *Spell-out of embedded CP without successive-cyclic movement*
        a. *Construction of embedded CP*
            [$_{CP1}$ that you ate **what**]
        b. *Resulting linearization*
            that < you < ate < **what**

Next in this derivation, the matrix clause is constructed, and then *what* moves in one step to the specifier of the matrix CP, as (9) diagrammed. After this movement occurs, spell-out of the matrix CP produces the ordering information in (11):

(11)   *Ordering in matrix CP*
        what < did < you < say < [embedded CP]

Notice that in (10b), *what* was determined to follow everything in the embedded CP. However, in (11) *what* was subsequently established to precede everything in the matrix CP, and thus to ultimately precede everything in the embedded CP as well, since the content of the matrix CP must precede that of the embedded one. Here we have a contradiction: this derivation has established that the moved phrase *what* must simultaneously follow and precede the content of the embedded

---

[10]What "<" encodes is relative linear precedence, not strict adjacency or concatenation. Thus an ordering [$\alpha < \beta$] is consistent with α later moving away from β, with the result that other material ultimately intervenes between them, as in [α γ β]. This is because α still precedes β after such movement, despite no longer being adjacent to it. While adjacency is not a primitive of CL, chapter 3 will argue for a circumstance under which adjacency also interacts with linearization.

23

CP. Given that *what* cannot be pronounced in a position that is consistent with both of these ordering instructions, CL predicts that this scenario gives rise to an irreconcilable linearization problem. That is, order preservation requires that both of these instructions for linearization be acknowledged, but because they contradict each other, the result is necessarily indigestible by PF.

Importantly, successive-cyclic movement through the linear edge of the embedded CP, as in (12), prevents this ordering contradiction from arising:

(12)   *Successive-cyclic movement from embedded CP*
        [$_{CP2}$ What did you say [$_{CP1}$ *t* that you ate *t*]]?

This is because by moving to the edge of the embedded CP (13a), *what* is linearized at the edge of this CP when it spells-out (13b) before later moving on to the matrix spec-CP (13c) and being linearized there (13d):

(13)   a.   *Movement to linear edge of embedded CP*
              [$_{CP1}$ What that you ate *t*]

        b.   *Ordering in embedded CP after movement to its edge*
              what < that < you < ate

        c.   *Movement from embedded CP edge to matrix spec-CP*
              [$_{CP2}$ What did you say [$_{CP1}$ *t* that you ate *t*]]?

        d.   *Ordering in matrix CP*
              what < did < you < say < [embedded CP]

Notice that the ordering information established by movement to the embedded CP edge (13b) and later movement to the matrix CP edge (13d) are in agreement: both sets of ordering information establish that *what* precedes their contents. This is consistent with a final representation where *what* is pronounced at the left edge of the sentence, preceding everything else in the derivation, as it does in reality.

In summary, if a moving constituent passes through the linear edge of each phase it exits, then phase-by-phase spell-out will determine that this constituent precedes the content of all phases it has passed through. This is ultimately consistent with the moved constituent being pronounced in its final landing site, where it precedes the content of all phases that it has crossed over. The CL theory thus predicts such a successive-cyclic derivation to yield a licit PF. This logic is not specific to *wh*-movement from CP, but generalizes to any sort of overt movement from any phase. We will see throughout this dissertation that taking advantage of linear edges is one way to derive licit movement from a phase. Importantly, CL also predicts that

there are other ways to accomplish cross-phasal movement. Next I describe another predicted movement strategy that is central to this dissertation.

3.2 Beyond successive-cyclicity

While CL predicts that successive-cyclic movement of the form described above is one way to achieve a coherent linearization for movement from a phase, the same principles that allow this result also predict other ways of achieving licit cross-phasal movement. This is because what is important for the CL approach is not successive-cyclicity per se, but rather avoidance of ordering contradictions. Consequently, CL predicts that non-successive-cyclic movement from a phase is possible, provided that it is accompanied by additional movement operations that serve to keep the result contradiction-free.

To illustrate this point, let's consider what the two phase theories discussed in this chapter predict about movement from a phase with multiple constituents in its edge. For the theory initiated by Chomsky's work, in which spell-out triggers the PIC but only applies to the phase's complement, notice that nothing in principle bans extraction of a lower specifier of a given phase across a higher one. This is illustrated with movement of β across α from the phase XP in (14) below. Since such movement is initiated from a complement-external position in XP, we do not expect the PIC to prevent this movement from occurring:[11]

(14)   *Prediction of the complement spell-out + PIC theory*
   $\checkmark$ $[_{YP_{[Phase]}}\ \beta\ \mathrm{Y}\ ...\ [_{XP_{[Phase]}}\ \alpha\ t\ \mathrm{X}\ ...\ ]]$

In contrast, under the CL approach such a derivation is predicted to run into a linearization contradiction, in the following way. First, upon the completion of the phase XP, β is linearized in its original position, following α, as in (15a) below. Second, β moves from XP in one step, crossing α and landing in a position preceding XP and therefore α as well. Thus when the next phase YP spells-out and linearizes β in its derived position, β is found to precede α, as shown in (15b). This result contradicts the relative order of these elements that was previously established upon the spell-out of XP:[12]

---

[11]Assuming that α is not independently an intervener for the relevant movement dependency with β, as would be the case if both α and β were interrogative *wh*-phrases, for instance.

[12]Of course, there would be no problem in (15) if β successive-cyclically moved to a position within XP preceding α before moving on, but the purpose of this discussion is to consider what we predict if a given instance of movement is, for one reason or another, not successive-cyclic. If α and β here are specifiers of XP, then the possibility of such successive-cyclic movement of β above α is in fact ruled out, if movement from one specifier to another of the same phrase is banned (Ko 2007,

(15)   *Prediction of CL: No direct extraction of lower specifier of a phase*[13]
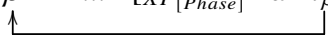
      a.   *Step #1: Ordering of phasal XP with two specifiers*

          $[_{XP_{[Phase]}} \quad \alpha \quad \beta \quad X \quad ... \;]$

          Linearization of XP: $\underline{\alpha < \beta} < X$

      b.   *Step #2: Movement from non-edge of XP yields contradiction*

          $[_{YP_{[Phase]}} \quad \beta \quad Y \quad ... \quad [_{XP_{[Phase]}} \quad \alpha \quad t_\beta \quad X \quad ... \;]]$
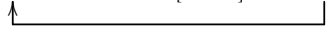
          * Linearization of YP: $\beta < Y < XP \rightarrow \underline{\beta < \alpha}$

Importantly, CL predicts that a configuration like (15b) can be rescued, provided that additional order-restoring movement occurs soon enough to avoid the linearization contradiction that would occur if (15b) were to remain as it is. Specifically, if α also moves into the higher phase YP and lands in a position preceding β, then the fact that β crossed α on the way out of XP no longer matters. This is because such movement of α causes it to precede β within YP just as it did within the previous phase XP, as illustrated in (16):

(16)   *Prediction of CL: Salvation by order-restoring movement*
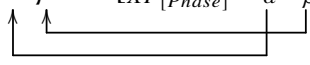
      a.   *Direct extraction from non-edge...*

          $[_{YP_{[Phase]}} \quad \beta \quad Y \quad ... \quad [_{XP_{[Phase]}} \quad \alpha \quad t_\beta \quad X \quad ... \;]]$

      b.   *...repaired by order-restoring movement of crossed material*

          $\checkmark \; [_{YP_{[Phase]}} \quad \alpha \quad \beta \quad Y \quad [_{XP_{[Phase]}} \quad t_\alpha \quad t_\beta \quad X \quad ... \;]]$

          Linearization of XP: $\alpha < \beta < X$
          Linearization of YP: $\underline{\alpha < \beta} < Y < XP$

   In this way, CL predicts that movement from a phase need not be successive-cyclic. Rather, it only needs to be order-preserving. While successive-cyclic movement via the linear edge of a phase is one way to derive licit cross-phasal movement, if a given movement step does not take advantage of the linear edge, the resulting configuration will still be legal provided that any crossed-over material also moves into the next phase, to precede what previously crossed it. Both of these strategies result in consistent ordering information across all phases involved.[14]

2014). This important concept is introduced in the next section.

[13]I follow Fox and Pesetsky (2005a,b) in using the symbol "→" to clarify a relevant consequence entailed by a given linearization statement. In (15b), this arrow points out the important consequence that $\beta < XP$ entails $\beta < \alpha$ here.

[14]Fox and Pesetsky (2005a,b) originally argued for the prediction in (16) on the basis of Holmberg's Generalization (Holmberg 1986, a.o.) about object shift in Scandinavian languages. In this

4   Locality constraints on movement

Throughout this dissertation, we will see that the predictions of CL laid out above interact with several independent constraints on possible derivations. One such factor that will reveal its influence at several points is a ban on specifier-to-specifier movement within the same phrase. As Ko (2007, 2011, 2014) points out, such a ban is predicted by the concept that movement to the specifier of a given head requires that head, or more specifically a movement-triggering feature on that head, to c-command the goal phrase to be moved (Chomsky 1995, 2000, 2005, a.o.). Since heads don't c-command their specifiers, for this theory it is not possible for a head to move a phrase from one of its specifiers to another.

This concept is illustrated in (17) below. In (17a), we see that the head α c-commands its complement ϵP, but not its specifiers βP and γP or any of their contents. Thus α cannot trigger movement of γP or its sub-part δP to a position above the higher specifier βP, as (17b-c) show. In contrast, α can trigger movement from within its complement to the periphery of αP, as diagrammed in (17d):

(17)   a.   *A head does not c-command its specifier(s)*
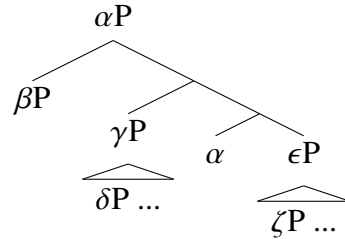
   b.   *A head cannot trigger movement of one of its specifiers*

---
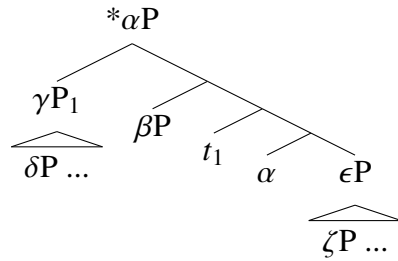
dissertation, we will see that this prediction extends well beyond Scandinavian, as expected if CL is a general property of human language.

As discussed in chapter 4, Bošković (2016) proposes a modified theory of the PIC which predicts the possibility of derivations like (16), though with a different timing of movement steps than assumed here. As we'll see throughout this dissertation, order preservation is necessary in these scenarios, a fact predicted by the framework advanced here, but not by Bošković (2016).

c. *A head cannot extract from one of its specifiers*

$$*\alpha P$$

Tree:
- αP with left daughter δP_2 and right branch
  - βP
    - γP (over t_2 ...)
    - α
    - ϵP (over ζP ...)

d. *A head can trigger movement from within its complement*

$$\checkmark \alpha P$$

Tree:
- αP with left daughter ζP_3 and right branch
  - βP
    - γP (over δP ...)
    - α
    - ϵP (over t_3 ...)
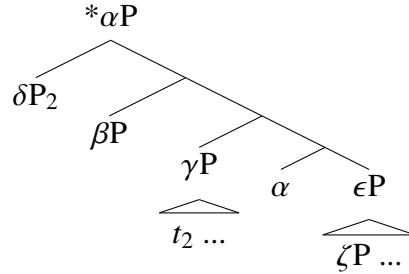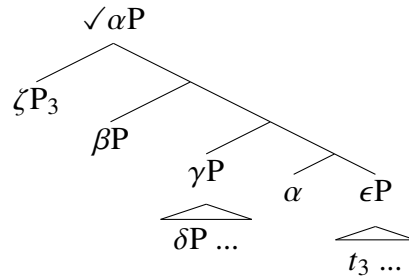
Given that a head also does not c-command any adjuncts it may have, movement of or from an adjunct within the local maximal projection is also ruled out. Thus what we really predict here is a ban on phrase-bounded movement operations making reference to what might generally be called "edge constituents". I will often speak of this constraint in terms of specifier-to-specifier movement, however, since this is what will generally be relevant to the phenomena investigated here.[15]

The ban on phrase-bounded edge-internal movement is at least partially subsumed by some formulations of *anti-locality*—the concept that movement must not be in some relevant sense too short (Bošković 1997; Ishii 1999; Grohmann 2003; Abels 2003, 2012; Erlewine 2016, a.o.). Anti-locality will be relevant at several points in this dissertation. Abels' proposals about anti-locality, in particular, also ban another relevant form of movement that was omitted from (17) above: movement of a head's complement into its edge, as diagrammed in (18) below. A ban on movement of this variety will also feature in this dissertation, primarily in chapter 2, though less frequently than the ban on edge-internal movement.[16]

---

[15]Richards (2004) argues that Bulgarian allows movement of a first *wh*-phrase to spec-CP, followed by extraction of a second *wh*-phrase out of the first, and into a second specifier of the same CP. Since this sort of movement is banned under the approach argued for here, such facts require a different account in the context of this dissertation. See for instance Frampton (2001) for a potentially compatible analysis which amounts to extraction from the lower copy of the outer *wh*-phrase. See chapter 2 for additional discussion of this topic.

[16]Put more simply, (17-18) describe a theory under which a head cannot trigger movement of

(18)   *No movement of a head's complement to its specifier*

$*\alpha P$

ϵP_4

ζP ...

βP

γP

δP ...

α t_4
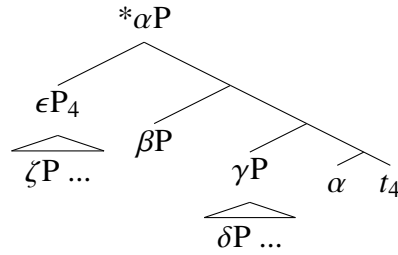
    This concludes the overview of the primary concepts that this dissertation argues for. In the following chapters, we will see a variety of circumstances in which the pressures of CL interact and conflict with the principles of locality just described. I will argue that the tension between these components of the grammar yields many correct predictions about the limitations of displacement and discontinuity. While numerous other aspects of syntactic theory will also be incorporated into this work, what has been said in this chapter is all that is needed to get started.

5 Chapter previews

Here I summarize the main topics and results of each of the following chapters. After this we will come to chapter 2, where the substance of this dissertation begins.

5.1 Chapter 2: Crossing and stranding at edges

Chapter 2 provides an initial argument for the concepts described above based on an investigation of pied-piping and stranding. In particular, here I examine the cross-linguistic character of configurations where a moved phrase leaves pied-piped material behind at an intermediate point in the path of movement. Such a derivation is schematized in (19) below, where movement of α pied-pipes and then strands β, which thus ends up in an intermediate position in the sentence:

(19)   *Intermediate stranding of β by movement of α*
         $[_{ZP} \ \alpha \ Z \ [_{YP} \ [\ t_\alpha \ \beta \ ] \ Y \ [_{XP} \ t_{\alpha\beta} \ X \ ]]]$
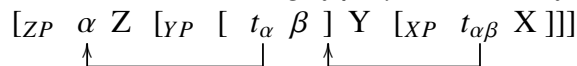
    Works on numerous languages, which this chapter summarizes, have argued

---

a phrase that it does not asymmetrically c-command. I will also occasionally refer to a locality constraint of a different variety recently explored by Erlewine (2016), with precedent in Bošković (2005) and preceding works cited therein: that when a given phrase XP is immediately dominated by another phrase YP, movement from spec-XP to spec-YP is illegally short, whereas when a distinct projection intervenes between XP and YP, movement from spec-XP to spec-YP is legal.

for the existence of such patterns. What is likely the most well-known instance of such stranding comes from McCloskey (2000), who shows that the West Ulster English variety of Ireland allows post-nominal *wh*-associated *all* to be stranded under A′-movement, either in its base position, or in a CP edge, as we see in (20):

(20)   *All-stranding in West Ulster English*
        What$_1$ (**all**) did he say [$_{CP}$ t_1 (**all**) that he wanted t_1 (**all**)]?
        (Adapted from McCloskey 2000, ex. 8)

This strandable *all* contrasts with English prepositions, which cannot be stranded in a clause edge, despite being strandable in situ (Postal 1972, 1974):
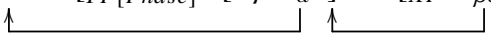
(21)   *No intermediate stranding of prepositions*
        (For) [which dog]$_1$ did they claim [$_{CP}$ (***for**) t_1 (that) I cooked an expensive dry-aged ribeye steak (for) t_1]?
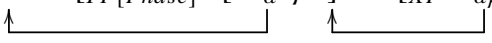
Notice that there is a word order difference between the two strandable elements just mentioned. West Ulster English strandable *all* follows the phrase that pied-pipes it, and is capable of intermediate stranding. In contrast, English prepositions precede the phrase they merge with, and are incapable of such stranding. In this chapter, I show that this correlation between word order and the possibility of intermediate stranding holds cross-linguistically, as the generalization in (22) below describes. This chapter's evidence for (22) consists of facts from English, Norwegian, Brazilian Portuguese, Wolof, Dutch, German, Afrikaans, Polish, and Russian, as well as various marginal cases from other languages.

(22)   *Intermediate Stranding Generalization (ISG)*
        Leftward movement of a phrase α can only intermediately strand an element β if β is (or can be) ordered to the right of α before stranding occurs.

Assuming that intermediate stranding is generally parasitic on successive-cyclic movement through a phase edge, I argue that this cross-linguistic generalization emerges naturally from CL. This is because intermediate stranding of a preposition, or any other element that precedes the phrase that pied-pipes it, will require the pied-piped element to be crossed over by the movement step that strands it in a phase edge (23a). In contrast, this issue does not arise for intermediate stranding of material that originally follows the phrase that pied-pipes it (23b). Since CL generally bans phase exiting movements that do not proceed from the linear edge of the phase, this contrast, and the generalization above, is straightforwardly predicted.

(23) a. ** Intermediate stranding with crossing at the edge*
      $[_{ZP}$ α Z $[_{YP[Phase]}$ [β $t_α$] Y $[_{XP}$ $t_{βα}$ X]]]

   b. ✓ *Intermediate stranding without crossing at the edge*
      $[_{ZP}$ α Z $[_{YP[Phase]}$ [$t_α$ β] Y $[_{XP}$ $t_{αβ}$ X]]]

Additionally, I show that anti-locality and the ban on phrase bounded edge internal movement accurately rule out potential derivations that would generate unattested intermediate stranding in a CL-obeying manner. In particular, these constraints on movement importantly prevent re-arranging the content of phase edges or of (some) pied-piped phrases as a means of avoiding the crossing problem illustrated in (23a). Consequently, the robustness of the ISG as a surface-level generalization about intermediate stranding stands as evidence for CL, as well as for these limitations on the locality of movement. This chapter will also examine further predictions about the positions that are licit for stranded material to surface in, though the cross-linguistic distribution of stranding positions remains more complex than this chapter will be able to fully account for.

5.2 Chapter 3: Possessor extraction in English

In chapter 3, I investigate a form of possessor extraction available in the colloquial English of some speakers. This extraction separates the possessor from the Saxon genitive morpheme [*'s*], as illustrated in (24a) below. While received wisdom states that English is among the many languages that disallow possessor extraction and therefore require pied-piping of the entire possessive DP instead (24b), my research has observed that many speakers permit configurations of the form in (24a). Such extraction is generally characterized as informal or colloquial.

(24) a. *Possessor extraction in colloquial English*
      % Mary is the one $[_{CP}$ **who**$_1$ they said $[_{CP}$ [__$_1$**'s new book**] is good]].

   b. *Standard English possessum pied-piping*
      Mary is the one $[_{CP}$ [**whose new book**]$_1$ they said $[_{CP}$ __$_1$ is good]].

This form of possessor extraction is subject to a number of detailed restrictions which are clear and consistent for speakers who accept the construction. These are unlikely to be a set of memorized quirks, since this phenomenon is far outside the domain of prescriptive English. Rather, they must stem from more general principles of grammar. In this chapter, I argue that these restrictions emerge straightforwardly from the concepts defended in the previous chapter, along with an additional condition relating to the morpho-phonological requirements of [*'s*].

Notice that in (24a), the possessum DP is stranded in an embedded clause. It turns out that such possessor extraction is in fact only possible for cross-clausal movement, and thus never occurs clause-internally. Another robust but surprising constraint on this construction is the fact that it disallows extraction from a non-subject possessum that remains in situ (25a), but rather requires such a possessum to be displaced to the edge of the local clause (25b):

(25)   *Obligatory displacement of non-subject possessum under PE*
         a.   * **Who**$_1$ do they think [$_{CP}$ Mary read [__$_1$**'s book**]]?
         b.   ✓ **Who**$_2$ do they think [$_{CP}$ [__$_2$**'s book**]$_1$ Mary read t_1]?

I show that these and numerous other constraints reveal the following generalization:

(26)   *Generalization about English PE and clause edges*
         A possessor must reach the left linear edge of the local CP before extracting from the possessum DP.

I argue that this generalization arises from the pressures of CL described above, and its interaction with an independent adjacency condition relating to the Saxon genitive morpheme [*'s*]. Several works have hypothesized that the typical illicitness of PE in English is attributable to a morpho-phonological constraint that penalizes movement which renders the possessor and [*'s*] non-adjacent at PF (Chomsky 1995; Radford 1997; Gavruseva and Thornton 2001, a.o.). I accept this general route of explanation for the lack of PE in standard English, and argue that the distribution of PE in colloquial English is accurately predicted by positing that this register has a version of this constraint that is evaluated phase-by-phase:

(27)   *Possessor-['s] Adjacency (Local version, available to PE speakers)*
         [*'s*] must be adjacent to its associated possessor at the spell-out of the minimal phase (vP, CP) containing [*'s*].

I will show that the tension between this adjacency constraint and the requirements of CL prevents English PE from occurring outside of the particular circumstances described by (26). The process of showing why this is the case will also lead us to further evidence for the ban on phrase-bounded specifier-to-specifier movement, which constrains the space of possible possessor extraction derivations in an important way which directly mirrors certain results of chapter 2.

Among other topics, this chapter also examines a little-analyzed relativization strategy in English which I will refer to as the *that's*-relative (Seppänen and Kjellmer 1995; McDaniel et al. 1998, 2002), previewed in (28):

(28)　　*That's-relatives*

      a.　　This is the girl [that's hat is red].

      b.　　That is the bike [that's wheel the kids broke].

I argue that in the context of this chapter's analysis of English possessor extraction, the existence of *that's*-relatives reveals new evidence for a proposal of Deal (2016) that TP is a phase in relative clauses. Other issues about phases, such as the question of the phasal status of DP, will also come to the fore in this chapter.

5.3　Chapter 4: Limiting discontinuity in Russian

In chapter 4, I switch focus to Russian. This language is ideal for testing the predictions at issue in this dissertation, given its richness in word order changing movement process. Here I integrate and extend the findings of two collaborative projects on Russian (Antonenko and Davis 2020; Bondarenko and Davis To appear), which when taken together reveal some general constraints on Russian sub-extraction that align directly with the concepts defended in the previous chapters. This chapter focuses on a variety of sub-extraction termed left branch extraction (LBE). LBE refers to extraction of an element that originates in the left periphery of the nominal domain, a process that is productive for many languages in the Slavic family, and beyond (Ross 1967; Corver 1990, 2007, a.o.). Some preliminary examples of Russian LBE are provided in (29):

(29)　　*LBE in Russian*

      a.　　**Kakuju**$_1$ ty　uvidel [$_{NP}$ t_1 košku]?
           what　　you saw　　　cat
           'What cat did you see?'

      b.　　**Miluju**$_1$ ty　uvidel [$_{NP}$ t_1 košku].
           cute　　you saw　　　cat
           'You saw a cute cat.'

In the first part of this chapter, we will examine LBE from nominal phrases with multiple left branches, expanding on facts observed in Antonenko and Davis (2020). In brief, while LBE is unrestricted for nominal phrases with only one left branch, LBE of a left branch that is preceded by another in the same nominal is typically banned, as previewed in (30). Here we see that LBE of an adjective across a demonstrative is unacceptable, since demonstratives originate preceding adjectives:

(30)   *Adjective LBE cannot cross co-occurring demonstrative*
       * **Bol'šogo**$_2$ ja videl [ètogo t_2 kota].
         big   I saw this   cat

       'I saw this big cat.'

However, if the crossed-over left branch also extracts, the result is acceptable, as (31) shows. Importantly, the final order of the two left branches must match the relative order that they had in their original positions. Thus as we see in (31), order-preserving LBE of both a demonstrative and adjective is acceptable (31a), while order-reversing LBE of these elements is not (31b):

(31)   *Deep LBE repaired by additional order-restoring LBE*
       a.   ✓ Ètogo$_1$ **bol'šogo**$_2$ ja videl [t_1 t_2 kota].
              this   big   I saw   cat

            'I saw this big cat.'
       b.   * **Bol'šogo**$_2$ ètogo$_1$ ja videl [t_1 t_2 kota].
              big   this   I saw   cat

            'I saw this big cat.'

As we will see, patterns like this are pervasive for Russian LBE. I argue that this result is correctly predicted by hypothesizing that the Russian nominal phrase is a phase, and thus under CL, a domain to which order preservation applies. Familiar principles of locality will also be central to understanding this generalization.

In the second part of this chapter, I examine an interaction between scrambling and LBE from subjects in Russian, building from Bondarenko and Davis (To appear). While LBE from Russian subjects is generally possible, such extraction usually fails when combined with scrambling of other material that originates in the same vP in the way shown in (32). Here a scrambled object intervenes between the subject and an element extracted from it:

(32)   *No object scrambling + LBE from subject*
       * **Každaja**$_2$ kota$_1$ [t_2 **devočka**] pogladila t_1.
         every.NOM cat.ACC   **girl.NOM** stroked

       'Every girl stroked a cat.'

Importantly, this combination of movement operations is only unacceptable if the scrambled phrase lands in a position in between the subject and the sub-extracted element. Thus if the scrambled object in (32) moves further, to a position above the extracted element, then the resulting configuration is acceptable (33):

(33)   *Further scrambling ameliorates LBE from subject*
         <u>Kota</u>$_2$ **každaja**$_1$ vćera   većerom [t_1 **devoćka**] pogladila t_2.
         cat.ACC every.NOM yesterday evening   girl.NOM   pet
         'Every girl stroked the cat.'

Bondarenko and Davis (To appear), and thus also this chapter, argue that this pattern
emerges from two concepts which Ko (2007, 2011, 2014) uses to account for similar
effects in Korean and Japanese: the phasehood of vP, and the ban on phrase-bounded
specifier-to-specifier movement. We will see that while this account makes accurate
predictions for sub-extraction in a wide variety of contexts in Russian, sentences
like (32) are sometimes acceptable when involving LBE of certain adjectives and
possessors. I suggest following Bondarenko and Davis (To appear) that this effect
stems from the external merge possibilities available for such elements. We will
also see that VS word order can improve configurations like (32), which I follow
Bondarenko and Davis (To appear) in hypothesizing is due to a mechanism of phase
extension (Den Dikken 2007; Gallego 2010; Alexiadou et al. 2014).

5.4   Chapter 5: Overlapping paths and parasitic gaps

Chapter 5 shifts focus away from sub-extraction, turning towards an examination of
configurations where multiple instances of (overt) A′-movement overlap. I argue that
the concepts defended throughout this dissertation shed light on several properties of
these configurations, as well as intertwined topics such as the formation of multiple
specifier structures, the pronunciation of movement chains, and the licensing of
parasitic gaps. The way that parasitic gaps interact with the semantics of movement
is particularly central to this chapter.

In English, the simplest way to form a configuration with overt overlapping A′-
chains is to combine topic/focus movement with *wh*-movement, as in (34), though
in this chapter we will see a variety of other contexts where this occurs:

(34)   *Topic/focus movement plus wh-movement in English*
         [This book]$_2$, who$_1$ should we talk to t_1 about t_2?

If both vP and CP are phases, then even mono-clausal sentences like (34) must
involve some successive-cyclic movement. Specifically, here the two internal argu-
ments undergoing A′-movement must have formed intermediate positions in the vP
edge. Since in this situation there are two moving phrases, there are two structures
for the vP that are conceivable for this derivation, illustrated in the simplified dia-
grams in (35) below: either the topic/focus phrase could form an outer spec-vP, with
the *wh*-phrase forming an inner one (35a), or the reverse might occur (35b).

(35)   *Two conceivable intermediate orders in the vP edge*

    a.   [$_{CP}$ *this book$_2$ who$_1$ ...* S [$_{vP}$ ***this book$_2$ who$_1$*** *v-V ... who$_1$ this book$_2$*]]

    b.   [$_{CP}$ *this book$_2$ who$_1$ ...* S [$_{vP}$ ***who$_1$ this book$_2$*** *v-V ... who$_1$ this book$_2$*]]

Notice that CL predicts that only the derivation in (35a) should be permitted, since only this derivation involves the two moving phrases having the same relative ordering both at the spell-out of vP, and in their final landing sites. I argue that this prediction is correct, and in doing so defend the following generalization:

(36)   *Overlapping Chain Generalization*
If two A′-moving phrases XP_1 and XP_2 form specifiers of an intermediate phase YP such that XP_1 c-commands XP_2, then XP_1 c-commands XP_2 in all subsequent positions these phrases occupy.

That (35a) accurately represents the intermediate order in this derivation is indicated by an asymmetry in parasitic gap licensing. While either of the varieties of A′-movement in (34) can license a parasitic gap when in isolation, when combined in one sentence, only the moved phrase with the higher position on the surface can serve as the licenser, as (37) exemplifies:

(37)   *A parasitic gap is licensed only by the higher of two overlapping movements*
[This book]$_2$, who$_1$ should we talk to t_1 about t_2...

    a.   ✓ ...before commenting on PG$_2$?

    b.   * ...before arranging a meeting with PG$_1$?

We will see that this finding is a very general one, applicable to all configurations with overlapping overt A′-chains. This asymmetry indicates that the higher phrase on the surface in such sentences was also the higher phrase within the vP, as CL predicts, given the following prediction from Nissenbaum (2000) that this chapter will justify in detail:

(38)   *Multiple Specifier Single Parasitic Gap Prediction*
In multiple specifier constructions in which XP_1 is the highest specifier of vP and XP_2 is a low specifier of vP, only XP_1 can license a single parasitic gap in an adjunct to that vP.[17]

While this result verifies the predictions of CL, it also raises a puzzle for the widely-adopted proposal of Richards (1997, a.o.) that non-initial movements to a given head "tuck-in" to a lower specifier. Under such a theory, we would expect

---

[17]This statement of the prediction is taken from Fox and Nissenbaum (2018), who use it to analyze certain constructions involving pied-piping that are beyond the scope of this chapter.

a derivation like (35b) to be required for the sentence in (34), rather than the derivation in (35a) which I argue in favor of. In response to this finding, I argue that tucking-in is not characteristic of multiple movement to one head, but rather of multiple movements driven by the same single feature (39). Conversely, I propose that multiple movements driven by separate features can form multiple specifier configurations either with or without tucking-in (40).[18]

(39)   *One feature attracts multiple goals → tucking-in*
$$[_{XP}\ \alpha_{[F1]}\ \beta_{[F1]}\ X_{[uF1]}\ ...\ t_\alpha\ t_\beta\ ]$$

(40)   *Attraction by separate features on one head → variability*

    a. $[_{XP}\ \alpha_{[F1]}\ \beta_{[F2]}\ X_{[uF1,uF2]}\ ...\ t_\alpha\ t_\beta\ ]$

    b. $[_{XP}\ \beta_{[F2]}\ \alpha_{[F1]}\ X_{[uF1,uF2]}\ ...\ t_\alpha\ t_\beta\ ]$
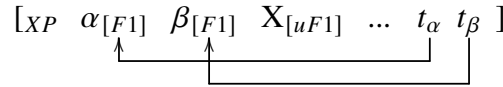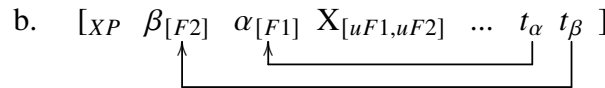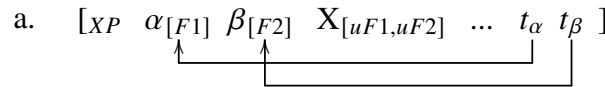
    The latter of these concepts is important for the analysis of a related phenomenon, the *path containment condition* (Pesetsky 1982, a.o.). This condition describes the fact that in various languages, English among them, overlapping A′-movements must involve nested dependencies, meaning that two overtly A′-moving phrases must end up in a final order that is the reverse of the order they had originally:

(41)   *Forced order reversal for overlapping movement paths*

    a.   [This book]$_2$, who$_1$ should we talk to t_1 about t_2?   [2 1 ... 1 2]
    b.   * [This student]$_1$, what$_2$ should we talk to t_1 about t_2?   [1 2 ... 1 2]

Richards (1997) argues that the path containment condition arises in languages that allow only one overt specifier of CP. Extending this proposal, I argue that the path containment condition emerges from the CL theory when combined with the hypothesis that in languages like English, lower specifiers of A′-movement are sometimes covert (Richards 1997; Nissenbaum 2000, a.o.). I will argue that this covertness results in a linearization problem in certain configurations, which is avoided by forming an alternative order of specifiers that culminates in a reversed final order. I will show that this understanding makes correct predictions about the path containment condition in English, its absence in the multiple *wh*-fronting languages Bulgarian and Romanian, and fits into a general theory of what governs

---

[18]The concept that the presence of multiple movement triggering features on one head removes the need to tuck-in has precedent in several preceding works (McGinnis 1998; Rackowski 2002; Doggett 2004; Rackowski and Richards 2005).

the distribution of crossing versus nesting movement paths.

5.5   Chapter 6: Ordering extraposition

While the preceding chapters all focus on leftward movement processes, chapter 6 shifts focus to extraposition, a rightward displacement phenomenon whose properties are interestingly different from the more familiar leftward sort. Here I specifically focus on extraposition of non-arguments, a classic instance of which is relative clause extraposition. This is illustrated in (42) below for English and Wolof (Niger-Congo), two languages which provide data that is essential to this chapter.

(42)   *Relative clause extraposition*

   a.   *English*
        I bought [a book t_1] yesterday [**that someone recommended to me**]$_1$.

   b.   *Wolof*
        Gis-naa [ab fas t_2] démb   [**wu nga sopp**]$_2$.
        see-1sg a   horse   yesterday that you like

        'I saw a horse yesterday that you like'

   The primary goal of this chapter is to justify and account for the generalization about extraposition in (43) below, which has been hypothesized, in different ways, by at least Jenks (2011, 2013a,b), Nissenbaum (2000), and Fox and Pesetsky (2009):

(43)   *Source ordering generalization (SOG)*
       Extraposition of a non-argument constituent α from a source phrase XP must result in α having the same relative linear order with respect to the rest of the content of XP that it would have had in the absence of extraposition.

As far as I am aware, no recent literature has gathered all of the currently known relevant facts about the relationship between extraposition and word order, and considered a linearization account of them in detail. The goal of this chapter is to fill this gap. Here I report the relevant data from the works cited above and from my own recent fieldwork on Wolof and Javanese in order to confirm the cross-linguistic nature of the generalization in (43), and then explain it.[19] [20]

---

[19]All Wolof data in this chapter, unless otherwise cited, comes from my fieldwork with speakers Lamine Diallo, Lamine Toure, and Aisha Toure at MIT in 2018, while all Javanese data reported here comes from my fieldwork with native speaker Ismartilah Drummond at MIT in 2019.

[20]This chapter owes a considerable debt to lecture notes from a syntax seminar taught by Danny Fox and David Pesetsky in 2003, and to numerous discussions with Danny Fox in particular, who deserves credit for the majority of the concepts that facilitate this chapter's findings. The proposals

The SOG describes the fact that only elements that would independently surface at the right edge of the source phrase in a given context are capable of extraposition—a generalization likely first observed by Nissenbaum (2000) on the basis of contrasts between English and Hebrew. The SOG is illustrated schematically in (44):

(44)   *Predictions of the SOG*

    a.   *Extraposition from right linear edge of source phrase*
       ✓ [$_{XP}$ X (YP) ... α] ... α

    b.   *No extraposition from linear interior of source phrase*
       * [$_{XP}$ α X (YP) ...] ... α

The concept that linear crossing interrupts extraposition as the SOG describes is evocative of other constraints on displacement that come into focus throughout this dissertation. A telling instance of the SOG comes from Wolof. In this language, indefinite determiners are pre-nominal, as in (42b) above, while definite ones and demonstratives are post-nominal, as in (45a) below, following relative clauses. My fieldwork has found that relative clause extraposition in Wolof is possible only with such elements that are pre-nominal, as (42b) showed, and not with those that are post-nominal, as we see in (45b):

(45)   *Post-nominal determiner/demonstrative blocks extraposition in Wolof*

    a.   Gis-naa [fas   [**wu nga sopp**] wi/wee] démb.
       see-1sg horse that you like   the/that yesterday
       'I saw the/that horse yesterday that you like'

    b.   * Gis-naa [fas t_1 wi/wee] démb   [**wu nga sopp**]$_1$.
       see-1sg horse   the/that yesterday that you like
       'I saw the/that horse yesterday that you like'

However, in Wolof a focused demonstrative precedes NP instead of following it as usual (46a), in which case relative clause extraposition becomes acceptable (46b):

(46)   *Pre-nominal demonstrative allows extraposition*

    a.   Gis-naa [(yii) góór [**yu njool**] (yii)].
       see-1sg (these) men that tall   (these)
       'I saw these / THESE men who were tall'

---

in sections 4, 5, and 6 of this chapter are all directly inspired by those lecture notes and by Danny's thoughts, though my decisions about how exactly these general ideas must be implemented and justified differs in various details.

b. Gis-naa [yii góór t_1] démb [**yu njool**]$_1$.
see-1SG these men yesterday that tall

'I saw THESE men yesterday who were tall'

Under CL, typical phrasal movement does not have to be sensitive to the premovement word order of the domains it passes through, since that movement can be successive-cyclic. The fact that the SOG holds thus indicates that extraposition is not successive cyclic extraction, since this would circumvent the crossing problem shown in (44b) above. I argue that a theory of extraposition as late external merge to a (covertly) moved phrase (Fox and Nissenbaum 1999; Nissenbaum 2000; Fox 2002; Overfelt 2015, a.o.) is best suited to account for the SOG. The account of extraposition in the works just cited depends on the proposal that covert syntactic movement and material is subject to linear ordering. Here I will offer an account of the SOG that removes this assumption, resulting in desirable consequences for several puzzles, such as the possibility of extraposition from overtly moved phrases.

5.6 Chapter 7: A digression about parasitic gaps in relative clauses

In chapter 7, I set aside CL and continue to explore several concepts from chapter 5 by way of a phenomenon which has received little attention in previous work—parasitic gaps in relative clauses licensed by extraction from the containing DP:

(47) *PG in relative clause licensed by extraction from DP*

a. Who$_1$ did Mary take [pictures of t_1 [that weren't that flattering to PG$_1$]]?
(Citko 2014, ex. 105)

b. That's the manager which$_2$ I know [an employee of t_2 [who's got a very intense grudge against PG$_2$]].

c. Let me tell you who$_3$ I've noticed [an aspect of t_3 [that makes me want to avoid PG$_3$]].

d. This person$_4$, I painted [a beautiful portrait of t_4 [that unfortunately didn't satisfy PG$_4$]].

As chapter 5 discusses, following Nissenbaum (2000), a parasitic gap should be able to be licensed as long as the constituent that contains it adjoins to a position targeted for (successive cyclic) A′-movement. In the context of this understanding of parasitic gaps, Citko (2014) cites example (47a) above as evidence that extraction from DP passes through spec-DP, assuming that the relative clause can be merged in the projection of D. However, restrictive relative clauses like those in (47) are

typically analyzed as semantic predicates that adjoin to NP (not DP) and combine with it via Predicate Modification (Heim and Kratzer 1998, p. 88), as in (48):

(48) *Restrictive relative clauses are sisters of NP*
      $\{x \mid x \text{ is cat}\}_{[\![NP]\!]} \cap \{x \mid x \text{ has orange fur}\}_{[\![RC]\!]}$
      $= \{x \mid x \text{ is cat and } x \text{ has orange fur}\}$

I argue that given these considerations, the facts in (47) actually indicate that extraction from DP can form an intermediate landing site in the NP edge, as in (49):

(49) *Extraction from DP via NP edge*
      $[\ \ WH_4\ \ ...\ \ \text{V}\ \ [_{DP}\ \ \text{D}\ \ [_{NP}\ \ t_4\ \ [_{N'}\ \ \text{N}\ \ t_4\ \ ]]]]$

    While extraction via the DP edge as well is not precluded by this analysis, the possibility of movement via the NP edge must be posited to capture the facts in (47). We expect extraction from NP to pass through its edge if NP is a phase. This concept has precedent in a variety of works on morpho-phonology: several studies in this domain argue that lexical categories like NP in fact consist of a category neutral root as well as a phase-demarcating categorizing head, in this case *n* (Marvin 2003; Embick and Marantz 2008; Embick 2010, a.o.) If NP is in fact *n*P and *n*P is a phase for morpho-phonology, then it should also be a phase for syntax. If this is so, movement from *n*P/NP should be required to pass through its edge. Such movement, if possible, should provide a position where a relative clause containing a parasitic gap can be merged. The sentences in (47) above show us that this is indeed the case.
    I will discuss several sources of evidence for this analysis of the sentences in (47), and explore a number of additional facts about this configuration. For instance, I will argue that certain interpretive constraints on the gaps in these sentences provide further evidence for a proposal from Nissenbaum (2000) that overlapping movement of multiple operators requires crossing paths:

(50) *Overlapping operator movements must form crossing paths*
      a.   ✓ OP$_1$ OP$_2$ ... t_1 t_2

      b.   * OP$_2$ OP$_1$ ... t_1 t_2

I will also argue that the possibility of movement through the NP edge challenges the claim in Bošković (2005) and related work that an interaction between the presence of D, and anti-locality, is responsible for blocking LBE in languages like English.

5.7 Chapter 8: Final remarks

In this final chapter, I summarize the main results of this dissertation, and discuss two more general topics that relate to the CL theory defended here. Both of these topics present a number of potential angles for future work.

First, I show that several phenomena that I attribute to the nature of linearization in this dissertation persist even when the relevant elements are covert. I hypothesize that a grammar in which linearization precedes the insertion of morpho-phonological content (Embick 2010; Arregi and Nevins 2012; Ostrove 2018), or in which preservation is evaluated in terms of structural relations (Emonds 1970, 1976; Williams 1999, 2003; Müller 2002) perhaps in addition to linear ordering, will successfully achieve most if not all of the results I argue for here. Second, I make explicit the predictions for CL about agreement, which under this theory should be in principle unbounded. As I discuss, there is precedent for a theory in which agreement is not bounded by spell-out, but by other locality concerns related to the nature of phases (Rackowski and Richards 2005; van Urk and Richards 2015; Halpert 2016, 2019; Branan 2018). In the context of CL, we expect an approach to the constraints of agreement that does not make reference to spell-out to be an accurate one.

In general, the results of this dissertation constitute evidence for a theory under which derivations proceed in discrete cycles or "chunks", and for which the information established at the end of each cycle must be respected for the remainder of the derivation. Provided that the arguments for such a theory are successful, there is much more that can be asked about the precise characteristics of the relevant cycles, and the nature of the information preservation requirement. I hope to show, however, that a cyclic grammar sensitive to the preservation of at least linear order information has a number of advantages.

Chapter 2

Crossing and stranding at edges

1 Introduction

In this chapter, I initiate this dissertation's exploration of displacement and discontinuity through an investigation of certain patterns involving pied-piping and stranding. As we will see shortly, numerous works on a variety of languages have identified patterns in which moved phrases can sometimes leave material behind at an intermediate point in the path of movement. If the syntactic derivation is bounded by phases, which movement from must successive-cyclically pass through the edge of, then the possibility of such stranding is expected. This is because the intermediate landing sites forced by phases provide positions where, in principle, we predict that stranding might occur. In this chapter, I examine the cross-linguistic distribution of such intermediate stranding, which I argue is constrained in such a way that reveals further evidence for the CL theory, along with several independently proposed limitations on movement in syntax.

The schema in (1) below illustrates the basic form of an intermediate stranding derivation. Here successive-cyclic movement of a phrase α intermediately strands an element β in the edge of the phase YP, in the following way. First, α moves to the YP edge while simultaneously pied-piping β. Second, α moves on alone, leaving β behind in the edge of YP. Thus movement of α through the edge of the phase YP feeds intermediate stranding of β.

(1)   *Intermediate stranding under successive-cyclic movement*
        $[_{ZP}\ \alpha\ \text{Z}\ [_{YP[Phase]}\ [\ t_\alpha\ \beta\ ]\ \text{Y}\ [_{XP}\ t_{\alpha\beta}\ \text{X}\ ]]]$

In this chapter, I show that such stranding obeys a cross-linguistic generalization about word order, stated below:

(2)   *Intermediate Stranding Generalization (ISG)*
        Leftward movement of a phrase α can only intermediately strand an element β if β is (or can be) ordered to the right of α before stranding occurs.

I argue that this generalization stems from a particular proposal concerning the nature of spell-out, along with independent constraints on movement in syntax.

1.1 Preview of results

Assuming that intermediate stranding typically occurs in phase edges for the reasons outlined above, I argue that the ISG in (2) holds because only those derivations that obey it avoid a fatal crossing problem. This problem is illustrated in (3a) below. Here β precedes α before intermediate stranding occurs. In this situation, when α strands β in the edge of the phase YP, movement of α from YP must cross over β. By contrast, in (3b) α originally precedes β. Thus here movement of α will not cross over β when β is stranded at the YP edge. In this chapter, I argue for a theory which rules in only non-crossing intermediate stranding derivations like (3b). Since (3b) corresponds to what the ISG describes, the ISG will thus be derived.

(3)   a.   * *Intermediate stranding with crossing at the edge*
               $[_{ZP} \ \alpha \ Z \ [_{YP[Phase]} \ [\ \beta \ t_\alpha \ ] \ Y \ [_{XP} \ t_{\beta\alpha} \ X \ ]]]$

        b.   ✓ *Intermediate stranding without crossing at the edge*
               $[_{ZP} \ \alpha \ Z \ [_{YP[Phase]} \ [\ t_\alpha \ \beta \ ] \ Y \ [_{XP} \ t_{\alpha\beta} \ X \ ]]]$

In particular, I argue that the ISG arises naturally from CL and its interaction with the ban on movement within phrase edges introduced in the previous chapter (Ko 2007, 2011, 2014, a.o.), as well as with anti-locality (Bošković 1997; Ishii 1999, Grohmann 2003; Abels 2003, 2012; Erlewine 2016, a.o.). These concepts together constrain intermediate stranding in precisely the way the ISG describes.

1.2 Chapter contents

Next, section 2 provides the empirical basis for the ISG, describing all intermediate stranding patterns that I am currently aware of. Section 3 overviews phase theory and the CL approach, and shows how CL derives the ISG. Section 4 demonstrates how independent constraints on movement interact with CL to accurately rule out various alternative derivations that would violate the ISG. Section 5 goes on to consider some factors that constrain the set of possible landing sites for stranded

material. Section 6 compares this chapter's approach to that of Bošković (2018), who predicts a result that is similar to the ISG, but different in several details.

2 Intermediate stranding cross-linguistically

This section provides the empirical motivation for the ISG, reporting all patterns I currently know of that may constitute intermediate stranding. Most of these patterns come from previous literature. While some of these scenarios are more straightforward than others, it will nevertheless be evident that a clear word order generalization can be stated about patterns of this nature.

2.1 Stranding in West Ulster English and a preposition puzzle

McCloskey (2000) discusses what is likely the most well-known case of intermediate stranding, involving the post-nominal quantifier *all* in West Ulster English, a dialect of Ireland. In this English variety A′-movement can pied-pipe *all*, strand it in its base position, or as (4) shows, strand it in a clause edge:

(4) *West Ulster English all-stranding in clause edge*
   a. Where$_2$ do you think [$_{CP}$ [t_2 **all**]$_1$ they'll want to visit t_1]?
   b. Who$_2$ did Frank tell you [$_{CP}$ [t_2 **all**]$_1$ that they were after t_1]?
   c. What$_2$ do they claim [$_{CP}$ [t_2 **all**]$_1$ (that) we did t_1]?
      (McCloskey 2000, ex. 9)

Based on such facts, McCloskey argues that A′-movement from CP pauses in the CP edge, and that the punctuated nature of such movement provides an intermediate landing site where this dialect's strandable *all* can be left behind. A clear piece of evidence that such patterns truly involve stranding is the fact that a stranded *all* can only appear in positions within the A′-movement path of the associated phrase:

(5) *All-stranding must occur within the movement path*
   a. What$_1$ did she buy t_1 **all** in Derry yesterday?
   b. * What$_1$ did she buy t_1 in Derry **all** yesterday?
   c. * What$_1$ did she buy t_1 in Derry yesterday **all**?
      (McCloskey 2000, ex. 19)

Decades earlier, Postal (1972, 1974) argued for the opposite of McCloskey's conclusion, based in part on the fact that English prepositions cannot be stranded in clause edges (6). Postal argues that such intermediate P-stranding would be

possible if movement from an embedded clause were truly successive-cyclic, given that English prepositions are in principle strandable:

(6)   *No intermediate stranding of English prepositions*

    a.   (To) [which writer]$_1$ do you think [$_{CP}$ **(*to)** t_1 (that) we should send the pen (to) t_1]?

    b.   (For) [which dog]$_1$ did they claim [$_{CP}$ **(*for)** t_1 (that) I cooked a steak (for) t_1]?

    c.   This is the person [(in) [whose pants]$_1$ you said [$_{CP}$ **(*in)** t_1 (that) I put eels (in) t_1]].

    d.   (With) [this poison]$_1$, I think [$_{CP}$ **(*with)** t_1 (that) we should kill the pterodactyls (with) t_1].

The same constraint is found in Norwegian, which like English allows a preposition to be either stranded in its base position or totally pied-piped, but not stranded in a CP edge, as (7) shows:

(7)   *Norwegian prepositions cannot be intermediately stranded*

    a.   (I) [hvilket rom]$_1$ trodde du [**(*i)** t_1 jeg satt (i) t_1]?
        (in) which room thought you (in) I sat (in)
        'In which room did you think I sat?'

    b.   (På) [hvilken bord]$_1$ trodde du [**(*på)** t_1 trollmannen sagde
        (on) which table thought you (on) the.wizard sawed
        kvinnen i to (på) t_1?
        the.woman in two (on)
        'On which table did you think the wizard sawed the woman in two?'

    c.   (Om) [hvilken bok]$_1$ trodde du [**(*om)** t_1 jeg snakka (om) t_1]?
        (about) which book think you (about) I spoke (about)
        'About which book did you think I spoke?'

    d.   (Om) hvem$_1$ trodde du jeg sa [**(*om)** t_1 han spurte (om) t_1]?
        (about) whom thought you I said (about) he asked (about)
        'About whom did you think I said he asked?'
        (Henrik Torgersen, p.c.)

Yet another example of this constraint is found in Brazilian Portuguese, which allows stranding of one preposition, *sobre* ("about"), but not in intermediate positions:

(8)   *No intermediate P stranding in Brazilian Portuguese*

    a.   (Sobre) o que$_1$ você acha (***sobre**) t_1 que eu li   (sobre) t_1?
        about   what   you   think about   that I   read about
        'What do you think that I read about?'

    b.   (Sobre) [esse assunto]$_1$, eu acho (***sobre**) t_1 que voce leu   (sobre) t_1.
        about   this   subject,   I   think about   that you   read about
        'This subject, I think you read about.'

    c.   o   assunto [$\varnothing_1$ que eu acho (***sobre**) t_1 que você leu   (sobre) t_1]
        the subject   that I   think about   that you   read about
        'the subject which I think you read about'

        (Filipe Hisao Kobayashi, p.c.)

The theories that the facts in (4) and (6-8) respectively suggest are in conflict: If A′-movement from CP is not successive cyclic, then what allows intermediate *all*-stranding in West Ulster English? If such movement really is successive-cyclic, then why is intermediate preposition stranding banned in English, Norwegian, and Brazilian Portuguese? I argue that word order is the key to this puzzle. Notice that prepositions, which can't be intermediately stranded, precede the phrase they merge to (9a) prior to stranding. The West Ulster English strandable *all*, by contrast, follows the associated phrase (9b), and can be intermediately stranded. This contrast parallels the schema in (3) above, and thus fits the ISG introduced in (2).

(9)   a.   *Intermediate stranding impossible*
        for which cat

    b.   *Intermediate stranding possible*
        what **all**

In this chapter, I maintain that A′-movement is indeed generally successive-cyclic when it exits a phase (CP, vP, etc.), and argue that the ban on intermediate stranding of prepositions, as well as the ISG more broadly, have a linearization explanation. In the remainder of this section, I discuss all of the additional intermediate stranding patterns that I am aware of, all of which we will see fit the ISG.[1]

---

[1]In the next chapter, which focuses on a form of possessor extraction in colloquial English, we will encounter another instance of stranding that obeys the ISG. In particular, this phenomenon sometimes requires stranding of the possessum ['s NP] in an intermediate position when the possessor extracts from it, a pattern that fits the ISG given that the English possessor precedes the possessum.

2.2 Stranding of adjuncts to *wh*-phrases in English

English allows an interrogative *wh*-phrase to be modified by *exactly / precisely* and similar elements. Zyman (To appear) argues that such elements should be considered adjuncts, and as is often (though not always) the case for English adjuncts, these elements can attach on either side of the phrase they merge to:

(10)   *Variable order of wh-adjoined exactly/precisely*

      a.   (**Exactly/precisely**) [how many cakes]$_1$ (**exactly/precisely**) did you say that we ate t_1?

      b.   Who said that you ate (**exactly/precisely**) [how many cakes] (**exactly/precisely**)?

Such adjuncts can be stranded in their base position or in a CP edge by *wh*-movement of their host phrase (Urban 1999; Stroik 2009; Zyman To appear):

(11)   *Adjunct stranding under wh-movement*
        What$_1$ did you suppose t_1 (**exactly/precisely**) (that) they wanted t_1 (**exactly/precisely**)?

Since these adjuncts are able to be adjoined to the right side of their host phrase, the possibility of stranding them in a CP edge fits the ISG. The same basic pattern holds for adjuncts of quantity like *to the nearest pound* in (12) below. Such adjuncts provide a clearer instance of intermediate stranding, since they are not potentially interpretable as adverbs of vP/VP, unlike *exactly / precisely*:

(12)   *Intermediate stranding of quantity adjunct*
        Tell me [$_{CP}$ (to the nearest pound) [**how much flour**]$_1$ (to the nearest pound) you said [$_{CP}$ t_1 **(to the nearest pound)** (that) the bakery wants t_1 (to the nearest pound)]].

Such intermediate adjunct stranding is unacceptable when the host *wh*-phrase does not move. We see this in (13b/d) below, where the relevant adjunct in the embedded clause periphery is construed with a *wh*-phrase that remains in situ, due to being the lower of two *wh*-phrases in a multiple *wh*-question. The unacceptability of these examples is expected, if this intermediate position of the adjunct can only be derived by stranding under movement of its host:

(13)   *Intermediate adjunct stranding is parasitic on movement*

      a.   [**How much flour**]$_1$ did you say t_1 (**to the nearest pound**) (that) they'll deliver t_1?

b. * Who said (**to the nearest pound**) that they will deliver [**how much flour**]?

c. [**How many donuts**]$_1$ did you say t_1 (**to the nearest dozen**) (that) the bakery will give away t_1?

d. * [Which bakery] reported (**to the nearest dozen**) that they will give away [**how many donuts**]?

Zyman (To appear) argues that evidence from vP fronting and sluicing shows that when an adjunct like *exactly* appears to have been stranded in its base position in VP, it has actually moved to a high rightward position. We might analyze this phenomenon as extraposition, which is independently attested for English adjuncts. The availability of such extraposition opens up the possibility of a non-stranding analysis for the intermediate stranding in spec-CP shown in this section: it could be the case that these examples involve extraposition of the adjunct from the *wh*-phrase and into the right periphery, followed by extraposition of the embedded CP to the right of that adjunct. If this hypothesis is right, these intermediate displaced adjuncts should require the embedded complementizer to be overt, since clauses with no overt complementizer resist extraposition (Stowell 1981). In actuality, such examples with no (obligatory) embedded complementizer are attested, as shown in (11-13) above. This observation strengthens the hypothesis that such facts are stranding-derived, and thus constitute relevant evidence for the ISG.

2.3 Q-stranding in Wolof

Torrence (2018) examines a variety of elements that appear in the periphery of clauses crossed by *wh*-movement in Wolof (Niger-Congo, Atlantic). Torrence reports that there are at least two morpho-syntactic classes of such elements. He argues that one of these classes, which he terms "Q-like", has a distribution indicative of stranding under movement. According to Torrence, these quantificational elements obligatorily follow the NP they merge with. This fact is evident when they are not stranded by movement, as in (14) below, where total pied-piping occurs:

(14) *Full pied-piping of Q-like element in Wolof*

a. [Ñ-an **ñ-epp**]$_1$ l-a Ayda wax ne l-a-a   dóór t_1?
   who   every   COP Ayda say that COP.1SG hit
   'Who all did Ayda say that I hit?'
   (Torrence 2018, ex. 38a)

    b.   [F-an **f-eeneen**]$_1$ l-a Ayda wax ne l-a-a dem t_1?
        Where other   COP Ayda say that COP.1SG go
        'Where else did Ayda say that I went'
        (Torrence 2018, ex. 38b)

These elements need not necessarily be pied-piped by *wh*-movement, but can also be stranded in situ (15), as well as in the edge of an embedded CP (16). Since these strandable elements originate to the right of NP, their ability to undergo intermediate stranding in the CP edge corresponds with the ISG.

(15)   *Q-stranding in base position in Wolof*

    a.   **Y-an**$_1$ la Binté waat ne nga lekk [t_1 **y-epp**]?
        what COP Binta swear that COP.2SG eat every
        'What all did Binta swear that you ate?'
        (Torrence 2018, ex. 44d)

    b.   **F-an**$_1$ l-a-ñu wax ne nga teg téére bi [t_1 **f-eneen**]?
        where COP.3PL say that COP.2SG put book the other
        'Where else did they say that you put the book?'
        (Torrence 2018, ex. 45b)

(16)   *Q-stranding in intermediate clause edge in Wolof*

    a.   **F-an**$_2$ l-a-ñu foog [t_2 **f-epp**]$_1$ ne la-a togg-e ceeb t_1?
        where COP.3PL think every that COP.1SG cook rice
        'Where all do they think that I cooked rice?'
        (Torrence 2018, ex. 29a)

    b.   **F-an**$_2$ l-a-ñu foog [t_2 **f-eeneen**]$_1$ ne la-a togg-e ceeb t_1?
        where COP.3PL think other that COP.1SG cook rice
        'Where else do they think that I cooked rice?'
        (Torrence 2018, ex. 29b)

    Torrence states that the stranding of these elements is parasitic on A′-movement, and thus they cannot appear in the edge of an intermediate CP that was not crossed by movement of a relevant noun phrase. This is as expected if the appearance of such elements in a clause-peripheral position, as in examples like (16) above, is indeed derived by intermediate stranding.

2.4   Stranding in spec-vP in Dutch

Barbiers (2002) shows that long-distance A′-movement from an embedded clause in Dutch can strand adpositions and various other elements in the matrix spec-vP,

as demonstrated in (17) below:

(17)   *Stranding in spec-vP in Dutch*

a.   Waar$_2$ had jij   dan [$_{vP}$ [t_2 **mee**]$_1$ gedacht dat je   de vis t_1 zou
where had you then   with thought that you the fish   would
moeten snijden]?
must   cut
'What had you thought to be forced to cut the fish with?'

b.   Waar$_2$ had jij   dan [$_{vP}$ [t_2 **voor bal**]$_1$ gedacht dat Ed t_1 zou
where had you then   for ball thought that Ed   would
kopen]?
buy
'What kind of ball had you thought that Ed would buy?'

c.   [Een boek]$_2$ had ik [$_{vP}$ [**maar** t_2]$_1$ gedacht dat Ed t_1 zou   kopen].
One book   had I   only   thought that Ed   would buy
'I had thought that Ed would buy only **one** book'
(Barbiers 2002, ex. 6)

Scrambled phrases in Dutch are generally "frozen" and hence behave like islands (see Corver (2017) and references therein), suggesting that these examples are not derived by first scrambling a phrase into the matrix spec-vP, and then sub-extracting from it. It is conceivable that the freezing effect is circumvented in these examples by first performing sub-extraction, and then scrambling the remnant of that extraction from the embedded clause and into the matrix spec-vP. However, Barbiers shows that long-distance A′-movement cannot normally be combined with long-distance scrambling in Dutch. Barbiers' conclusion is that the facts in (17) are best analyzed as stranding under successive-cyclic movement.

Most of the elements that Barbiers shows undergoing such intermediate stranding in spec-vP originate to the right of what strands them, as in (17a-b). The exception is (17c), involving a stranded *maar* ("only"), which Barbiers shows as originating to the left of the moving element that stranded it. This example thus appears to violate the ISG. However, in other work, Barbiers (1995) shows that such configurations where *maar* has a numeral as its focus associate allow *maar* to be either pre- or post-nominal. We see this in (18) below, where *maar* may appear either on the left or the right of the NP containing the numeral *twee* ("two") that it focuses:

(18)   *Dutch "maar" in pre- or post-nominal position*

[(**Maar**) twee jongens (**maar**)] weten het antwoord.
(only)   ᴛᴡᴏ boys   (only)   know the answer

'Only two boys know the answer.'
(Adapted from Barbiers 1995, ex. 31c-d)

Since *maar* can be ordered to the right of NP in these contexts, (17c) is not an exception to the ISG.

Further Dutch stranding facts also fit the ISG. In Dutch, inanimate pronouns within PPs take on a special form (termed *R-pronoun*) with which many prepositions are inverted to alternative postpositional forms. While prepositions in Dutch cannot be stranded by A′-movement, their postpositional forms used with R-pronouns can be (van Riemsdijk 1978), as exemplified by (19-20) below:

(19)   *No preposition stranding in Dutch*

    a.   Ik snij het brood **met een mes**.
        I   cut the bread with a   knife

        'I cut the bread with a knife.'

    b.   * [**Welk mes**]$_1$ snij je   het brood [**met** t_1]?
        which knife cut you the bread with

        'Which knife are you cutting the bread with?'
        (Coppe van Urk, p.c.)

(20)   *Postposition stranded by moved R-pronoun in Dutch*

    a.   Ik snij het brood **daar-mee**.
        I   cut the bread there-with

        'I am cutting the bread with that.'

    b.   **Waar**$_1$ snij je   het brood [t_1 **mee**]?
        where cut you the bread   with

        'What are you cutting the bread with?'
        (Coppe van Urk, p.c.)

Importantly, the P "with" in (19-20) above is realized as *met* when it is a preposition, and *mee* when it is a postposition. Example (17a) above showed that the postposition *mee* is capable of intermediate stranding. As expected given the ISG, its prepositional variant *met* cannot be intermediately stranded. Hence an example analogous to (17a) that uses *met* instead of *mee* such as (21) below is unacceptable:

(21) *No preposition stranding in spec-vP in Dutch*

    * [Welk mes]$_2$ had jij   dan [**met** t_2]$_1$ gedacht dat je   de vis t_1 zou
     which knife had you then with   thought that you the fish   would
    moeten snijden?
    must   cut

    'Which knife did you think then that you would have to cut the fish with?'
    (Coppe van Urk, p.c.)

The respective distribution of prepositions and postpositions in Dutch thus corresponds precisely to what the ISG leads us to predict.

2.5 Afrikaans postposition stranding

Du Plessis (1977) shows that, like the related Dutch, Afrikaans cannot strand prepositions under A′-movement:

(22) *No preposition stranding in Afrikaans*

    a.   [**Vir wat**]$_1$ werk ons nou eintlik t_1?
          For what work we now actually?
          'For what do we actually work?'

    b.   * **Waar**$_1$ werk ons nou eintlik **vir** t_1?
           What   work we now actually for?
          'For what do we actually work?'
          (du Plessis 1977: pp. 724)

Also like Dutch, Afrikaans has strandable postpositions that occur with R-pronouns. du Plessis shows that these elements can be stranded at clause edges:

(23) *Afrikaans postposition stranding*

    a.   Waar$_1$(voor) dink julle [$_{CP}$ t_1 (**voor**) werk ons t_1 (**voor**)]?
          where(for)   think you [(for)   work we   (for)]?
          'For what do you think that we work?'

    b.   Wat/waar$_1$ dink julle die bure   [$_{CP}$ t_1 (**oor**)   stry   ons t_1
          What   think you the neighbors [(about) argue we
          (**oor**)]?
          (about)]?
          'What do you think the neighbors think we are arguing about?'
          (Adapted from du Plessis 1977, ex. 5, 12-13)

The elements that can undergo such intermediate stranding in Afrikaans are, as postpositions, ordered to the right of what strands them by leftward movement. Hence these facts also fit the ISG.[2]

2.6 Quantifier float under *wh*-movement in German

Doliana (2019) analyzes instances of the German quantifier *alles* that appear discontinuous from their *wh*-moved associate, as exemplified in (24) below:

(24)   *Quantifier float under wh-movement*
         Wen$_1$ (alles) hast du t_1 (**alles**) angerufen?
         who.ACC (all)   have you   (all)   called?

         'Who all did you call?'
         (German, Doliana 2019, ex. 1/2)

Doliana argues that an *alles* displaced from its host phrase is only able to appear in positions that its host is independently able to occupy, and thus argues that a displaced *alles* must have been derived by stranding. For instance, Doliana shows that *alles* cannot appear in a position outside of the movement path of its associated *wh*-phrase, as in example (25):

(25)   *Floated quantifier cannot appear outside of movement path of its host*
         * Wem$_1$ hat   der Peter t_1 erzählt, [$_{CP}$ dass die Maria **alles** die
           Who.DAT have.3SG the Peter   told   that the Maria all   the
         Susi   geholfen hat]?
         Susi.DAT helped   have.3SG

         'Who all did Peter tell that Maria helped Susi?'
         (German, Doliana 2019, ex. 22)

Importantly, Doliana also shows that *wh*-movement from an embedded clause can strand *alles* within the matrix clause, as in example (26) below. He argues that due to the clause-bounded nature of scrambling in German such a displaced *alles* (or a remnant phrase containing it) could not have moved to such a position by itself, and that therefore facts like (26) should be analyzed as intermediate stranding in spec-vP under successive-cyclic *wh*-movement:

---

[2]Rackowski and Richards (2005) and Den Dikken (2009) both argue that this Afrikaans pattern actually involves stranding in a clause-medial position. What matters for the present chapter is not where exactly these elements end up, but only that their word order properties fit the ISG.

(26) *Quantifier float in matrix clause by long-distance movement*

Wem$_1$ hat der Peter [$_{vP}$ t$_1$ (**alles**) gemeint, [$_{CP}$ dass die Maria t$_1$
Who.DAT have.3SG the Peter (all) told that the Maria

(**alles**) geholfen hat]]?
(all) helped have.3SG

'Who all did Peter say that Mary helped?'
(German, Doliana 2019, ex. 24)

If this conclusion is right, then in German we find another instance of ISG-respecting stranding, since the strandable quantifier *alles* originates to the right of its host.

2.7 Stranding by left branch extraction in Polish

Wiland (2010) analyzes intermediate NP stranding under *wh*-movement in Polish. Generally, Polish *wh*-movement permits both pied-piping of the entire nominal phrase containing a *wh*-element, as well as left branch extraction of the minimal *wh*-element, stranding NP below:

(27) *Pied-piping versus left branch extraction in Polish*

Jaki$_1$ (**samochód**) Paweł kupił swojej żonie t$_1$ (**samochód**)?
What car Pawel bought his wife car

'What car did Pawel buy his wife?'
(Wiland 2010, ex. 1-2)

Wiland shows that such left branch extraction can strand NP at various intermediate points in the sentence, as we see in (28) below. Assuming that Polish V moves to v, Wiland argues that the stranding positions shown in (28) are the specifiers of VP, vP, and CP. Wiland proposes that these three phrases are all phases.[3]

(28) *Intermediate stranding of NP under left branch extraction in Polish*

   a. *Intermediate stranding in spec-VP*

   Jaki$_2$ Paweł kupił [$_{VP}$ [t$_2$ **samochód**]$_1$ swojej żonie t$_1$]?
   What Pawel bought car his wife

   'What car did Pawel buy his wife?'
   (Wiland 2010, ex. 3)

---

[3]Wiland's proposal that VP is a phase is not shared by most work on phase theory, but this concept has some precedent, as discussed in section 5.3 below.

  b. *Intermediate stranding in spec-vP*

    Jaki$_2$ Paweł [$_{vP}$ [t_2 **samochód**]$_1$ kupił swojej żonie t_1]?
    What Pawel   car   bought his wife

    'What car did Pawel buy his wife?'
    (Wiland 2010, ex. 4)

  c. *Intermediate stranding in spec-CP*

    ? **Jaki**$_2$ *pro* myślisz [$_{CP}$ [t_2 **samochód**]$_1$ (*że) Paweł kupił
     What (you) think   car   (*that) Pawel bought

    swojej żonie t_1]?
    his wife

    'What car do you think that Pawel bought his wife?'
    (Wiland 2010, ex. 5)

Wiland notes that prima facie, it is possible to analyze these examples as scrambling followed by sub-extraction of the *wh*-element from the scrambled phrase. However, Wiland goes on to show that unlike *wh*-movement, Polish scrambling is clause-bounded. From this, he argues that (28c)[4] above and (29) below (at least) must truly involve stranding by *wh*-movement, and not a scrambling derivation, since in these examples NP is stranded outside of the region in which scrambling is permitted.

(29) *Long-distance wh-movement with stranding in matrix spec-vP in Polish*[5]

   % **Jaki**$_2$ Maria [$_{vP}$ [t_2 **samochód**]$_1$ myślała [$_{CP}$ że Paweł kupił swojej
    What Maria   car   thought that Pawel bought his

   żonie t_1]]?
   wife

   'What car did Maria think Pawel bought his wife?'
   (Wiland 2010, ex. 6)

Since the Polish NP is ordered to the right of the *wh*-element undergoing sub-extraction in the above examples, these stranding configurations fit the ISG.

---

[4]In particular, Wiland shows that clause-internal scrambling to the clause periphery in Polish lands in a position below C. Wiland argues that the stranded NP in (28c) sits in spec-CP (with C obligatorily null due to the Doubly Filled Comp Filter) since placing the complementizer left of the stranded NP in this example is unacceptable, though this word order should be permitted if (28c) were derived by scrambling. Thus Wiland argues that this example must involve not scrambling, but stranding in the clause's periphery under successive-cyclic movement.

[5]As the marking "%" on example (29) encodes, such a configuration is not acceptable for all speakers. Wiland does not offer an explanation for this fact.

2.8 Strandable ambivalent adpositions in Russian

Podobryaev (2009) compares two types of adpositions in Russian. First Podobryaev shows that it is not possible to strand typical prepositions in Russian, as is the case in many languages:

(30)   *No preposition stranding in Russian*

         a.   [**O**   **čem**]$_1$ ty   govoriš t_1?
              About what   you talk?
              'About what are you talking?'

         b.   * **Čem**$_1$ ty   govoriš **o** t_1?
               What you talk   about?
              'About what are you talking?'
              (Podobryaev 2009, ex. 1)

However, Russian also has what Podobryaev terms "ambivalent Ps", which can follow or precede their complement NP:

(31)   *Variable word order of Russian ambivalent Ps*

         a.   **navstreču** Pete
              towards   Petya

         b.   Pete **navstreču**
              Petya towards

         c.   **nazlo**   tebe
              to.spite you

         d.   tebe **nazlo**
              you to.spite
              (Podobryaev 2009, ex. 15-16)

Podobryaev shows that these ambivalent Ps may be stranded:

(32)   *Pied-piping and stranding of Russian ambivalent Ps*

         a.   (Navstreču) **komu**$_1$ (navstreču) ty   bežal t_1 (**navstreču**)?
              (Towards)   whom (towards)   you ran   (towards)?
              'Towards whom did you run?'

         b.   (Nazlo)   **komu**$_1$ (nazlo)   ty   èto sdelal t_1 (**nazlo**)?
              (To.spite) who   (to.spite) you this did   (to.spite)?
              'To spite whom have you done it?'
              (Podobryaev 2009, ex. 18-19)

Intermediate stranding of these adpositions in a clause edge is also possible (33):[6]

(33)   *Intermediate stranding of Russian ambivalent Ps*

    a.   ? Komu$_2$ Vasja xotel [t_2 **navstreču**]$_1$ čtoby Petja pobežal t_1?
          Who Vasya want   towards   that Petja run
          'Toward whom did Vasya want that Petja would run?'

    b.   ? Komu$_2$ Lena xotela [t_2 **nazlo**]$_1$ čtoby Maša   pobedila t_1?
          Who   Lena wanted   to.spite that Masha win
          'In spite of whom did Lena want that Masha would win?'

    c.   Čego$_2$ ty   dumaeš' [t_2 **radi**]$_1$ on prišël t_1?
          what you think   for   he came
          'For what do you think that he came?'
          (Tanya Bondarenko, Anton Kukhto, Mitya Privoznov, p.c.)

Since these ambivalent adpositions can be ordered to the right of what strands them in an intermediate position, such stranding fits the ISG.[7]

It is also possible for the adposition to end up in this same inter-clausal position even when there is no *wh*-movement to strand it there. Such examples require the moved adposition to receive a focused interpretation, however, as in (34) below.[8]

---

[6]This observation was made by Tanya Bondarenko and Mitya Privoznov, who confirmed that such sentences are possible, though subject to inter-speaker variation. The examples in (33a-b) use a subjunctive embedded clause because these are easier to extract from in Russian (Bailyn 2012). Speakers who permit extraction from finite clauses with an overt complementizer (*čto*) allow similar examples with movement from a finite clause.

[7]Sabine Iatridou (p.c.) reports a pattern in Greek with a similar character to this Russian one. In certain contexts like (i) it is possible for a preposition (here "above") to either precede or follow its complement in Greek, and in this case leaving the preposition behind in an intermediate CP edge under *wh*-movement is marginally acceptable (ii):

(i)   Evala   to vivlio [(**pano**) sto trapezi (**pano**)].
    put.1SG the book above   the table   above
    'I put the book on the table'

(ii)   ? Se pio   trapezi nomizis **epano** oti evale   to vivlio?
      to which table   think.2SG above that put.3SG the book
      'On which table do you think he/she put the book?'

[8]Anton Kukhto and Mitya Privoznov report that the examples in (34) are most acceptable if the nominals left behind by the adposition's movement precede the verb, as in (i):

(i)   a.   Vasja xotel **navstreču**$_1$ čtoby Petja **Maše**$_2$ pobežal t_2 t_1.
        Vasya want towards   that Petja Masha ran
        'Vasya wanted that Petja would run TOWARDS Masha (not any other direction).'

(34) *Non-pied-piping adposition movement in Russian*

    a. Vasja xotel **navstreču**$_1$ čtoby Petja pobežal Maše t_1.

        Vasya want towards   that Petja ran   Masha

        'Vasya wanted that Petja would run TOWARDS Masha (not any other direction).'

    b. Lena xotela **nazlo**$_1$ čtoby Maša pobedila Naste t_1.

        Lena wanted to.spite that Masha win   to Nastya

        'Lena wanted that Masha would win IN SPITE OF Nastya (not for her benefit).'

        (Tanya Bondarenko, Anton Kukhto, p.c.)

I argue that (33) is true intermediate stranding under A′-movement, whereas (34) involves scrambling of a remnant PP that has been evacuated by NP. While such PP scrambling evidently has a concomitant effect on interpretation, the fact that this semantic effect is absent from examples like (33) suggests that (33) does not involve scrambling, but rather pied-piping of PP along with an independent movement chain.

2.9 Interim summary

In this section I have reported all instances of intermediate stranding that my research so far has uncovered. All of these facts fit the ISG, repeated below:

(35) *Intermediate Stranding Generalization (ISG)*
Leftward movement of a phrase α can only intermediately strand an element β if β is (or can be) ordered to the right of α before stranding occurs.

The remainder of this chapter focuses on demonstrating how CL, plus certain independently proposed constraints on movement, straightforwardly derives the ISG.[9]

---

    b. Lena xotela **nazlo**$_1$ čtoby Maša **Naste**$_2$ pobedila t_2 t_1.

        Lena wanted to.spite that Masha to Nastya win

        'Lena wanted that Masha would win IN SPITE OF Nastya (not for her benefit).'

While I do not have an explanation for this preference, this fact does not pose a problem for any of the concepts under consideration here.

    [9]Ko (2011) argues that in Korean, object scrambling with the subject remaining in situ allows stranding of a numeral quantifier in spec-vP, as (i) shows. This pattern potentially instantiates another instance of intermediate stranding:

(i) **Kong-ul**$_2$ amato   [$_{vP}$ [t_2 **sey-kay**]$_1$ haksayng-tul-i t_1 patassulkesita].

    Ball-ACC probably   3-thing student-PL-NOM received

    'The students probably received three balls'

    (Ko 2011, ex. 24)

3 Two phase theories and their predictions

This section compares the predictions about intermediate stranding made by the phase theory in Chomsky (2000, 2001, a.o.) against those of CL. I will argue that the latter theory is better equipped to account for the word order generalization about intermediate stranding illustrated in the previous section.

3.1 Phases in Chomsky (2000, 2001, a.o.)

To make explicit what it predicts for the distribution of intermediate stranding, let's briefly review the approach to phases in Chomsky's work. Chomsky argues that syntactic derivations are mapped to phonology (PF) and interpretation (LF) incrementally, phase by phase. Minimally, vP and CP are phases. Chomsky hypothesizes that when the operation *spell-out* performs this mapping, the content of the spelled-out constituent becomes inaccessible to the rest of the syntactic derivation. Chomsky argues that spell-out applies to only the complement of phase heads. Consequently, this theory predicts that moving from a phase directly from its complement isn't possible, since the material in the phase's complement will undergo spell-out before such movement can apply (36a). However, movement to the edge (specifier) of the phase before its complement spells-out is predicted to allow further movement from the phase (36b).

(36)   *Must exit phase complement via the phase edge*

  a. * $[_{ZP}$ α Z $[_{YP[Phase]}$ Y $[_{XP}$ t $]]]$

  b. ✓ $[_{ZP}$ α Z $[_{YP[Phase]}$ t Y $[_{XP}$ t $]]]$
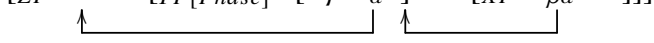
In this way, Chomsky's theory of phases predicts that moving phrases must pass successive-cyclically through the specifier of each phase crossed, in order to avoid being prematurely trapped by spell-out.
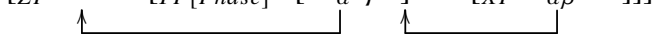
Similar strings are possible in Japanese (Shigeru Miyagawa, Takashi Morita, p.c.), whose syntax is highly similar to that of Korean. Ko shows that the numeral quantifier constructions in Korean/Japanese that allow stranding involve the numeral quantifier being ordered to the right of NP. Thus if examples like (i) are indeed instances of intermediate stranding, they fit the ISG. A reviewer for a previous version of this work points out that the great productivity of scrambling in Korean/Japanese makes it difficult to determine whether such examples actually constitute intermediate stranding rather than a more complex derivation involving something like remnant scrambling. Due to the possibility of such a confound, I have omitted this example from the main text of this chapter.

3.1.1 Predictions for intermediate stranding

Under the phase theory described above, anything which is in (or can reach) a complement-external position within a phase should, in principle, be available for further movement. Word order should not be at issue here, only structure. Therefore unless more is added to this theory, it predicts that both of the hypothetical intermediate stranding scenarios in (37) below should be licit.[10] These two scenarios are structurally comparable, but differ in word order. In reality, we've seen that in all attested intermediate stranding patterns the stranded material originally was (or could have been) ordered to the right of what stranded it. This fact is described by only the ISG-matching schema in (37b):

(37) *Two structurally comparable hypothetical forms of intermediate stranding*

    a.   *Intermediate stranding with crossing at the edge (ISG-violating)*
$$[_{ZP}\ \alpha\ Z\ [_{YP[Phase]}\ [\ \beta\ t_\alpha\ ]\ Y\ [_{XP}\ t_{\beta\alpha}\ X\ ]]]$$

    b.   *Intermediate stranding without crossing at the edge (ISG-obeying)*
$$[_{ZP}\ \alpha\ Z\ [_{YP[Phase]}\ [\ t_\alpha\ \beta\ ]\ Y\ [_{XP}\ t_{\alpha\beta}\ X\ ]]]$$

    The linear nature of the ISG is thus not captured by a theory of phases and movement cast purely in structural terms. As we'll see, the connection between linear word order and the availability of intermediate stranding emerges naturally under CL, for which syntactic movement is constrained by the linear ordering information that phase-by-phase spell-out generates.

3.2 Why CL predicts the ISG
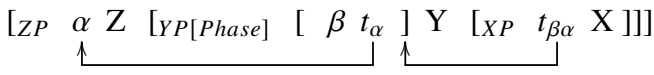
As discussed extensively in the previous chapter, the CL approach to phases derives successive-cyclic movement from the order-preserving effect of moving from a phase via its linear edge. Under this theory, if movement exits a phase from a position that is not at the linear edge, hence crossing over some material in the phase on the way out, incoherent ordering information is generated (unless appropriate order-restoring movements subsequently apply). With this system in mind, recall the generalization about intermediate stranding that this chapter argues for:
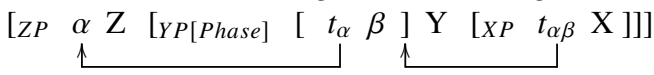
(38) *Intermediate Stranding Generalization (ISG)*
Leftward movement of a phrase α can only intermediately strand an element β if β is (or can be) ordered to the right of α before stranding occurs.

---

[10]Bošković (2018) extends Chomsky's theory to an account of movement from moved phrases which predicts something resembling (but distinct from) the ISG. See section 6 below for discussion.

The ISG is an automatic consequence of CL only permitting movement from a phase via its linear edge. If a successive-cyclically moving phrase intermediately strands material that precedes it, then that phrase illegally crosses over the material it strands in the edge as it moves on into the next phase (39a). In contrast, if the material being stranded at the edge follows the phrase that strands it by moving on (39b), then such problematic crossing at the phase edge doesn't occur:

(39)   a.   * *Intermediate stranding with crossing at the edge*
       $[_{ZP}$ α Z $[_{YP[Phase]}$ [β $t_α$] Y $[_{XP}$ $t_{βα}$ X]]]

   b.   ✓ *Intermediate stranding without crossing at the edge*
       $[_{ZP}$ α Z $[_{YP[Phase]}$ [$t_α$ β] Y $[_{XP}$ $t_{αβ}$ X]]]

To see concretely why this is so, compare the ordering information generated in the derivation of (39a) with that of (39b). In (39a), the constituent βα first moves to the edge of the phase YP. In this situation, when YP undergoes spell-out, the result is the ordering β < α < Y < X. Next, the element α is extracted into spec-ZP. Spell-out of ZP then generates the ordering information α < Z < YP. Here a contradiction arises: within YP, it was previously established that β precedes α. However, after movement of α from YP, α is linearized as preceding the content of YP, including the stranded element β which YP contains. Because α cannot both precede and follow β, this configuration is not pronounceable. In contrast, this linearization problem does not arise in (39b). Since in this derivation α precedes β in the first place, no contradiction is caused by extraction of α from YP after α pied-pipes β to the YP edge: α simply precedes β for the entire derivation. In this way, CL accurately permits only non-crossing intermediate stranding derivations like (39b), which corresponds to what the ISG describes.

Importantly, the crossing problem that derives the ISG applies only at phase edges. Thus base position preposition stranding in languages like English and Norwegian, for instance, is correctly permitted. Only after undergoing an initial step of successive cyclic movement does preposition stranding become impossible.

4   The role of locality constraints on movement

In this section, I argue that CL interacts with independently proposed locality constraints in a way that correctly rules out several potential ISG-violating derivations. This section will also clarify why the ISG contains the disjunction that if β is or can be ordered to the right of α, then α can intermediately strand β.

4.1   No phrase-bounded specifier-to-specifier movement

Here I adopt the hypothesis that movement from one specifier to another of the same phrase is unavailable. As summarized in the previous chapter, Ko (2007, 2011, 2014) observes that such a ban is predicted by the concept that movement to the specifier of a given head requires a movement-triggering feature on that head to c-command the goal phrase to be moved (Chomsky 1995, 2000, a.o.): since heads don't c-command their specifiers, it is not possible for a head to target and move a phrase from one of its specifiers to another. This constraint correctly rules out circumventing the ISG by re-arranging the content of the phase edge before stranding occurs, as described below.

If intermediate stranding fails when a moving phrase crosses what it strands while exiting the phase, then ISG-violating intermediate stranding should be permitted if the moving phrase can reach a higher position within the phase edge, above any pied-piped material that initially preceded it. In (40) below, this hypothetical configuration is illustrated with pied-piping *wh*-movement of a prepositional phrase into spec-CP (40a), followed by extraction of the *wh*-phrase from the complement of the prepositional phrase and into a higher specifier of the same CP (40b):

(40)   *Hypothetical edge-internal movement after pied-piping*

a.   *Step 1: Pied-piping into phase edge*
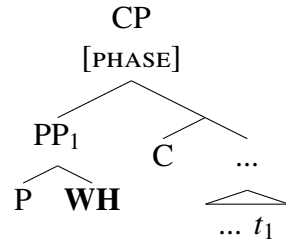
b.   *Step 2: Extraction to second specifier of the same phase*
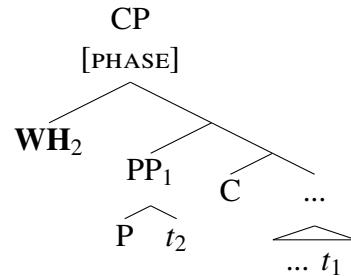
Since the edge-internal movement in (40b) brings the *wh*-phrase to a position within CP preceding the pied-piped preposition, then this movement, if available, should allow the *wh*-phrase to subsequently extract from CP without crossing P. Such movement would thus strand the preposition in spec-CP. As we've seen, such ISG-

violating stranding is not possible in reality. However, if following Ko the edge-internal movement required for this hypothetical derivation is independently ruled out, then we correctly avoid predicting the possibility of such derivations.[11]

The problematic configuration in (40b) could also be derived by first performing *wh*-movement to spec-CP without pied-piping PP, and then moving the remnant PP from its base position and into a lower specifier of the same CP via tucking-in (Richards 1997, a.o.), as we see in (41):

---

[11]As mentioned in chapter 1, Richards (2004) argues that Bulgarian allows movement of a first *wh*-phrase to spec-CP, followed by extraction of a second *wh*-phrase out of the first, and into a second specifier of the same CP. For instance, in (i) below, Richards argues that the larger *wh*-phrase first moves to spec-CP, and then the inner *wh*-phrase *po kakvo* ("of what") is extracted to either a higher or lower additional specifier of CP:

(i)   (Bulgarian, Richards 2004, ex. 14)

a.   [**Po kakvo**]$_2$ [kolko   studenti t_2 ot   Bulgaria]$_1$ vidja t_1?
      of what   how-many students   from Bulgaria you-saw
      'How many students of what from Bulgaria did you see?'

b.   [Kolko   studenti t_2 ot   Bulgaria]$_1$ [**po kakvo**]$_2$ vidja t_1?
      how-many students   from Bulgaria of what   you-saw
      'How many students of what from Bulgaria did you see?'

Since Richards' analysis of such sentences is incompatible with the ban on phrase-bounded edge-internal movement argued for in this dissertation, such facts require an alternative account in this context. See Frampton (2001) for an analysis compatible with this dissertation which in essence amounts to extraction from the lower copy of the outer *wh*-phrase.

Another route of analysis would be to posit that phrase-bounded edge-internal movement is in principle allowed, and identify a reason why it is apparent only in languages like Bulgarian. As discussed in greater detail in chapter 5, Bulgarian is a multiple *wh*-movement language, and hence, according to Richards, is unlike English in allowing multiple overt specifiers to inhabit the same projection. We might hypothesize that only multiple *wh*-movement languages (or languages that permit multiple overt specifiers in some form) will actually allow phrase-bounded edge-internal movement to occur in contexts with overt movement such as those focused on in this chapter. This hypothesis leads us to expect that in other multiple *wh*-movement languages we should also find presumed instances of phrase-bounded edge-internal movement of the variety Richards (2004) identifies for Bulgarian. However, as far as I know, no research so far has attempted to test this prediction. If other multiple *wh*-movement languages do not in fact uniformly allow configurations like (i), it may be that Bulgarian must ultimately be understood as exceptional.

(41)   *Hypothetical PP remnant movement to lower spec-CP*

   a.   *Step 1: P-stranding A′-movement*

```
            CP
          [PHASE]
         /      \
      WH₁        \
            C      ...
                  /   \
                ...     PP
                       /  \
                      P    t₁
```
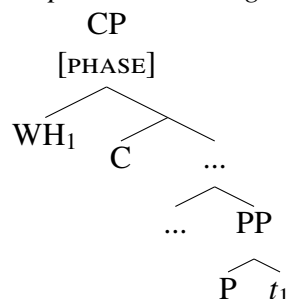
   b.   *Step 2: Movement of remnant PP*

```
            CP
          [PHASE]
         /      \
      WH₁        \
            PP₂    \
           /  \  C   ...
          P   t₁    /  \
                  ...   t₂
```
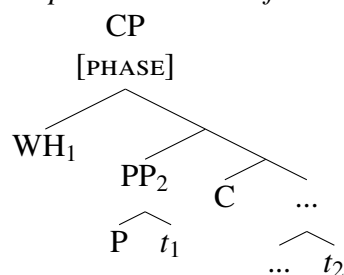
Given that this *wh*-movement and subsequent PP remnant movement would both be A′-movement to the same position, the possibility of such a derivation is likely excluded by the finding of Müller (1998) that it is impossible to extract from a given constituent and also move the resulting remnant, when both the extraction and remnant movement would be movements of the same type. See Takano (1993) and Kitahara (1994) for discussion of similar considerations.[12] This constraint applies

---

[12]The impossibility of either sub-extraction from a pied-piped PP or leftward movement of a remnant PP interestingly contrasts with the possibility of strings that can be analyzed as either sub-extraction from an extraposed PP or rightward extraposition of a remnant PP (Huck and Na 1990; Sheehan 2010, a.o.). While rightward movement combined with extraction in this way has been observed to be somewhat marginal, it is attested, and according to Huck & Na is especially acceptable under focus, as in the following example:

(i)   Here's an article in the Tribune by Trevor, of all people; he's someone [who₂ I'd expect to read a story t_1 in the paper [**ABOUT** t_2]₁].
   (Adapted from Huck and Na 1990 pp. 66)

There are multiple ways that this chapter's proposals can accommodate this fact. On one hand, if extraposition precedes extraction, extraction will be licit provided that the extraposed phrase lands in a position that is not in the phase edge. In this case, the phase head will c-command that phrase and can trigger extraction from it. On the other hand, if extraction precedes extraposition of the remnant, then following the findings of Müller (1998) mentioned above the derivation will be licit if these two movements are syntactically distinct, due to being driven by different features, for instance.

to any remnant constituent we might consider, not just PP, and thus broadly prevents later remnant movement of the relevant sort from ever being available to derive the appearance of ISG-violating stranding.[13]

4.2 The role of anti-locality

A related constraint on movement in syntax is anti-locality, the concept that movement must not be too short (Bošković 1997; Ishii 1999; Grohmann 2003; Abels 2003, 2012; Erlewine 2016, a.o.). Many formulations of anti-locality at least partially subsume the ban on phrase-bounded specifier-to-specifier movement just discussed. Additionally, Abels' version of anti-locality also bans movement of a head's complement to its specifier. I adopt this constraint on movement as well, since it correctly rules out the possibility of deriving certain ISG-violating configurations via movement within the pied-piped constituent.

For instance, a *wh*-phrase complement of a prepositional phrase might conceivably move to spec-PP prior to pied-piping the PP into spec-CP, as diagrammed in (42) below. In the resulting configuration, the *wh*-phrase occupies the left linear edge of PP and of the containing CP:

(42)   *Hypothetical movement internal to pied-piped PP*

$$\begin{array}{c}\text{CP}\\ [\text{PHASE}]\end{array}$$

PP$_2$ C ...

WH$_1$ P t_1 ... t_2
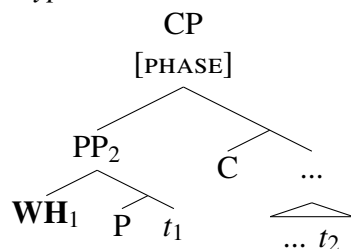
After these movements occur, the *wh*-phrase could be extracted from CP without crossing over P on the way out, thus undesirably deriving intermediate stranding of

---

[13]This chapter's analysis of the ban on intermediate preposition stranding automatically predicts a generalization from Bruening (2018), stated as follows:

(i)   A PP can be extracted from as long as it has not moved out of VoiceP (=vP); a PP that has moved out of VoiceP is an island to extraction.
(Bruening 2018, ex. 61)

Under the framework developed in this chapter, before the vP spells-out, there is an opportunity for a DP complement of P to reach a position preceding P before the relative order of the two is fixed. However, by the time a given PP has moved beyond the local vP, it will already have been linearized within that vP. Consequently, if P-stranding did not occur within the vP and thus the spell-out of vP established the ordering P < DP prior to movement of PP from that phase, then order preservation will prevent P from ever being stranded at a later point in the derivation.

the preposition, as the diagram in (43) shows:

(43)   *Hypothetical intermediate P-stranding after movement within PP*
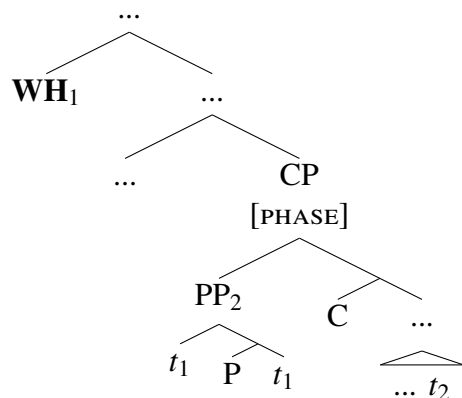
However, since the required movement from complement to specifier of PP is banned by anti-locality, a configuration like (42) cannot be derived, and thus cannot serve as the input for an ISG-violating derivation like (43).[14] The same concerns prevent such highly local movement from ever being available to derive ISG-violating extraction of a complement to a pied-piped phrase.

4.2.1   On the nature of PP

The ruling-out of PP-internal movement in the scenario just considered is in correspondence with Abels (2003), who argues that movement from complement to specifier of PP is banned by anti-locality, but that not all languages require such movement in order to extract from PP. In particular, Abels (2003) argues that if PP

---

[14]Abels' anti-locality as applied in this section raises a question about *swiping*, which inverts P and its *wh*-complement under sluicing (Ross 1969; Merchant 2002):

(i)   I know John went somewhere, but I don't know **where to**.

This chapter does not allow swiping to be derived by movement of the *wh*-phrase within PP, or from PP to a second specifier of the embedded CP. Since the inversion characteristic of swiping is not possible without ellipsis, maintaining that such movements are unavailable appears tenable:

(ii)   I know John went somewhere, but I don't know (to) where (*to) he went.

The fact that swiping is generally exclusive to single-word *wh*-phrases may provide evidence that swiping is derived by a PF process available under ellipsis rather than by usual syntactic movement:

(iii)   * I know John read about a few languages, but I don't know **how many languages about**.

The nature of swiping is certainly relevant to this chapter, but I must leave this topic aside for now.

is not a phase in a given language, there is no need for extraction of the complement of a PP to pass through the PP edge, hence in such language P-stranding can be derived by extraction directly from the interior of PP (44):

(44)   *P-stranding from a non-phasal PP bypasses PP edge* (Abels 2003)

```
              CP
            /    \
        WH_1      \
             C     ...
                  /  \
               ...    PP
                     /  \
                    P    t_1
```
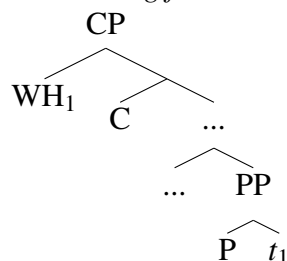
In contrast, Abels (2003) hypothesizes that in languages where P-stranding is banned, PP is a phase. In this case, the complement of P would need to exit PP via its edge, but since such movement via spec-PP is banned by anti-locality, P-stranding is impossible for such languages (45):

(45)   *Anti-locality prevents stranding of phasal PP* (Abels 2003)

```
              CP
            /    \
        WH_1      \
             C     ...
                  /  \
               ...    PP
                     [PHASE]
                    /   \
                 *t_1    \
                      P   t_1
```
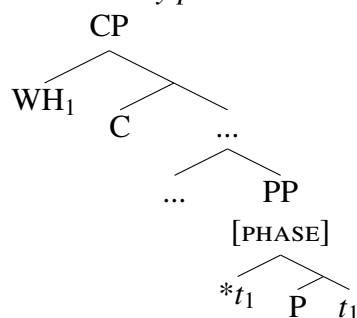
While Abels (2012) also argues for complement-to-specifier anti-locality, Abels (2012) seeks to remove the variability in phasehood proposed in Abels (2003). In particular, Abels (2012) takes PP to be a phase in all languages, and hypothesizes that in P-stranding languages PP is more structurally complex, such that extraction via its edge is possible. This analysis is diagrammed in (46) below:

(46)   *Phasal PP with more structure permits extraction via its edge* (Abels 2012)
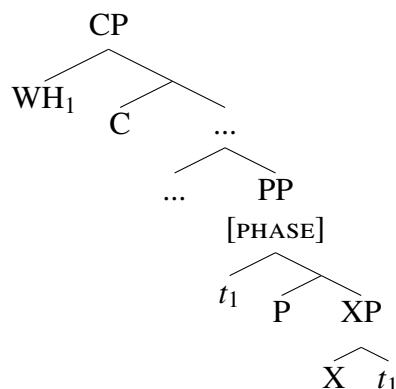
This proposal from Abels (2012) is not compatible with the present chapter, for which the possibility of movement via the edge of PP in languages like English would wrongly rule in the possibility of intermediate preposition stranding, as the text surrounding examples (42) and (43) above has already discussed. Abels notes that there is not much direct evidence favoring his (2012) approach over his (2003) one, given that P-stranding languages do not display any obvious surface-evident indication of greater structure within the PP. The arguments of the present chapter, if correct, can be taken as evidence that Abels (2003) is in fact on the right track in positing that P-stranding languages have a non-phasal PP.[15]

---

[15]Sabbagh (2007) proposes that PP is a phase, chiefly in order to account for the fact that rightward movement supposedly cannot strand a PP:

(i)   * Jamie walked [$_{PP}$ into t_1] suddenly, [the dean's office]$_1$.
        (Sabbagh 2007, ex. 2c)

Sabbagh argues that the phasehood of PP also explains the unacceptability of the following example, involving rightward movement from PP, but not P-stranding per se:

(ii)   * Josh edited a review [$_{PP}$ of an article t_1] for Jamie yesterday, [about verb movement]$_1$.
        (Adapted from Sabbagh 2007, ex. 2d)

In my judgment, this sentence is perfectly acceptable. If my judgment is correct, Sabbagh's appeal to the phasehood of PP may be too general. If so, it could be that the English PP is not a phase, and that something independent blocks rightward movement of the complement of P.

   While the English PP may not generally be a phase, if the phasehood of PP is in principle variable, it could be that some PPs in English are phases and thus disallow P-stranding. This intuition facilitates a hypothesis for the fact that even in English, not all prepositions are strandable, such as the preposition *since* (von Fintel and Iatridou 2019):

(iii)   a.   [Since what year]$_1$ have you been living here t_1?
        b.   *?? [What year]$_1$ have you been living here since t_1?

4.3   When anti-locality is irrelevant

While we have just seen a scenario in which anti-locality asserts its influence, anti-locality is predicted to be irrelevant if the pied-piped constituent is structurally larger, however. For instance, in the schema in (47) below, a *wh*-phrase pied-pipes ZP and the containing phrase YP into a CP edge. Since movement of the *wh*-phrase from ZP to the edge of YP does not violate anti-locality, such movement to the linear edge of this two-layered pied-piped constituent is possible. After such movement as well as pied-piping into the CP edge occur, the *wh*-phrase ultimately occupies the linear edge of the pied-piped phrase and of the containing CP phase, as (47) shows:

(47)   *Legal movement within larger pied-piped constituent*
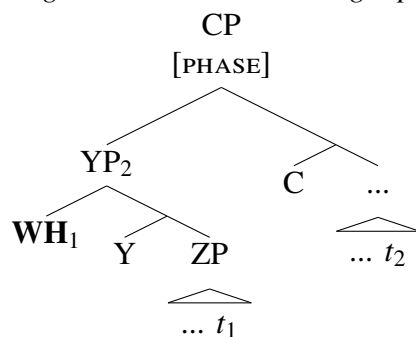
In principle the *wh*-phrase could subsequently move from the CP in (47), intermediately stranding the pied-piped YP in the CP edge. Since the *wh*-phrase would not cross any material as it exits this structure, such stranding should be licit (48):

(48)   *Intermediate stranding after movement within pied-piped constituent*
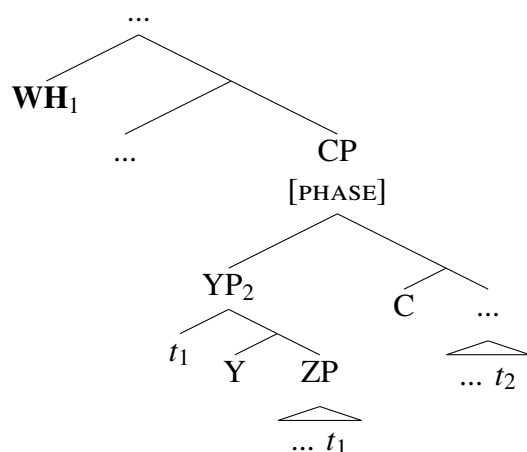
If phasehood really is subject to variability, it raises the question of what determines that variability. While I cannot answer this question here, see section 5.3 below for further discussion.

This prediction provides an analysis for intermediate stranding of material that can be ordered on either side of what it merges with, as we saw in section 2 for certain adjuncts of *wh*-phrases in English, *maar* ("only") in Dutch, and ambivalent adpositions in Russian. If the word order variability these elements display stems from the possibility of movement within the relevant constituent, then such movement should provide a means of deriving ISG-obeying intermediate stranding. This expectation is connected to the disjunction in the definition of the ISG that if an element β is or can be ordered to the right of α, then α can intermediately strand β. For the account developed here, this aspect of the ISG emerges from the fact that if a given phrase can move to the left edge of a containing constituent, then such movement should be available to feed later intermediate stranding of that constituent.

To see precisely how the derivation of such an instance of intermediate stranding would proceed, let's consider Russian ambivalent Ps once more. Podobryaev (2009) argues that anti-locality bans movement within typical Russian PPs, but that the strandable variably ordered Ps discussed in section 2.8 are more structurally complex, such that movement within them is possible. Podobryaev suggests, consistent with the diachronic facts, that such PPs are derived from nouns via the N to P incorporation in (49) below. If these adpositions indeed involve two phrasal layers, then anti-locality respecting A′-movement to the edge of such an adposition should be licit, as we see with a moved *wh*-phrase in (49):

(49)   *Anti-locality respecting movement to edge of complex Russian adposition*
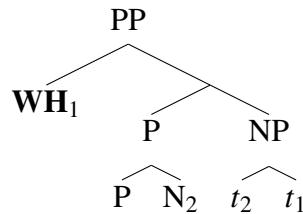
Such movement is predicted to feed intermediate stranding in the following way. After adposition-internal movement of the *wh*-phrase in (49), further *wh*-movement pied-pipes the adposition to a phase edge, as in (50a). Here the *wh*-phrase occupies the linear edge of the adposition and of the containing phase, thus the *wh*-phrase can move on, stranding the adposition in the phase edge as in (50b).

(50)   a.   *Pied-piping of complex PP after movement within it*

```
                        CP
                      [PHASE]
                    /         \
              PP₃              /    \
            /      \          C      ...
        WH₁       /    \            /\
                 P      NP        ... t₃
                /\      /\
               P  N₂   t₂  t₁
```
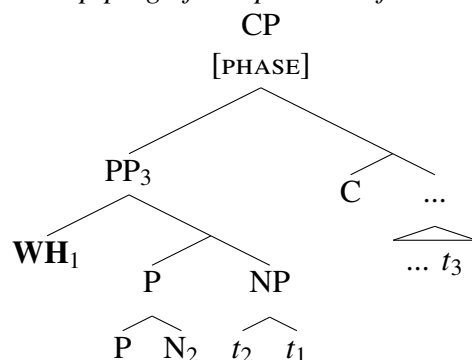
   b.   *Intermediate stranding of complex PP after pied-piping*

```
             ...
           /     \
       WH₁        \
         /           \
      ...              \
              ┌──────────────────────────┐
              │              CP           │
              │            [PHASE]        │
              │          /         \      │
              │     PP₃             /  \   │
              │    /    \          C    ...│
              │  t₁    /    \          /\  │
              │       P      NP      ... t₃│
              │      /\      /\            │
              │     P  N₂   t₂  t₁         │
              └──────────────────────────┘
```
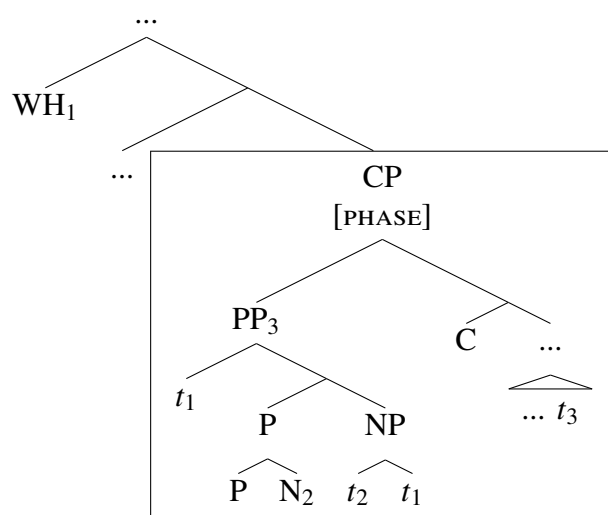
Thus the possibility of moving to the left edge of the complex adposition provides the means for intermediate stranding of this constituent to succeed, as we have seen it does in reality. In contrast, this account predicts that if the extracting *wh*-phrase here had remained to the right of the rest of the content of the complex adposition by not moving to its edge, then such intermediate stranding would fail.

Note that while pied-piping a constituent large enough to permit movement through its edge is one predicted way to yield legal intermediate stranding, the same result is predicted for a constituent whose internal order is simply free due to optionality (or under-specification) of linearization rules. This alternative could be correct for the stranding of the English adjuncts discussed in sub-section 2.2 above: given that English adjuncts can often be freely ordered either left or right of the phrase they adjoin to, this may simply be a matter of free ordering feeding intermediate stranding.

Heck (2009) observes several contexts with so-called secondary *wh*-movement

of a *wh*-phrase to the edge of the constituent it pied-pipes, as seen in certain French relative clauses (51), and with possessors in Mayan languages like Tzotzil (52):

(51)　*Secondary pied-piping in French*
　　　un homme [[$_{DP}$ dont$_2$　le　comportement t_2]$_3$ t_3 devient　inquiétant]
　　　a　man　　　of.who the behavior　　　　becomes alarming

　　　'a man whose behavior becomes alarming'
　　　(Heck 2009, ex. 33a)

(52)　*Secondary pied-piping in Tzotzil*
　　　[Buch'u$_2$ x-ch'amal t_2]$_4$ i-cham t_4?
　　　whose　child　　　died

　　　'Whose child died?'
　　　(Heck 2009, ex. 30a)

A reviewer for an earlier version of this work points out that this chapter predicts such secondary *wh*-movement to feed intermediate stranding of the pied-piped phrase. While I do not know if this prediction is correct, if intermediate stranding is impossible in such circumstances, it would not be in conflict with the present chapter's arguments to posit that the relevant constituents are islands, despite permitting movement to their edge. This is true of relative clauses, at least, under an analysis in which relativization involves movement of an operator to the clause edge.

4.3.1　Extraction from a moved DP in spec-CP

These concerns lead to an understanding of configurations in English involving A′-extraction from a constituent that has undergone topicalization or *wh*-movement to an embedded spec-CP (53). Such patterns have been discussed by at least Pesetsky (1982), Chomsky (1986), Lasnik and Saito (1999), and Sauerland (1999).

(53)　*Extraction from a moved DP in spec-CP*
　　　? [What student]$_2$ did Ann ask [$_{CP}$ [what picture of t_2]$_1$ to put up t_1]?
　　　(Sauerland 1999, 24c)

This is not an instance of intermediate stranding per se, since unlike the scenarios examined earlier in this chapter the two moved phrases here must both move for independent reasons to separate CPs. Nevertheless, the present approach accurately predicts the possibility of such sentences, in the following way.

　　First, the inner *wh*-phrase *what student*$_2$ moves to the spec-DP of the outer *wh*-phrase *what picture of*$_1$. This movement from the complement of NP to spec-DP does not violate anti-locality:

(54)   *DP-internal movement*
        $[_{DP1}$ $[_{DP2}$ what student$]_2$ what picture of $t_2]$

The outer *wh*-phrase then moves to the embedded spec-CP:

(55)   *Movement of outer DP to embedded spec-CP*
        $[_{CP}$ $[_{DP1}$ $[_{DP2}$ What student$]_2$ what picture of $t_2]_1$ to put up $t_1]$?

The inner *wh*-phrase is then the leftmost phrase in the embedded CP since it occupies the specifier of the outer *wh*-phrase, which in turn occupies the specifier of the embedded CP. From this peripheral position, the inner *wh*-phrase can move on into the matrix clause, creating the structure in (53) above.

    In short, since by hypothesis extraction via the linear edge of the outer *wh*-phrase is available, we expect the possibility of a derivation in which this *wh*-phrase moves and a second *wh*-phrase is later extracted from it in a linearization-respecting way.[16]

4.4   On quantifier float under A-movement

It is in the context of A′-movement that the type of stranding this chapter focuses on emerges most clearly. However, many languages also allow the apparent stranding of quantifiers under A-movement, a phenomenon often referred to as *floating quantification*, exemplified in (56):

---

[16]In my judgment, a sentence like (53) is clearly unacceptable for extraction from a moved PP:

(i)   * Who$_2$ did Ann ask $[_{CP}$ $[_{PP}$ to which relative of $t_2]_1$ we should send the letter $t_1]$?

The ban on phrase-bounded specifier-to-specifier movement rules out a derivation where *who* extracts from the PP in (i) to a higher specifier of CP in order to precede the preposition before moving on into the matrix clause. However, it is not obvious why *who* could not alternatively move to spec-PP before the PP moves to the CP edge, and then extract from it. If DP is a phase (as is important for the analysis of extraposition in chapter 6), such movement to spec-PP would have to be preceded by a step of movement through spec-DP. Movement from spec-DP to spec-PP would be banned by the version of anti-locality in Bošković (2005); Erlewine (2016), and Brillman and Hirsch (2016), which disallows movement from the specifier of a given phrase to that of the immediately dominating one (as discussed further in chapter 4).

    This understanding accurately predicts no issue for extraction from a DP complement of P:

(ii)   Guess who$_1$ I shared a burrito $[_{PP}$ with $[_{DP}$ a friend of $t_1]]$ today!

Since the English PP is not a phase, as argued earlier in this chapter, extraction from the PP in (ii) can proceed directly from spec-DP. In the absence of movement of this PP to a phase edge, there is no reason for extraction via spec-PP to take place.

(56)   *Floated quantifiers*

   a.   *English*
        The students have **all** had lunch.

   b.   *French*
        ([Sportiche 1988](), ex. 2b)
        Les enfants ont   **tous** vu   ce   film.
        the kids   have all   seen this film

Some such patterns may present exceptions to the ISG at first glance, given the ban on movement within phrase edges adopted earlier in this section. To illustrate the potential problem, it will suffice to consider the English example in (56a). Here the subject DP is separated from its associated quantifier *all* by an intervening auxiliary. If *all* was originally merged to the subject which then A-moved to spec-TP, then the fact that *all* cannot follow a DP that has not moved, like those in (57) below, suggests that *all* originally preceded the subject in (56a) prior to A-movement.[17]

(57)   *Unmoved DP must be preceded by "all"*

   a.   I saw (all) the cats (*all).
   b.   I gave a cookie to (all) the kids (*all).

Sentences like (56a) have a word order consistent with *all* having been stranded in spec-vP by A-movement of the subject from its θ-position, as illustrated in (58) below. Notice that if *all* really must have originally preceded the subject DP, then the subject's A-movement from vP would cross over this stranded quantifier on the way out of vP, as (58) shows:

---

[17]In contrast to what we see for lexical nominal phrases in (57), pronouns can precede an associated *all* even in non-subject positions:

(i)   a.   The investigator talked to **them all**.
      b.   If just one of us steals office supplies, they'll fire **us all**!

This may indicate that there is indeed a peripheral landing site available in DPs/QPs headed by *all*, but that for reasons yet to be identified, this position is only available for pronouns.

   Tangentially, I observe that in such contexts *all* cannot precede the object pronoun, unless *of* is included, whereas this is only optional for lexical nominals:

(i)   a.   They'll fire **all *(of) us/you/them**!
      b.   I ate **all (of) the cookies**.

While it is beyond the scope of this chapter to analyze this fact, it is clear that lexical nominals and pronouns do not behave syntactically identically in this context.

(58)   *Stranding analysis of quantifier float by subject movement*

```
                              TP
                  ┌───────────┴───────────┐
               DP₁                   ┌─────┴─────┐
             ┌──┴──┐               TP          AuxP
        The students           ┌───┴──┐      ┌───┴────┐
                             Aux₂    T      t₂        vP
                             have                   [PHASE]
                                              ┌────────┴────────┐
                                             QP                 │
                                           ┌──┴──┐         ┌────┼────┐
                                           Q    t₁         v       VP
                                          all                   ┌──┴──┐
                                                                V     DP
                                                               had   lunch
```
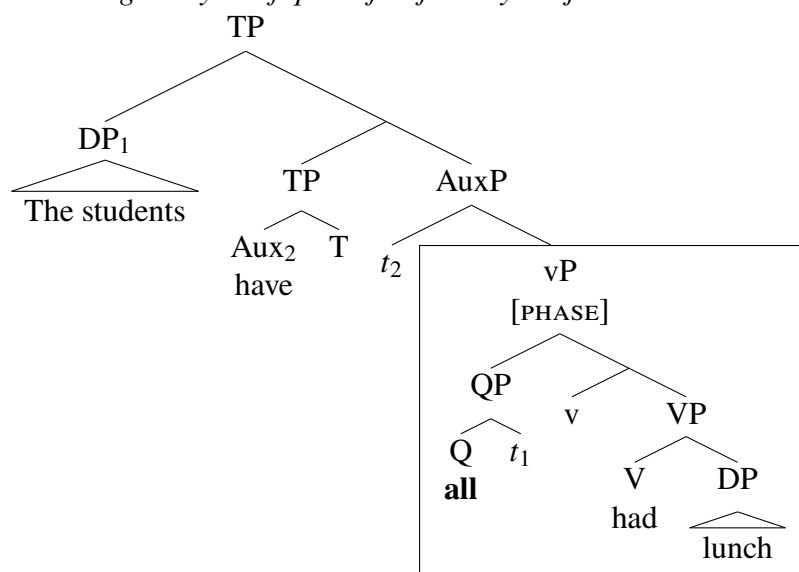
Given the ban on phrase-bounded specifier-to-specifier movement, it is not possible for the subject to move to a higher spec-vP above the quantifier to avoid this crossing problem. Thus movement of the subject to spec-TP must cross the stranded quantifier in the linear edge of vP. Therefore the possibility of sentences like (56a) appears to constitute a violation of the ISG, if the right analysis for them is (58).[18]

However, if the floated quantifier in (56a) was not derived by stranding, then this sentence does not constitute an exception to the ISG. A non-stranding approach to quantifier float has been explored by several works (Dowty and Brodie 1984; Bobaljik 1995; Doetjes 1997; Fitzpatrick 2006, a.o.). Fitzpatrick (2006) in particular follows preceding works in taking quantifier float under A-movement to be essentially adverbial, involving an adjunct containing the quantifier and a null PRO co-indexed with the A-moved phrase, as shown in the diagram in (59) below. Importantly for the present chapter, Fitzpatrick goes on to argue that while quantifier stranding under A′-movement is genuine sub-extraction, apparent quantifier float under A-movement is always adverbial in this way.

---

[18]However, if Bošković (2004) is right that floated quantifiers do not appear in θ-positions, then the quantifier in (56a) must occupy a position other than spec-vP. If this is so, then such sentences cannot instantiate violations of the ISG.

(59)   *Adverbial analysis of floated quantifier: No stranding, no ISG violation*
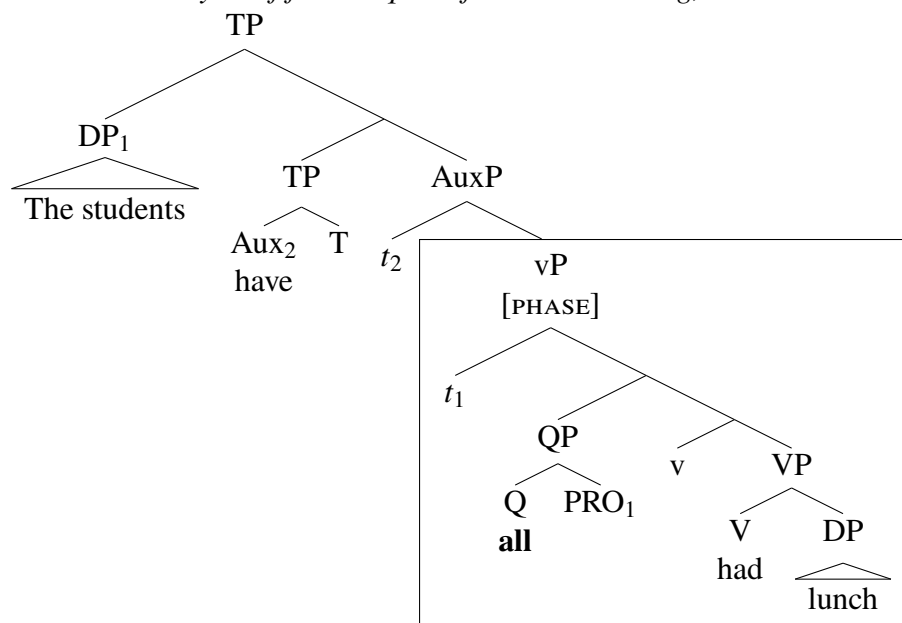
If Fitzpatrick is correct, apparent quantifier stranding / floating under A-movement in fact never involves sub-extraction, and therefore, cannot instantiate intermediate stranding of the variety that the present chapter examines. From this perspective, quantifier float poses no challenge to the ISG.

5   On when an edge is available for stranding
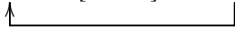
So far, this chapter has focused on understanding what sorts of elements can be stranded in intermediate positions. Another relevant topic in this domain is the question of what positions are, in principle, available to be stranded in. While the present chapter cannot provide a full understanding of this issue, this section will discuss several relevant predictions and possibilities.
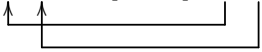
5.1   Predicting the distribution of stranding in spec-vP

As previewed in the introduction, the linearization-based explanation for successive-cyclic movement in the CL theory also predicts that certain exceptions to successive-cyclicity are possible, as long as additional movements occur that keep linearization coherent. In particular, the prediction is that if a given phase-exiting movement crosses over some material in that phase on the way out, the derivation will still remain contradiction-free if the crossed-over material also moves into the next phase,

to a position preceding what originally crossed it (60):

(60) a. *Illicit crossing at the edge...*
          $* [_{YP[Phase]} \; \beta \; [_{XP[Phase]} \; \alpha \; \beta \;]]$

      b. *...repaired by restoring original order*
          $\checkmark [_{YP[Phase]} \; \alpha \; \beta \; [_{XP[Phase]} \; \alpha \; \beta \;]]$

I argue that we find a verification of this prediction in the distribution of *exactly* and other adjuncts to *wh*-phrases in English, which as shown in section 2.2 above, can be stranded at clause edges:

(61) *Adjunct stranding at clause edge by wh-movement*
      a. **What**$_1$ did you suppose t_1 (**exactly/precisely**) (that) they wanted t_1?
      b. **How much saffron**$_1$ did the chef say t_1 (**to the closest gram**) (that) we need t_1?

If vP is a phase in addition to CP, then we also in principle expect these adjuncts to be able to be stranded in spec-vP, and thus appear in a position between the subject and verb. However, this linear position is also a possible location for adverbs:

(62) *Adverbs may appear between subject and verb*
      a. I **quickly** ate the beans.
      b. She **precisely** drew a cube.

For this reason I focus on strandable adjuncts like that in (61b), which can't be parsed as adverbs. Example (63) below attempts to strand such an adjunct in spec-vP, which proves to be unacceptable (though see Zyman (To appear) for an alternative view):

(63) *No adjunct stranding in spec-vP*
      a. How much flour (to the nearest pound) did you [$_{vP}$ (***to the nearest pound**) tell me [$_{CP}$ (to the nearest pound) that the bakery [$_{vP}$ (***to the nearest pound**) asked you for (to the nearest pound)]]]?
      b. Tell me [$_{CP}$ how many grams of tranquilizer (to the third decimal place) the researchers [$_{vP}$ (***to the third decimal place**) reported [$_{CP}$ (to the third decimal place) that they [$_{vP}$ (***to the third decimal place**) used to sedate the tiger (to the third decimal place)]]]].

The impossibility of such stranding in spec-vP is expected, given the prediction of CL discussed at the beginning of this section. To see why, consider the interaction of successive-cyclic A′-movement with A-movement of the subject. CL requires an

A′-moving non-subject argument to temporarily land in the most peripheral position of the vP phase on its way to spec-CP. This will be a spec-vP above the external argument (EA) in situ in its θ-position. No linearization problem arises when the subject later A-moves to spec-TP across that outer spec-vP formed by successive-cyclic A′-movement, provided that the content of that outer specifier moves to spec-CP. After such movements, the relative order of the moved phrases established in vP and CP is the same, yielding a coherent linearization, as (64) exemplifies with movement of an object *wh*-phrase:

(64)   *WH < EA order maintained in vP and CP*
         [$_{CP}$ WH C EA T [$_{vP}$ ~~WH~~ ~~EA~~ v-V ~~WH~~]]

However, if this *wh*-movement strands something in that outer spec-vP, then movement of the subject across that position encounters a crossing problem. This is shown in (65), where we see that while there is no issue if the moving *wh*-phrase pied-pipes the element α to spec-CP, there is predicted to be a problem if α is stranded in vP and is thus crossed by A-movement of EA:

(65)   *Conflict between EA movement and stranding in the vP edge*
         [$_{CP}$ WH-($\checkmark\alpha$) C EA T [$_{vP}$ ~~WH~~-(*α) ~~EA~~ v-V ~~WH~~-($\checkmark\alpha$)]]

Thus we have a reason for the unacceptability of the instances of vP level stranding attempted in the transitive clauses in (63) above. As expected, the same stranding restriction also holds for unergative clauses, which also involve an A-moving EA:

(66)   *No adjunct stranding in spec-vP: Unergative predicate*
    a.   How many bad jokes (to the nearest dozen) did the audience [$_{vP}$ (**to the nearest dozen**) laugh in spite of (to the nearest dozen)]?
    b.   How much money (to the nearest million) did the governor [$_{vP}$ (**to the nearest million**) resign for (to the nearest million)]?

    If all vPs are phases (Sauerland 2003; Legate 2003; Ko 2014) then under CL, theme subjects of unaccusatives and passives are expected to pass through the edge of vP, given that V moves to v in English (Larson 1988; Chomsky 1995; Kratzer 1996, a.o.). Such subject movement is necessary to ensure that the theme subject precedes V within vP, as it will later after movement to spec-TP (67):

(67)   *V movement and theme subject movement within vP*
       [$_{vP}$ S v-V [$_{VP}$ V̶ S]]
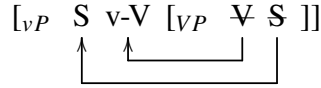
Any A′-movement in such contexts will form a higher spec-vP above the moved theme subject, just as occurs with EAs, which originate in spec-vP rather than moving there. This being the case, later movement of a theme subject to spec-TP must cross over anything stranded in the vP edge by A′-movement, just as we've seen with EAs. Thus intermediate stranding in the edge of vP should not be possible in passive and unaccusative derivations. This prediction is correct:

(68)   *No adjunct stranding in spec-vP: Passive*[19]
       [How many boats]$_1$ (to the nearest hundred) has the Micronesian navy [$_{vP}$ t_1 (**to the nearest hundred**) been provided with t_1 (to the nearest hundred)]?

(69)   *No adjunct stranding in spec-vP: Unaccusative*
       [How many firefighters]$_1$ (to the nearest dozen) did the house [$_{vP}$ t_1 (**to the nearest dozen**) burn down despite the efforts of t_1 (to the nearest dozen)]?

These contexts where A-movement of the subject blocks intermediate stranding in the vP edge would be acceptable if it were possible to rearrange the specifiers of vP, as in (70). Here successive-cyclic movement of a *wh*-phrase pied-pipes α to the vP edge, above the subject. Then the subject moves over the *wh*-phrase and α, and then the *wh*-phrase moves over the subject, stranding α below:

(70)   *Successive-cyclic movement followed by rearranging in the vP edge*
       [$_{vP}$ WH S [W̶H̶-α] S̶ v-V [$_{VP}$ [W̶H̶-$\alpha̶$]]]
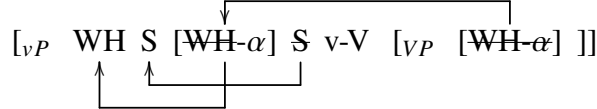
If these movements occurred within vP, they would yield a vP structure that is consistent with the final ordering that will be produced in CP, and importantly, movement of the subject to spec-TP would not cross the stranded α with these rearrangements in place. However, given the ban on phrase-bounded edge-internal movement, such a derivation is not available.

The same concerns lead to a solution for a puzzle from McCloskey (2000). Recall McCloskey's argument that the strandable *wh*-associated *all* in West Ulster English can be left behind at clause edges because movement from CP successive-cyclically passes through its edge (71):[20]

---

[19] I assume that the edge of the clause-internal phase in passives is left of the passive auxiliary *be*, which following Harwood (2015), is merged in v.

[20] A reviewer for a previous iteration of this work notes that if such stranding is really in spec-CP as

(71)   *All-stranding in spec-CP*
        What$_1$ (all) did he say [$_{CP}$ t_1 **(all)** that we should buy t_1 (all)]?
        (McCloskey 2000, ex. 8)

McCloskey also notes that if vP is a phase, West Ulster English is predicted to allow *all*-stranding in its edge, contrary to fact, as we see in (72) below. McCloskey's analysis of West Ulster English suggests that V moves to a head above vP, thus his examples showing this gap in the stranding paradigm attempt *all*-stranding after V:

(72)   *No all-stranding in spec-vP*
        What$_1$ did he **tell**$_2$ [$_{vP}$ t_1 **(*all)** t_2 his friends [$_{CP}$ t_1 (all) that he wanted t_1]]?
        (McCloskey 2000, ex. 14e)

The impossibility of such stranding follows directly from the concerns just discussed, given A-movement of the subject from vP. The movement of V from vP that McCloskey posits for West Ulster English provides a second reason why such stranding should be banned: there is no position in vP where V can precede the specifiers of vP. Hence movement of V from vP will necessarily cross over any specifiers of vP, forcing them to be evacuated, thus blocking stranding there. This is illustrated in (73) below.[21] [22]

McCloskey (2000) argues, the fact that an overt C can co-occur with it is surprising, given something like the Doubly Filled Comp Filter. It is conceivable that the Doubly Filled Comp Filter only applies when a full phrase is sitting in spec-CP, rather than something like a remnant quantifier. The same might be said of examples like (61) above. The reviewer also points out that the Doubly Filled Comp Filter fails to apply in the matrix CP of a *wh*-question, where C is filled by an auxiliary (*What **will** John buy?*). Evidently, the Doubly Filled Comp Filter is only applicable to the complementizer *that*, and not in all situations. These facts indeed raise intriguing questions about the nature of the Doubly Filled Comp Filter, but are tangential to this chapter's focus.

[21] This analysis predicts that stranding in spec-vP is possible only when what is stranded isn't later crossed by non-successive-cyclic movement from vP. A situation that may represent a verification of this prediction is the Korean example shown in footnote 9 above. This could instantiate stranding in spec-vP with a subject that is left in situ, which thus should not interrupt stranding. Any movement of V should also not interrupt stranding here, since Korean is head-final, meaning that any movement of V will be linearly rightward.

[22] If vP is a phase in English, and there is no way for V to precede an external argument subject at the vP level for the reasons just discussed, then we make the more general prediction that V should never be able to precede such a subject in English. This prediction appears to be violated by sentences like (i). If (i) was derived by raising of *John* from the infinitive followed by remnant movement of a constituent containing the infinitive, then here the content of the vP in which *John* originated comes to precede *John* after movement occurs, in such a way that CL should not permit:

(i)   Tell me [how likely t_1 to **win** the race]$_2$ **John**$_1$ is t_2.

However, CL does not predict the impossibility of this sentence if it instead involves *John* controlling a PRO in the infinitive rather than having raised from it:

(73)   *V movement from vP blocks stranding in the vP edge*

$[_{CP}$ wh_1 ... $[_{XP}$ X ... $[_{vP}$ t_1 (***all***) v V t_1]]]
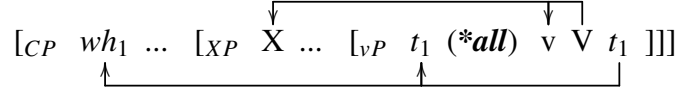
5.1.1   On the origination of expletive *there*

Several works argue that expletive *there* is externally merged in spec-vP (Biberaur and Richards 2005; Deal 2009, a.o.) before A-moving to spec-TP. If this is so, we expect A-movement of the expletive to result in a crossing effect that makes it impossible for A′-movement in English to strand an adjunct in spec-vP in expletive constructions. This prediction is correct, as we see in (74):

(74)   *No adjunct stranding in spec-vP in expletive constructions*

    a.   [How many kids]$_1$ have there $[_{vP}$ t_1 (***exactly***) been t_1 in the office today]?

    b.   [How many kilos of gold]$_1$ (to the nearest hundred) have there $[_{vP}$ t_1 (***to the nearest hundred***) been t_1 consumed in the production of fancy pens]?

These facts only stand as evidence for A-movement of the expletive if vP is a phase in expletive constructions. If it is not, there would be no reason to expect successive-cyclic movement through, or stranding in, this phrase's edge. Thus evidence for vP phasehood in this environment is necessary. As I discuss extensively in chapter 5, Nissenbaum (2000) argues that parasitic gaps in sentential adjuncts are licensed by successive-cyclic movement through spec-vP. If such a parasitic gap can be licensed in a given environment, it thus suggests that successive-cyclic movement

---

(ii)   Tell me [how likely PRO$_1$ to win the race]$_2$ John$_1$ is t_2.

In such a case, *John* did not originate in and therefore was not previously ordered with respect to the content of the infinitive that is carried along by remnant movement here, and thus there would be no basis for an ordering contradiction to arise in this sentence. Independent evidence that such sentences do in fact involve control comes from the fact that they are incompatible with idiomatic interpretations (Lasnik and Saito 1999), as shown in (iii). The sentences in (iii) are acceptable only under literal interpretations, not idiomatic ones:

(iii)   a.   Guess [how likely to get out of the bag] the cat is.
         (Not interpretable as *Guess how likely the secret is to be revealed*.)
    b.   Tell me [how likely the shit is] to hit the fan today.
         (Not interpretable as *Tell me how likely there is to be a disaster today*.)

While *likely* is typically able to be involved in a raising derivation rather than a control one, we expect the raising parse of these sentences and thus an idiomatic interpretation for them to be ruled out, given that the relevant derivation would violate CL as just described.

through spec-vP occurred. Legate (2003) uses this logic to diagnose the phasehood of verb phrases in several contexts, but Legate did not perform this test in expletive constructions. When we do, I argue that we find successful parasitic gap licensing:

(75)   *Parasitic gaps licensed by A′-movement in expletive constructions*
   a.   Who$_1$ was there a big rumor about t_1 [after the police arrested PG$_1$]?
   b.   Who$_1$ was there a party for t_1 [before the boss promoted PG$_1$]?

To the extent that this constitutes evidence for a vP phase in expletive constructions, the facts in (74) indicate movement of the expletive from vP.[23]

5.2   Unexpected absences of stranding in spec-CP

The previous subsection analyzed a circumstance under which stranding in spec-vP is banned. There are also patterns where material that appears to be capable of being stranded in situ fails to be stranded in the CP edge, even though such stranding would fit the ISG. Several such patterns are shown below:

(76)   *Base position stranding but no intermediate stranding at clause edge*
   a.   *Combien split in French*
        Combien$_1$ (de lirves) crois-tu t_1 (***de livres**) que je devrais lire
        How.many of books believe-you   of   books that I should read
        t_1 (de livres)?
           of books
        'How many books do you believe that I should read?'
        (Vincent Rouillard, p.c.)

   b.   *Possessor extraction in Greek*
        Pianou$_1$ (to vivlio) ipe o   Yanis t_1 (***to vivlio**) oti (i   Maria)
        Whose (the book) said the Yanis   (**the book**) that the Maria
        diavase / diavase (i   Maria) t_1 (to   vivlio)
        read   / read   the Maria   (the book)?
        'Whose book did Yanis say that Maria read?'
        (Sabine Iatridou, p.c.)

---

[23]A criticism of the logic used by Legate (2003) is that while PG licensing may demonstrate the possibility of successive-cyclic movement through vP, it does not demonstrate its necessity, which we do expect to find if such movement is phase-motivated. The edge of vP is, however, not final a landing site for any independently attested movement process in English, meaning that it is not obvious what motivation there is for movement via this position other than forced successive-cyclicity. In footnote 12 of chapter 5 we will see evidence for the obligatoriness of such movement from the behavior of parasitic gaps in configurations combining typical leftward A′-movement with heavy NP shift, in the context of the theory of parasitic gaps in Nissenbaum (2000).

c. *How much ... of split*
How much (of the chocolate cake) did you say (***of the chocolate cake**) that I ate (of the chocolate cake)?

d. *ago-stranding*
How long (ago) did you say (***ago**) that you went to France ($^?$ago)?

There is no additional movement across the CP edge in these examples that should force spec-CP to be evacuated in the way that we have just seen for vP. Thus these examples present a puzzle. Note that such patterns are a puzzle not only for the CL theory defended in the present chapter, but also for any theory that takes intermediate stranding to be derived by movement through phase edges.[24]

Kayne (2002) suggests that French examples like (76a) above do not involve extraction of *combien* ("how many"), but rather movement of a constituent evacuated by everything except for *combien*. (See Corver (2007) for more on such remnant movement derivations.) Under this analysis, apparent base position stranding of *de livres* ("of books") in (76a) actually is derived by movement of *de livres* to a low position in the clause (77a). Subsequent A′-movement of the phrase that *de livres* once occupied creates the appearance of *combien* having extracted (77b):[25]

(77) *Movement of "de livres" (a) followed by wh-movement of the remnant (b)*

a.

---
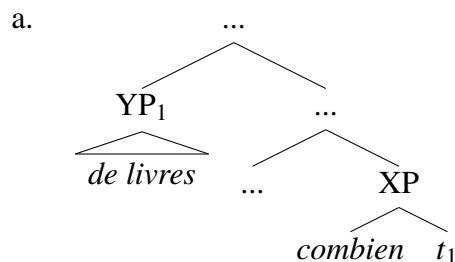
[24]The *of*-phrase in examples like (76c) seems to need to be rightmost (i). This fact is consistent with the displaced *of*-phrase having been derived by PP extraposition. This analysis provides a possible way for (76c) to be accounted for: if the *of*-phrase was never in fact stranded by movement because it can only be displaced by extraposition, then there is no expectation that stranding in spec-CP should be possible for this phrase.

(i)   How much did you put (*of the chocolate cake) in the fridge (of the chocolate cake)?

[25]Note that the derivation in (77) does not violate CL, provided that these movements occur within the same phase (presumably vP here). Prior to (77a), *combien* precedes *de livres*. The movement of *de livres* over *combien* in (77a) reverses their order, but the next movement in (77b) restores the original ordering of these elements, such that there is no basis for an ordering contradiction.

b.
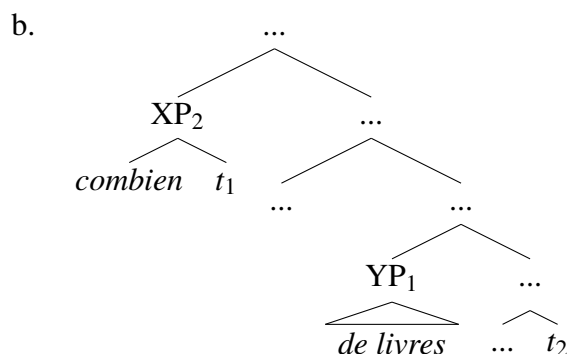
The displaced elements in (76) are plausibly non-constituents, or perhaps left branches that should be immobile under the Left Branch Condition (Ross 1967, a.o.). Since French generally obeys the Left Branch Condition, a remnant movement derivation is precisely what we would expect to be responsible for the exceptional displacement of *combien*, which is the only element in French for which the Left Branch Condition appears to be inapplicable. If a derivation along such lines is plausible for the examples in (76), a lack of intermediate stranding in spec-CP is expected of these patterns: under the remnant movement analysis, an element that appears to have been stranded in its base position in fact was not. Rather, it evacuated the moving phrase at an earlier point. Since stranding is not involved in such a derivation, there is no pied-piping/stranding at issue in the first place, and hence, no expectation that stranding in spec-CP should be possible.

5.3   On possible landing/stranding sites and phase theory

The previous two subsections have discussed scenarios where stranding in a particular edge fails. While the analyses I have offered may apply to some such patterns, they are unlikely to be correct for all, since the cross-linguistic variance in intermediate stranding patterns is quite rich. For instance, recall that in West Ulster English as reported by McCloskey (2000), *wh*-adjoined *all* can be stranded in spec-CP, but not spec-vP. Henry (2012) corroborates the existence of such a variety, which she terms West Derry City English. However, Henry shows that there is in fact great variation within West Ulster English dialects. For instance, Henry shows that two other varieties, which she refers to as South Derry English and East Derry English, permit what appears to be *all*-stranding in spec-vP:

(78)   *Spec-vP intermediate stranding in South Derry English*
     a.   What$_2$ did he [t_2 **all**]$_1$ do t_1 on holiday?
           (Henry 2012, ex. 25)

    b.   Where$_2$ does she [t_2 **all**]$_1$ see her students t_1?
           (Henry 2012, ex. 29)

(79)   *Spec-vP intermediate stranding in East Derry English*

    a.   What$_2$ did he [t_2 **all**]$_1$ do t_1 in Derry?
           (Henry 2012, ex. 52)

    b.   Who$_2$ did he [t_2 **all**]$_1$ say was elected t_1 in the council elections?
           (Henry 2012, ex. 56)

If the syntax of these varieties is basically the same as that of the one studied by McCloskey and of mainstream English, then the analysis of subsection 5.1 above incorrectly rules out such stranding. This problem also extends to the examples of spec-vP intermediate stranding we have seen above in Dutch (17) and Polish (28b, 29). Further variation within West Ulster English presents yet more puzzles. According to Henry, South Derry English allows base position *all*-stranding in addition to spec-vP *all*-stranding, but does not permit *all*-stranding in the CP edge. The Strabane variety is evidently even more restrictive, permitting *all*-stranding in only the base position. The least restrictive is East Derry English, which Henry shows allows *all*-stranding in the base position, spec-vP, and spec-CP.

    In response to these diverse facts, Henry argues that the set of positions in which a language permits stranding is simply a matter of choice. That is, while the syntactic principles endowed by Universal Grammar determine a consistent set of landing sites through which successive-cyclic movement passes, languages may opt to permit stranding in only some of them. While this could ultimately be the correct understanding, a more predictive theory is desirable if possible.[26]

    A different strategy is taken by Barbiers (2002), who analyzes stranding in a position consistent with spec-vP in Dutch, as described in section 2.4 above. As Barbiers shows, this stranding pattern is highly restricted: only stranding in the matrix spec-vP in a long-distance movement derivation is permitted. Stranding in the embedded spec-vP is banned, as is stranding in the embedded CP edge. Part of Barbiers' approach to these facts is to connect the distribution of stranding to the distribution of phases: if a constituent is not a phase, it is not a domain for successive-cyclic movement, and hence not a possible location for intermediate stranding. In general, relating the distribution of intermediate stranding and of

---

[26]McCloskey (2000) suggests that prosodic differences between West Ulster English and mainstream English can predict the fact that the latter does not permit *all*-stranding under A′-movement. If this is correct, it is conceivable that prosodic differences between dialects could also be responsible for some of the variation discussed here. More generally, it is possible that some of the cross-linguistic variation in intermediate stranding is due to interface factors of this sort: if a particular instance of syntactically licit stranding violates an independent phonological/prosodic constraint, that instance of stranding is expected to be blocked.

phases in this way has the potential to lead to a more predictive theory of whether or not a given language will permit stranding in a particular position.

A difficulty for this analytic direction is that it likely depends on positing considerable cross-linguistic variation in the set of phases. This is a complex issue, since the current literature offers many proposals about where phases can be found. As Ko (2014) points out, beyond the original phases vP and CP, most constituents in the clause have been taken to be a phase at some point, including VP (Fox & Pesetsky 2005a, Wiland 2010, Ko 2011), Asp(ect)P (Bobaljik and Wurmbrand 2013; Bošković 2014; Harwood 2015), and TP (Deal 2016; Wurmbrand 2013, 2017). Further, some works argue that phasehood can change during a derivation, for instance, in response to movement of a phase head (Den Dikken 2007; Gallego 2010; Alexiadou et al. 2014). While these possibilities can accommodate many different patterns of intermediate stranding,[27] pursuing a phase-centric approach to the cross-linguistic variance in stranding will require case-by-case analysis of each pattern and the language in which it is situated. Such a task is beyond the scope of this chapter. Regardless, the present chapter predicts that any instance of intermediate stranding must obey the ISG, an expectation that is confirmed by the known facts.

6   Against the labeling account in Bošković (2018)

As analyzed in this chapter, intermediate stranding is an instance of movement from a moved phrase: a first step of successive-cyclic movement pied-pipes material into a phase edge prior to a second step of movement stranding it there. As Bošković (2018) notes, movement from moved phrases has been shown to be quite restricted by many works. While such sub-extraction is largely banned in many languages, others permit it to some extent. Working within the phase theory in Chomsky (2000, 2001) and the labeling framework of Chomsky (2013), Bošković (2018) makes a prediction about when movement from moved phrases is allowed. While his results partially overlap with the ISG, they differ in several details, as we'll see next.

---

[27]For example, a way of approaching the cross-linguistic variance of intermediate stranding in the verbal domain would be to enrich the set of phases here. Legate (2014) argues for a voiceP distinct from vP, and that voiceP is a phase instead of vP. However, if work in Distributed Morphology (Halle and Marantz 1993, a.o.) is right that categorizing heads (v, n, a, etc.) are phase heads (Marvin 2003; Marantz 2007; Embick and Marantz 2008), then vP should be considered a phase as well. Furthermore, as just mentioned, several works propose that VP is a phase. If VP, vP, and voiceP are all potential phases, then there are several possible landing sites in the verbal domain based on which a variety of different stranding patterns might be derived. Movement within or from this domain could further restrict stranding in some cases, following the arguments of section 5.1 above.

6.1   Movement from moved phrases in Bošković (2018)

For Chomsky (2013), when a phrase XP is merged with another phrase YP, XP must move away unless Y agrees with XP such that the two share a common feature, based on which the mother of XP and YP is labeled. For this theory, successive-cyclic movement does not involve agreement and hence does not feed labeling, which is in part why successive-cyclically moving phrases must keep moving until an appropriate landing site is found.[28] Bošković (2018) argues that this system predicts a desirably restricted distribution of movement from moved phrases, when combined with two assumptions: that only phrases may move (Rackowski and Richards 2005; Legate 2014; Harwood 2015, a.o.),[29] and that un-labeled constituents cannot move. To illustrate how these concepts ban movement from moved phrases in some situations, let's first examine the ban on extraction from subjects, which holds in many languages (though not all), including English:

(80)   *No sub-extraction from subject*
        ?* I wonder [who$_1$ [friends of t_1] hired Mary].

Assuming that DPs are phases, and that the subject originates vP-internally, the unacceptability of (80) is derived as follows for Bošković (2018): First, prior to A-movement of the subject, the *wh*-phrase *who* does a first step of successive-cyclic movement to the edge of the subject DP that contains it. Because successive-cyclic movement does not feed labeling, this movement effectively de-labels the subject DP. This prevents it from moving to spec-TP, and hence the derivation fails:

---

[28]In particular, the claim is that successive-cyclically moved material must evacuate its intermediate positions in order to prevent a labeling failure. However, the possibility of intermediate stranding in positions created by successive-cyclic movement would appear to contradict such a theory.

[29]In the context of CL, the concept that only phases can move is by itself enough to predict the ISG. If all constituents that can move are phases, then by the time a given constituent has moved it will already necessarily have been constructed and spelled-out, and thus any material that has not already reached its left linear edge will be unable to extract from that constituent later on. I have not taken such an approach in this chapter for a few reasons. First, the concept that only phases can move is not a conclusion that straightforwardly emerges from independent evidence, but is rather a conjecture, and therefore should not be assumed unless necessary. Such a concept is not in fact necessary to predict the ISG, since this chapter's account has required no reference to it. Second, it is simply not a foregone conclusion that all mobile phrases are phases. The English PP can move, but as discussed extensively in this chapter, assuming that the English PP is a phase does not lead to correct predictions about the distribution of preposition stranding. Further, though DPs can certainly move, we will see that a straightforward consequence of the results of chapter 3 is that the English DP is not a phase—a conclusion with precedent in previous literature. While I make an attempt to unite the results of chapter 3 with the concept that the English DP is a phase, this is an issue of some complexity that this dissertation does not decisively resolve. Given that the phasehood of DP remains debatable, the concept that only phases can move is, in turn, not a trivial assumption.

(81) *Successive-cyclic movement within subject bleeds movement to spec-TP*

$[_{TP}$ * T $[_{vP}$ $[_{??P}$ *who* $[_{D'}$ friends of *t*]] v-V ...]]
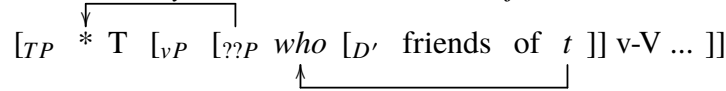
In general for this approach, successive-cyclic movement to the edge of any phase de-labels it, preventing it from moving and thus automatically precluding the possibility of deriving movement from a moved phase. Bošković (2018) argues that for this reason, movement from moved constituents cannot usually occur.

While specifiers formed by successive-cyclic movement yield the labeling problem just discussed, this issue should be irrelevant for specifiers that agree with the head of the phrase they merge to: such agreement should trigger labeling, and allow movement of the containing phrase. Bošković argues that this is correct. In particular, he argues that the ban on movement from moved phrases dissolves for specifiers that are externally merged in, and can remain in, the edge of a phase. This is because in the context of the labeling theory, any specifier that is able to remain in situ must have undergone agreement, or else it would have to move away.

Much of Bošković's supporting evidence for this claim comes from Serbo-Croatian. In this language the specifiers of the nominal phrase (and adjuncts, which Bošković assumes to be structurally equivalent to specifiers) agree with N in case and ϕ-features. Indeed, these elements can be extracted, as exemplified below with possessor extraction from a subject:

(82) *Possessor extraction from subject in Serbo-Croatian*
Jovanov$_2$ je $[_{NP}$ t_2 **prijatelj**]$_1$ vjerovatno t_1 otpustio Mariju.
John's.NOM is friend.NOM probably fired Mary.ACC
'John's friend probably fired Maria.'
(Bošković 2018, ex. 25c)

In general, Bošković (2018) makes the following prediction:

(83) *Prediction for movement from moved phrases in Bošković (2018)*
Movement from a moved phrase is possible only for a specifier that has agreed with the containing moved phrase.

Since specifiers are (at least in the basic case) linearized left of their sister, (83) predicts that left-adjoined phrases will be those that we see successfully extracting in scenarios of movement from a moved phrase. This prediction is thus partially overlapping with the ISG, since the ISG states that intermediate stranding is only possible when the extracted phrase was able to be initially linearized leftward of what it strands. The ISG and (83) differ in several respects, however.

6.2 Comparison with this chapter's analysis

One difference between (83) and the ISG is that the former is concerned only with extraction of specifiers, whereas the ISG is only concerned with extraction of leftward-linearized elements. Many of the scenarios discussed in section 2 above do not, prima facie, involve extraction of specifiers. Since mere word order is all that the ISG is defined in terms of, it thus describes the facts with less analytical commitments than (83).

More significantly, (83) predicts that movement from a moved phrase requires the extracted and stranded phrases to have an agreement relationship, while the ISG does not require this. Importantly, many of the intermediate stranding scenarios shown in section 2 do not involve any surface-evident agreement between the extracted phrase and stranded material, though it happens that some do, like quantifier stranding in Wolof. The ISG has a strong advantage on this issue if Preminger (2019) is right that there can be no agreement which is systematically morpho-phonologically null across its entire paradigm. The ISG is fully compatible with Preminger's results, since it has nothing to do with agreement. In contrast, Bošković (2018) frequently posits agreement where there is no independent evidence for it, given that the labeling theory requires it. This issue is relevant, for instance, to the examples of intermediate stranding in spec-vP in Dutch from Barbiers (2002), several of which we saw in section 2.4, exemplified once more below:

(84)   *Intermediate postposition stranding in Dutch*   [=(17a)]
        Waar$_2$ had jij   dan [$_{vP}$ [t_2 **mee**]$_1$ gedacht dat je   de vis t_1 zou
        where had you then   with   thought that you the fish   would
        moeten snijden]?
        must   cut
        'What had you thought to be forced to cut the fish with?'

Here an adposition that was inverted in the context of an R-pronoun is intermediately stranded. Bošković suggests that since R-pronouns and their concomitant P-inversion occur with a restricted set of elements, some agreement relationship must be involved in such Dutch configurations. Bošković cites van Riemsdijk (1997) for a notion of *R-feature* that might be applicable, but it is not obvious whether such a feature can really be equated with agreement. We also saw in section 2.4 that this Dutch pattern is not exclusive to postpositions like *mee* in (84). For instance, in (17b) above the phrase *voor ball* ("for ball") is stranded, and in (17c) *maar* ("only") is stranded. Bošković's proposal requires claiming that these examples of stranding also involve agreement, which there is no independent evidence for.

This issue also arises for Serbo-Croatian. Bošković shows that this language allows an intensifier to be extracted from a scrambled adjective, as in (85) below:

(85)   *Intensifier extraction from scrambled adjective in Serbo-Croatian*
        Izuzetno$_2$ su [$_{AP}$ t_2 skup]$_1$ kupili [t_1 automobil].
        Extremely are   expensive bought   car

        'They bought an extremely expensive car.'
        (Bošković 2018, ex. 30)

The prediction in (83) requires Bošković to assume that there is agreement between
the intensifier and adjective, since otherwise, this movement from the adjectival
phrase should be impossible. Bošković notes that the intensifier can remain in situ
in the adjectival phrase, which in the context of the labeling theory, implies that
label-facilitating agreement occurred. There is, however, no direct evidence for such
agreement in Serbo-Croatian. In contrast, notice that (85) obeys the ISG, given that
the intensifier's base position is leftward of the adjective that it strands (though this
example is of multiple scrambling rather than stranding per se).

    Finally, it is unclear how Bošković (2018) would account for the movement from
a moved phrase in spec-CP shown in (53) above, repeated below:

(86)   *Extraction from a DP in spec-CP*   [=(53)]
        ? [What student]$_2$ did Ann ask [$_{CP}$ [what picture of t_2]$_1$ to put up t_1]?
        (Sauerland 1999, 24c)

Bošković briefly mentions such examples and assumes that they are unacceptable,
but as we've seen, there are multiple works reporting the possibility of such sen-
tences. As described above in section 4.3.1, such examples are correctly predicted
to be possible under the proposals of the present chapter.

    A similar counterexample to Bošković (2018) comes from Zyman (2019), who
reports that extraction from subjects in languages like English and French is not
altogether banned, but becomes improved when the right sort of material intervenes
between the subject and extracted element. If correct, such facts also favor the
present chapter over Bošković (2018), since only the latter predicts an outright ban
on extraction from subjects. The difficulty of extraction from subjects is often
interpreted as an instance of freezing (Corver 2017, a.o.). Unlike Bošković (2018),
the present chapter doesn't propose a general cause for freezing: CL by itself does not
make any commitments about freezing, but this chapter argues that CL does capture
a particular instance of freezing when combined with independent constraints on
movement. Overall, the approach taken in this chapter expects effects like freezing to
emerge not from phase theory itself, but rather from the interaction of phase theory
with independent syntactic constraints. The freezing effect of criterial positions
is another independent factor that may be responsible for the island-hood of some
moved phrases (Epstein 1992; Rizzi 2006; Corver 2017, a.o.).

7 Conclusion

This chapter has proposed a generalization about word order in configurations with stranding in an intermediate position, and argued that this generalization emerges from the CL theory of phases, in combination with several independently supported constraints on the locality of movement.[30]

(87) *Intermediate Stranding Generalization (ISG)*
      Leftward movement of a phrase α can only intermediately strand an element β if β is (or can be) ordered to the right of α before stranding occurs.

Furthermore, this chapter has considered how the set of positions available for intermediate stranding might be determined, though predicting the space of cross-linguistic variance in where such stranding can occur remains a task for future work. This concern is related to but logically separate from the ISG, however, which has so far proven to be robust. The general mechanisms that I have argued here derive the ISG will serve as the backdrop for explaining the intricacies of another circumstance involving pied-piping and stranding that the next chapter focuses on—possessor extraction in colloquial English.

---

[30]The understanding of the ISG argued for in this chapter entails another consequence not mentioned above: that successive cyclic movement should only occur at phase edges. This is in contrast to a theory in which movement paths can form intermediate landing sites in any phrase they cross. If successive cyclic movement through non-phases were possible, then it should be possible for a moving phrase to intermediately strand a preposition, for instance, by pied-piping the preposition into the edge of a non-phase before then moving on into a phase edge, as in (i):

(i) $[_{ZP[Phase]}$ WH$_2$ Z $[_{YP}$ [P t_2]$_1$ Y $[_{XP}$ t_1 X]]]

The second step of movement in (i), which strands the preposition, both occurs before the linearization of the containing phase and does not violate the ban on phrase-bounded edge-internal movement. For this reason, such an intermediate stranding derivation should be fully licit, if it is possible. The fact that intermediate preposition stranding is (as far as I know) wholly banned indicates that such intermediate movement through a non-phase is not available—a conclusion supported by the general cross-linguistic robustness of the ISG.

Chapter 3

Possessor extraction in English

1 Introduction

This chapter examines a case of *possessor extraction* (PE), the A′-movement of a possessor from the possessed nominal phrase. This instance of PE, which emerges from a little-explored corner of colloquial English, shows a great deal of stable intricacy which will lead us to a detailed understanding of the concepts defended in the previous chapter. Thus this construction, despite being somewhat peculiar at first glance, will provide a rich body of evidence for the framework that I develop throughout this dissertation.

For many English speakers, A′-movement of a possessor requires pied-piping of the containing possessum, as (1) shows. Such pied-piping is often thought to be the only possibility for English.

(1)   *Standard English possessum pied-piping*
        Mary is the author $[_{CP}$ [**whose new book**]$_1$ they said $[_{CP}$ t_1 is good]].

This view is challenged by examples like (2) below, which are the subject of this chapter. In (2) we see an equivalent of (1) available in the colloquial language of some speakers, in which the possessor extracts, stranding the Saxon genitive morpheme [*'s*] as well as the rest of the possessum in an embedded clause. This initial English PE example is appropriately marked with "%", as PE is not available to all speakers, though I omit this from subsequent examples.

(2)   *PE in English*
        % Mary is the author $[_{CP}$ **who**$_1$ they said $[_{CP}$ $[t_1$**'s new book**] is good]].

In (2), [*'s*] becomes phonologically dependent on the verb *said* in the absence of the moved possessor. It is easy to see that this "-s" really must be a stranded Saxon

genitive [*'s*]. The past tense and plural subject of the relative clause in (2) where PE is taking place eliminate the possibility of this element being subject agreement. The fact that the possessor is female also removes any possibility of this being a reduced resumptive *his*.

Though not all native speakers accept such PE examples, the present study of 34 speakers, mostly residents of the Boston area, resulted in 19 judging PE of this form to be acceptable.[1] Those who accept the construction frequently note that it is distinctly colloquial, a fact that may contribute to its rarity in written form.[2] Some speakers are aware that they are capable of English PE, while many others are surprised to notice its acceptability when it is brought to their attention. While PE has been well established in various languages[3] like Hungarian (3), the possibility of PE in English has received little attention.

(3)   *Hungarian PE*
       Ki-nek$_1$ ismer-té-tek   [a t_1 vendég-é-∅-t   ]?
       who-DAT know-PST-2PL [the   guest-POSS-3SG-ACC]
       'Whose guest did you know?' [Lit: 'Whose did you know guest?']
       (Szabolcsi 1984, ex. 14)

---

[1] My informants are mostly American, though my set of speakers who accept PE includes two Canadian, one Australian, and one British individual. There is no clear generalization about the age/origin/background of speakers who allow PE. I suspect that my result that 19/34 speakers accept PE is somewhat inflated due the fact that those capable of it often made an effort to seek me out, whereas those who do not accept it unsurprisingly did not do so.

The sentences reported here are the aggregate of notes taken from interviews during 2016-2018. Since these PE sentences are most suited to colloquial language, I typically interviewed individuals verbally face-to-face. Generally, I constructed some sample PE sentences and asked that they be judged, though individuals capable of the phenomenon often readily volunteered more.

[2] English PE can be found in informal writing, as the following examples from the Internet show:

(i)   a.   She raised her eyebrows while her other brunette friend, **who I heard's name** is Caroline...
             (https://www.quotev.com/story/5110453/THE-GREAT-McCANN/23)
       b.   ...the rizinosaurus, **who you said's major downfall** would be it's size...
             (http://www.topix.com/forum/science/dinosaurs/TAIDJ8LEBGL3O0D5I/p2)
       c.   So **who do you think's car** it is.
             (https://www.wattpad.com/133087986-stranger-c-d-2)
       d.   Noelle has helped me in the past, along with another woman **who I believe's name** is Rosie.
             (https://www.dbchocolate.com/Hazelnut-Truffles_p_835.html)

[3] In addition to the Finno-Ugric Hungarian, PE is attested in Austronesian languages like Chamorro (Chung 1991), Mayan languages such as Tzotzil (Aissen 1996) and Chol (Coon 2009), as well as much of the Slavic family (Ross 1967; Bošković 2005, a.o.). Romance and Germanic languages permit some extraction of post-nominal/PP possessors (Gavruseva 2000).

The only literature I know to have considered PE in English is Gavruseva and Thornton (2001), discussed below.[4] While I argue that English PE is true movement, we will see that its distribution is quite restricted, which plausibly contributes to the fact that it has not been widely noticed. An analysis of this construction's restrictions and their consequences for syntactic theory is the focus of this chapter.

1.1 Background

The possibility of this form of PE contrasts with the known impossibility of extracting *whose* (which I take to be *who* + [*'s*]), or any possessor marked with [*'s*] in English:

(4) *No extraction of whose*

    a.   * Mary is the author [$_{CP}$ **whose**$_1$ they said [[t_1 **new book**] is good]].

    b.   * **Whose**$_1$ did you say we should buy [t_1 **cookies**]?

If a possessor DP is the specifier of a possessive D whose exponent is [*'s*] in English (Abney 1987; Corver 1992; Chomsky 1995; Munn 1995, a.o.), then *whose* and elements like it are not syntactic constituents, and therefore their immobility is expected. However, the specifier of [*'s*] is surely a phrase, which as such is in principle extractable:[5]

(5) *A structure for possessive DPs headed by* [*'s*] *(where DP$_{[Poss]}$ = possessor)*

DP

DP$_{[Poss]}$ D   NP

            [*'s*]
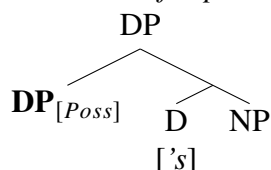
While the possessor DP in the specifier of [*'s*] can indeed be moved for some speakers as (2) above showed, the marking "%" on (2) reminds us that for many English speakers, such movement is not possible.

    Various works attribute the typical illicitness of PE in English to a requirement of PF which rejects movement that separates a possessor from the possessive D

---

[4]The only mention of this construction in other work that I am aware of comes from Heck (2009), who in footnote 64 credits Andrew McIntyre for the observation of the sentence "*a person who I think's reputation would be better if they stopped dribbling incessantly*". Heck assumes that this sentence is formed by placing a parenthetical in DP, an analysis that the present chapter argues against in section 3 below.

[5]Adger (2003) argues that possessive [*'s*] is in fact genitive case marking assigned to a DP that inhabits the specifier of a null possessive D. The strandability of possessive [*'s*] that this chapter reports mandates in favor of the view that this element is indeed a head in its own right, though as Adger discusses, its distribution is without a doubt complex.

(Chomsky 1995; Radford 1997; Gavruseva 2000; Gavruseva and Thornton 2001, a.o.). Indeed, Gavruseva argues that PF adjacency requirements of this variety play an important role in constraining PE cross-linguistically. In this chapter, I accept this general line of explanation for those English speakers who reject PE. However, I argue that the nature of PE in English does not indicate that the speakers who permit it lack such an adjacency requirement, but rather shows that such speakers can satisfy this requirement more locally. This analysis accurately predicts a constrained form of PE in English, as we'll see.

Gavruseva and Thornton's (2001) investigation of PE focused on *whose*-questions in child speech, where PE is quite frequent in long-distance questions. Gavruseva and Thornton argue that PE is possible in child English because children have not yet acquired the PF constraints that require pied-piping, and consequently block PE. This perspective on the acquisition path leads us to expect a total lack of PE in a mature English grammar. However, in a control study on adults reported in the same work, Gavruseva and Thornton (pg. 255) observed PE in adult speech. 11% of their adult data comprises PE of the form shown in (2) above.[6] The authors suggest that this 11% is the result of production errors. However, almost all such instances of PE gathered in their adult study were from two speakers, Cristy and Kath. Cristy produced PE about half as often as pied-piping, while Kath produced PE even more often. These speakers evidently have PE as a productive option. Indeed, in this work I argue that PE is a reality for some English speakers, though subject to restrictions, several of which Gavruseva and Thornton also identified. The explanation of these constraints and their significance comprises the core of this chapter.

1.2 Results in preview

I argue that English PE obeys the following generalization, which subsumes an array of restrictions on this construction that will be demonstrated in the next section:

(6)   *Generalization about English PE*
        A possessor must reach the left linear edge of the local CP before extracting from the possessum DP.

A particularly clear manifestation of this generalization is evident with PE from non-subjects, exemplified in (7) below. This example shows that PE from an object possessum is unacceptable if this possessum is stranded in its base position (7a). Rather, the possessum must be displaced to the edge of the local CP (7b).

---

[6]A very small percentage of their adult data is comprised of other unusual extraction configurations (for instance, movement of *whose*) which I take to be genuine errors.

(7)   *Displacement of non-subject possessum under PE*
    a.   * **Who**$_1$ do they think [$_{CP}$ Mary read [t_1**'s book**]]?
    b.   ✓ **Who**$_2$ do they think [$_{CP}$ [t_2**'s book**]$_1$ Mary read t_1]?

I argue that the generalization in (6), which describes (7) and related restrictions, is a natural result of the interaction of two concepts: the CL theory of spell-out, and a version of the independently proposed requirement that the possessor and possessive D [*'s*] be adjacent at PF. In particular, I argue that while English speakers that reject PE enforce adjacency between possessor and [*'s*] absolutely, in contrast, speakers who permit PE can satisfy this requirement locally, as (8) describes:

(8)   *Possessor-['s] Adjacency (Local version, available to PE speakers)*
        [*'s*] must be adjacent to its associated possessor at the spell-out of the minimal phase (vP, CP) containing [*'s*].

The interaction of CL with (8) will be vital to this chapter's explanation of the intricate properties of English PE. Under this account, CL is a part of the grammar of both speakers who permit PE and those who do not. The difference between these two groups lies in how they enforce a PF condition. This understanding maintains a uniform syntax, with variation attributed to the PF interface.

1.3   Chapter contents

Section 2 describes the facts about English PE, which section 3 argues is true extraction. Section 4 overviews the concepts used in section 5 to build an account of this phenomenon, which as we'll see, yields new evidence for the CL approach. Section 6 turns to a PE-based account of the under-analyzed *that's*-relative construction in English (Seppänen and Kjellmer 1995; McDaniel et al. 1998, 2002), which supports the proposal of Deal (2016) that TP is a phase in relative clauses. Section 7 considers what this analysis indicates about the phasehood of DP, an issue of some complexity. Section 8 discusses some residual considerations for this analysis of English PE, and is followed by the conclusion in section 9.

2   The restricted distribution of English PE

This section describes the facts about English PE. This exposition involves some preliminary analysis, leading to the generalization in (6) above, which this chapter will derive in section 5 below.

Gavruseva and Thornton's study of PE in child speech focuses on questions, but English PE is possible in any A′-movement context, as (9) shows:

(9)   *The generality of English PE*

      a.   *Matrix question*
           Who$_1$ do you think [[t_1**'s kid**] ate the most cake]?

      b.   *Embedded question*
           I can't remember [**who**$_1$ I said [[t_1**'s friend**] is coming over]].

      c.   *Relative clause*
           This is the student [**who**$_1$ they suspect [[t_1**'s answers**] were copied]].

      d.   *Free relative*[7]
           I'll speak to [**whoever**$_1$ you suggest [[t_1**'s idea**] is the best]].

      e.   *Cleft*
           It's Michelle [**who**$_1$ we heard [[t_1**'s cat**] is the cutest]].

      f.   *Topic / focus movement*
           John's life is certainly boring, but let me tell you about my cousin Jim.
           Now [**this guy**]$_1$, I think [[t_1**'s story**] will entertain you].

Most of the PE examples shown so far have involved extraction of *who*. Other possessors can be extracted also, as (10) below shows.[8] Extraction of phonologically larger possessors is often judged as more difficult. For maximal clarity of judgments, most of the PE sentences reported in this chapter involve extraction of *who*.

---

[7]The possibility of forming a free relative by PE contrasts with the reported incompatibility of free relatives with pied-piping *wh*-movement (Horvath 2006), though this topic is more involved than can be usefully addressed in this chapter.

[8]The possibility of examples like (10c) below is interesting in light of the fact that *which* cannot occur in possessor position in non-extraction circumstances:

(i)   * This is the shop **which's** pastries are tastiest.

In (10c), extraction of *which* evidently avoids whatever causes the degradation of (i). Other *wh*-possessors that cannot be selected by [*'s*] are also comparatively acceptable when extracted:

(ii)   a.   * I don't remember **where's** prices you said are lowest.
        b.   ? I don't remember **where** you said**'s** prices are lowest.

(iii)   a.   * Tell me **what's** flavor you think is weird.
        b.   ? Tell me **what** you think**'s** flavor is weird.

This amelioration by movement is reminiscent of work by Ott (2011), who discusses certain patterns involving split topicalization and quantifier float in Germanic, which are puzzling in that they have the characteristics of A′-movement despite there being no obvious prior stage of the derivation from which they could have been formed. Ott argues that such circumstances arise when external merge creates constituents which are unstable due to being unable to be labeled, and hence cannot remain as they are, but rather must be rendered discontinuous in order for labeling to succeed. If the impossibility of strings like *which's*, *where's* and *what's* can be attributed to a labeling issue, then Ott's findings present a potential route of explanation for this fact.

(10)   *Extraction of other possessors*

     a.   [**Which student**]$_1$ did he claim [[t_1's idea] was stolen]?

     b.   [**How many people**]$_1$ do you think [[t_1's files] fell off the shelf]?

     c.   ? I went to the place [**which**$_1$ she said [[t_1's pastries] are tastiest]].

Further examination of this construction reveals numerous restrictions. First, notice that all the English PE examples shown so far have been multi-clausal. This is not a coincidence. As (11) shows, English PE is banned for clause-bound movement:

(11)   *PE unavailable for clause-bound movement*

     a.   * **Who**$_1$ will [t_1's friend] arrive tomorrow?

     b.   * John is the one [**who**$_1$ you stole [t_1's pie]]!

     c.   * Tell me [**who**$_1$ they reviewed [t_1's book].

Additionally, most of the English PE examples presented so far show extraction from a subject. As previewed in (7) above, non-subject[9] possessum DPs exited by PE must be displaced to the edge of the local CP, as (12) shows once again:

(12)   *Displacement of non-subject possessum under PE*

     a.   * **Who**$_1$ do they think [$_{CP}$ Sue found [t_1's cat] today]?

     b.   ✓ **Who**$_2$ do they think [$_{CP}$ [t_2's cat]$_1$ Sue found t_1 today]?

     c.   * Mary is the person [**who**$_1$ I heard [$_{CP}$ John ate [t_1's food]]].

     d.   ✓ Mary is the person [**who**$_2$ I heard [$_{CP}$ [t_2's food]$_1$ John ate t_1]].

     e.   * Tell me [**who**$_1$ they said [$_{CP}$ we should meet [t_1's friend]]].

     f.   ✓ Tell me [**who**$_2$ they said [$_{CP}$ [t_2's friend]$_1$ we should meet t_1]].

Example (12) demonstrates the obligatoriness of this displacement for PE from the object of a transitive clause, but this requirement importantly holds for PE from any non-subject. Examples (13-14) below show that this displacement is also required for PE from either non-subject argument of a ditransitive:

(13)   *PE from direct object of a ditransitive*

     a.   **Who**$_2$ do they think [[t_2's book]$_1$ we should give Mary t_1]?

     b.   **Who**$_2$ do they think [[t_2's book]$_1$ we should give t_1 to Mary]?

(14)   *PE from indirect object of a ditransitive*[10]

     a.   **Who**$_2$ do they think [[t_2's kid]$_1$ we should give t_1 the prize]?

     b.   **Who**$_2$ do they think [[t_2's kid]$_1$ we should give the prize to t_1]?

---

[9]I use the term "non-subject" to refer to all DPs whose base position prior to A′-movement is not spec-TP, but a lower, post-verbal position.

[10]Though A′-movement of indirect objects is independently ruled out for some speakers.

Example (15) below shows that the same holds for expletive associates. Such arguments are post-verbal by default, though under PE they must end up at the edge of CP, as in (15c). This example is marked, but clearly improves on (15b), which lacks the needed displacement.[11]

(15)   *PE from expletive associate*
     a.   Mary said [there were [someone's books] on the table].
     b.   * **Who**$_1$ did Mary say [there were [t_1**'s books**] on the table]?
     c.   ? **Who**$_2$ did Mary say [[t_2**'s books**]$_1$ there were t_1 on the table]?

Gavruseva and Thornton noticed the obligatory displacement of non-subject possessum DPs under PE (specifically of objects) in their study as well, in both children and adults. They hypothesize that this is caused by the possessor's movement pied-piping the possessum to the specifier of the embedded CP, and then stranding it there by subsequent extraction. This is the view that the present chapter defends. If this analysis is correct (as argued further in section 8.1) then such stranding provides evidence that movement from CPs successive-cyclically passes through their edge, joining similar arguments from previous literature on stranding in mainstream English (Urban 1999), West Ulster English (McCloskey 2000), and a variety of other languages which the previous chapter discussed. Work in this vein identifies elements that can be stranded in a CP edge, as well as in their base position. We have just seen, however, that English PE from a non-subject possessum does not allow the possessum to be stranded in its base position (in contrast to PE in Hungarian, as shown in (3) above). This fact suggests that English PE is more complex than typical cases of stranding. This chapter provides an explanation for why this is so.

2.1 The possessor extracts from DP via the linear edge of CP

We've seen that PE from non-subject possessum DPs requires displacement of the possessum to the edge of the local CP. At first glance, this fact suggests that PE is only possible from the structurally highest DP in the clause. Before any A′-movement, this is whatever DP ends up in spec-TP. If the DP exited by PE is not in spec-TP, it consequently must be pied-piped to spec-CP with the moving possessor prior to PE. This description is consistent with what has been shown so far.

---

[11]A reviewer for a previous version of this work notes that (15c) may be degraded because a phrase's possessor influences its ability to be an expletive associate: since *who* is not independently a licit *there*-associate (i), a phrase originally containing *who* may be similarly degraded in this position. As far as I know this is true of *wh*-phrases generally.

(i)   Which teacher said there is a student/*who in the room?

If this were a sufficient description, however, PE of a post-nominal possessor from a subject should be acceptable, contrary to fact (16):

(16)   *PE unavailable for post-nominal possessors*
    a.   * **Who**$_1$ did you say [$_{CP}$ [**a cookie recipe of** t_1**'s**] is getting popular]?
    b.   * That's the senator [**who**$_1$ they think [$_{CP}$ [**a friend of** t_1**'s**] reported his corruption]].

We have already seen examples of PE from subjects, so there is no benefit to attributing the deviance of (16) to the known fact that extraction from subjects is frequently degraded. The examples in (16) improve if [*'s*] is absent (to the extent that extraction from indefinite subjects is relatively tolerable), indicating that this morpheme's requirements are influential in constraining PE, as this chapter argues.

Notice that in (16), overt material within the possessum intervenes between the trace of PE at the right side of this DP, and the left edge of the local CP. The fact that PE in this situation is unacceptable suggests the generalization in (17a) below, graphically represented in (17b), which I'll show to be correct:

(17)   *Generalization about English PE*   [=(6)]
    a.   A possessor must reach the left linear edge of the local CP before extracting from the possessum DP.

    b.   [$_{CP2}$ DP$_{[Poss]}$ ... [$_{CP1}$ (*α) [$_{DP}$ (*α) t 's NP]]]

Given (17), we expect the unacceptability of (16), since here there is material separating the trace of PE from the left linear edge of the embedded CP. Additional facts about English PE shown in the remainder of this section have the same explanation.

Before moving on to the remaining facts, notice that the generalization in (17) is consistent with the displacement of non-subject possessum DPs under PE that we saw earlier in this section: without such displacement prior to PE, (17) could not be met in such configurations. This generalization also clarifies the impossibility of PE for clause-bound movement, shown in (11) above: if (17) holds, PE cannot become evident unless the possessor exits a possessum in an embedded CP edge.

2.1.1 Preposition stranding and PE

PE is not possible from a DP within a PP (18b/d), unless that PP is stranded in its base position prior to PE (18a/c). As with other non-subject DPs, PE from the DPs in (18) requires them to be displaced to the edge of the local CP, presumably via pied-piping with the possessor's movement as discussed above. These DPs originate inside of PPs, and even though P may typically be pied-piped along with A′-movement of its

DP complement in English, in (18) only P-stranding is permitted.

(18)   *Pied-piped P with displaced non-subject blocks PE*

        a.   Who$_2$ do they think $[[t_2$'s house$]_1$ we should leave **from** $t_1]$?

        b.   * Who$_2$ do they think $[[$**from** t_2's house$]_1$ we should leave $t_1]$?

        c.   Who$_2$ did they say $[[t_2$'s cat$]_1$ we should give the prize **to** $t_1]$?

        d.   * Who$_2$ did they say $[[$**to** t_2's cat$]_1$ we should give the prize $t_1]$?

This is what we expect given (17): if P is pied-piped along with movement of the non-subject possessum to the CP edge, then P illegally intervenes between the left edge of the embedded CP and the trace of PE within the possessum DP.

2.1.2   Complementizers and PE

The distribution of complementizers in PE derivations also fits the linear generalization in (17). Recall that in English, long-distance *wh*-movement of non-subject DPs is compatible with an overt complementizer *that* in the embedded clause:

(19)   *Overt complementizer* that *with non-subject extraction*
        Who$_1$ do they think $[_{CP}$ (**that**) Mary likes $t_1]$?

Subject extraction, however, is not compatible with an overt complementizer *that*, a phenomenon well-known as the *that*-trace effect:

(20)   *Appearance of* that-*trace effect with subject extraction*
        Who$_1$ do they think $[_{CP}$ (***that**) t_1 likes Mary]?

PE from a subject is also incompatible with the complementizer *that* (21). This fact is interesting because here we have extraction from, but not (cross-clausal) movement of, the subject. Thus this fact does not necessarily instantiate the *that*-trace effect:[12]

(21)   *Overt C* that *disallowed with PE from subject*

        a.   Who$_1$ do you think [(***that**) $[t_1$'s friend] always laughs at bad jokes]?

        b.   Who$_1$ did they say [(***that**) $[t_1$'s window] broke during the storm]?

In (21), the presence of the C *that* to the left of the possessum subject separates the trace of PE within this DP from the left linear edge of the local CP. Thus the unacceptability of the complementizer here is expected, since this violates (17).

---

[12]This fact is compatible with accounts of the *that*-trace effect as a linear filter on an overt C *that* adjacent to a trace (Bresnan 1972, a.o.). I do not analyze complementizer-trace effects in this chapter, since the account presented here predicts (21) automatically.

Non-subject DPs exited by PE also cannot be preceded by an overt C *that* (22). However, if the possessum in such configurations is stranded in spec-CP as discussed earlier in this section, then C should be unable to precede it anyway, independent of the fact that a complementizer in this position would also violate (17). An overt C *that* to the right of the possessum here is also impossible, as expected given the Doubly-Filled Comp Filter (Chomsky and Lasnik 1977).

(22)   *Overt C* that *disallowed with PE from a pied-piped non-subject possessum*
    a.   Who$_2$ did you say [(**that**) [t_2's cat]$_1$ (*that) John saw t_1]?
    b.   Who$_2$ does he think [(**that**) [t_2's cat]$_1$ (*that) he wants to take home t_1]?

2.1.3   Adverbs and PE

Typically, high adverbs may occur on either side of the English subject:

(23)   *Variable high adverb position*
    a.   (Fortunately) Mary (fortunately) has money.
    b.   (Usually) John (usually) eats a burrito for lunch.

PE from a subject is incompatible with such an adverb appearing on the left side of that subject, but is possible with the adverb to its right:

(24)   *An adverb cannot precede a subject exited by PE*
    Who$_1$ did you say [(**usually/*fortunately**) [t_1's friend] (usually/fortunately) has an extra taco]?

This is expected if (17) holds, since an adverb to the left of the possessum subject linearly intervenes between the trace of PE in this DP and the edge of CP. PE from a non-subject, involving stranding in the CP edge as we've seen, behaves the same:

(25)   *An adverb cannot precede a non-subject exited by PE*
    Who$_2$ did you say [(**allegedly/*fortunately**) [t_2's cat]$_1$ (allegedly/fortunately) John found t_1]?

2.2   The puzzle

This section has shown that English PE is subject to the generalization in (17), which describes the fact that no overt material may intervene between the trace of PE within DP, and the left linear edge of the local CP. This chapter will argue that this generalization emerges from the interaction of CL, and an adjacency requirement of [*'s*] that PE speakers can satisfy at a local (phase-bounded) level of the derivation.

2.2.1 Against a discriminating [*'s*]

Some of the unacceptable examples of English PE shown in this section might, at first glance, be attributed to restrictions on what [*'s*] can cliticize to. Upon closer inspection, however, it is unclear how to state what the relevant restrictions would be. We have seen that [*'s*] can attach to verbs in PE contexts (as in example 2, and many more) but not adverbs (24, 25), or functional heads like P (18) or C (21, 22).[13] These facts do not yield an obvious generalization about what [*'s*] may attach to in PE derivations. It is also not obvious why [*'s*] would be discriminating in PE contexts, even though in typical non-PE contexts it is not selective, and can cliticize to adverbs and functional heads, among other elements (Zwicky 1987).

(26)   *Typical indiscriminating cliticization of* ['s]
    a.   [the person you're talking **to**]'s jacket
    b.   [the man who left **yesterday**]'s book

For an account along such lines, it would also remain puzzling that [*'s*] can cliticize onto verbs in some PE contexts, but not in those like (7a) where an object possessum is stranded in its base position, which as we've seen are unacceptable.

Given these issues, this chapter accounts for the facts about English PE without positing cliticization restrictions on [*'s*]. With this alternative hypothesis set aside, the next section makes the case that English PE is truly movement, setting the stage for the core analysis of this chapter.

3   English PE is true extraction

Recall that English PE only occurs in long-distance A′-movement contexts, unlike PE in languages like Hungarian. This fact might be thought to show that English PE is an illusion created by a DP-internal parenthetical clause, between the possessor and [*'s*]. This parenthetical would make the resulting construction always appear multi-clausal, since it contributes an additional verb to the surface string. For instance, the initial PE example in (2) could be true extraction in a bi-clausal context (27a), or a single clause with a parenthetical *they said* in the possessed DP (27b):

---

[13]PE in ditransitives can also yield strings where [*'s*] is adjacent to a pronoun:

(i)   Who$_1$ did they tell **me/you/**$^{??}$**your friend** [[t_1 's name] is Bill?

As this example shows, a lexical noun in the same position is less acceptable. I suspect that this is attributable to a garden path effect.

(27) *String:* Mary is the author who they said's new book is good

    a. *Extraction analysis*
       Mary is the author [$_{CP2}$ who$_1$ they said [$_{CP1}$ [$_{DP}$ t_1's new book] is good]].

    b. *Parenthetical analysis*
       Mary is the author [$_{CP}$ [$_{DP}$ who (**they said**)'s new book] is good].

As (28) below shows, parentheticals are not independently permitted in this DP-internal position (Emonds 1976, a.o.), weakening the parenthetical analysis of PE:

(28) *Parentheticals are not permitted between DP$_{[Poss]}$ and* ['s]

    a. [$_{DP}$ John (***in my opinion/*of course**)'s idea] is funny.
    b. [$_{DP}$ A friend of the teacher (***Mary said/*in fact**)'s] visited yesterday.

Similarly, if the strings which appear to involve PE are parenthetical-derived, the same configurations should be possible not only with parentheticals containing an embedding verb like *they said* in (28b) above, but also for verb-less parentheticals, two of which (28) also shows. As expected if PE is true extraction without any parenthetical involved, placing such a parenthetical in the relevant position in a PE sentence results in unacceptability:

(29) *Apparent PE impossible with verb-less parentheticals*

    a. Mary is the author who (***of course**)'s new book is good.
    b. Mary is the author who (***in fact**)'s new book is good.

Even if these arguments were not conclusive on their own, further evidence shows that the parenthetical analysis of English PE is inferior to the movement analysis.

3.1 Failures of parenthetical subtraction

Parentheticals are optionally inserted into what are otherwise well-formed sentences. Therefore if PE constructions in fact involve parentheticals, subtracting the content that is supposedly parenthetical should yield a licit configuration. In reality, this subtraction test often fails in PE sentences, revealing that no parenthetical is involved.

    Consider the PE question in (30) below. Subtraction of the supposed parenthetical here yields an impossible string, whether or not the auxiliary *do* is counted as part of the parenthetical:[14]

---

[14]Thought strictly speaking it should be be counted, since the auxiliary is required for a parenthetical in a question:

(i) Whose book, (**do you think / *you think**), did Mary buy?

(30)   *Failed parenthetical subtraction:* Who do they think's cat he saw?
    a.   Who **do they think**'s cat he saw? → * Whose cat he saw?
    b.   Who do **they think**'s cat he saw? → * Who do's cat he saw?

Similar facts become evident when we consider the phenomenon of "Free Deletion in Comp" (Chomsky and Lasnik 1977), which can derive examples like (31), where the *wh*-operator responsible for relativization is optionally silent:

(31)   *Relativization with null operator*
    a.   the one [$who_1/\varnothing_1$ I like t_1]
    b.   the cat [$which_1/\varnothing_1$ I saw t_1]

Relativizing PE with no overt *wh*-operator is also possible (32).[15] Removal of the supposed parenthetical material from such sentences yields an unacceptable result:[16]

(32)   *PE of null relativizing* wh-*operator*
    a.   the person [$\varnothing_{WH1}$ **I said** [[t_1's cat] meows too much]]
        → * The person['s cat meows too much]
    b.   the person [$\varnothing_{WH1}$ **I said** [[t_1's cat]$_2$ you saw t_2]]
        → * The person['s cat you saw]
    c.   the person [$\varnothing_{WH1}$ **I think** [[t_1's friend]$_2$ you talked to t_2]]
        → * The person['s friend you talked to]

    A similar effect is apparent with control. Example (33) below shows a PE sentence in which the subject of the supposed parenthetical controls into a *before* adjunct. Removal of the supposed parenthetical yields an unacceptable sentence where PRO is un-controlled:

(33)   *Adjunct control with PE*
    [Which author]$_2$ **did you**$_1$ **say** [[t_2's book] looked good] [before **PRO**$_1$ ordering it]?

The hypothetical form in (30b) is nevertheless attempted for the sake of thoroughness.

[15]These PE examples with a null relative operator behave the same as PE examples with an overt operator, including, as we see in (32b/c), the property of pied-piping non-subject possessum DPs to the local CP edge. The analysis in section 5 will attribute this pied-piping to the interaction of CL and a requirement of adjacency between [*'s*] and the possessor, but such adjacency is not obviously relevant when, as in (32), the possessor is covert. I hypothesize that the possessor here was originally overt, before later being removed from the surface string via Free Deletion in Comp. See chapter 8 for further discussion of this and analogous facts involving movement of covert phrases.

    The apparent relevance of null operators to what I argue to be linearization-sensitive phenomena will emerge more than once in this dissertation. This issue will be addressed directly in chapter 8.

[16]Though the post-subtraction strings in (32) are acceptable on a distinct but irrelevant parse.

→ * Which author's book looked good [before **PRO**$_{???}$ ordering it]?

To sum up, given that parenthetical subtraction fails in numerous PE configurations, I argue that this construction is not derived by insertion of a parenthetical.

3.2 PE and variable binding

Much work has argued that a pronoun bound by a quantifier must be interpreted in a position c-commanded by that quantifier (Reinhart 1983, a.o.), and also observed that a quantifier cannot bind a pronoun in a parenthetical (Safir 1986; Potts 2005; de Vries 2007). This is illustrated in (34) below, where the quantifier *nobody* can bind the co-referent pronoun in (34a), but not (34b), where that pronoun is inside of a parenthetical clause:

(34) *No variable binding into a parenthetical*

  a. Nobody$_1$ claimed that he$_1$ was thinking about Hank.
    (de Vries 2007, ex. 29a)

  b. * Nobody$_1$ was, **he$_1$ claimed**, the dumbest guy in the room.
    (de Vries 2007, ex. 30b)

As (35) below shows, in a PE sentence, it is possible for a quantifier to bind a pronoun that inhabits the content intervening between the extracted possessor and possessum, which is the material that would be taken to be parenthetical under the analysis being argued against in this section. The possibility of such binding is accurately predicted if there is indeed no parenthetical involved in such PE sentences:[17]

(35) *Variable binding possible in relevant segment of PE sentences*

  a. [No nurse]$_1$ is happy for a patient [who **she$_1$ suspects**'s wounds will get infected].

  b. [Every boy]$_1$ feels sorry for someone [who **he$_1$ heard**'s lunch got stolen].

---

[17]de Vries argues that parentheticals are opaque for a variety of other dependencies, such as anaphor binding. In the context of this proposal, the fact that a pronoun intervening between the extracted phrase and possessum in a PE sentence can license an anaphor within the possessum constitutes further evidence against the parenthetical analysis, though such examples are marked:

(i) *Anaphor binding in possessum by matrix subject*

  a. ? Tell me who you$_1$ think[['s picture of yourself$_1$] is the best].
  b. ? Tell us who you$_1$ think[['s drawing of yourself$_1$] you consider the worst].

3.3   No PE from opaque complements

If English PE is an illusion caused by a DP-internal parenthetical, then we expect the set of verbs that are licit in parentheticals, and the set of verbs that can appear to be crossed by illusory PE, to be the same. This is not the case. Consider *whisper*, which is productive in parenthetical clauses:

(36)   *Productive parenthetical* whisper
Mary (John **whispered**) wants (John **whispered**) a kitten (John **whispered**).

This verb is among the manner of speech verbs (*mutter, stammer, mumble, groan*) whose complements are not transparent for extraction, though these verbs are productive in parentheticals (Zwicky 1971, p. 255). As (37) shows, pied-piping possessor movement from the complement of such a verb is unacceptable (37a). An equivalent PE configuration is no better (37b):

(37)   *No extraction (PE or otherwise) with verbs with opaque complements*
a.   *Pied-piping possessor movement*
That's the person [[whose cat]$_1$ I thought/said/***whispered/*groaned** [t_1 was cute]].

b.   *PE*
That's the person [who$_1$ I thought/said/***whispered/*groaned** [[t_1's cat] was cute]].

In contrast, verbs with transparent complements (*think, say, claim, suspect, tell, believe, hear,* etc.) are acceptable with PE, as we see throughout this chapter.

3.4   PE is blocked by islands

Placing an island boundary between the extracted possessor and ['s] results in unacceptability, as expected if this construction involves genuine extraction.

(38)   *Adjunct island constraint*
a.   * Who$_1$ did you go home [because I said [[t_1's friend] is ugly]]?
b.   * [Which kid]$_1$ did you leave class [before I said [[t_1's shoes] are ugly]]?

(39)   *Complex NP constraint*
a.   * Who$_1$ did you hear [the rumor that they said [[t_1's friend] will get hired]]?

b.   * [Which teacher]$_1$ do you know a student [that said [[t_1's class] is hard]]?

(40)   *Coordinate structure constraint*

   a.   * Who$_1$ did [[she say [[t_1's friend] is nice]], and [he say [Bill is mean]]]]?

   b.   ✓ Who$_1$ did [[she say [[t_1's cat] is nice]], and [he say [[t_1's dog] is mean]]]]?

See also section 8.1, which shows that "weak" islands also block English PE.

3.5   Parasitic gap licensing

The extracted possessor can license a parasitic gap. Since a parasitic gap is granted an interpretation only when the constituent that contains it is structurally crossed by A′-movement of the licensing phrase (Engdahl 1983, a.o.), this fact would be unexpected if the possessor is not actually undergoing movement in such examples:

(41)   *Parasitic gap licensed by extracted possessor*

   a.   ? Who$_1$ do you think [t_1's research] is good, [despite not thinking [PG$_1$'s paintings] are nice]?

   b.   ? That's the guy [who$_1$ I said [t_1's table manners] were awful [after having lunch with PG$_1$ today]].

   To conclude this section, we have now seen several sources of evidence that English PE truly involves A′-movement, and not a deceptive parenthetical construction.[18] [19] With this result established, next I summarize the concepts which I argue derive the details of this phenomenon.

---

[18]Given that some English speakers can extract possessors via A′-movement, it would also be reasonable to expect such speakers to permit possessor raising via A-movement, though this is not possible (i). I hypothesize that because English possessors are always Case-licensed DP-internally, A-movement of possessors from DP is blocked, since English lacks phenomena like hyper-raising.

(i)   * John$_1$ was washed [t_1's hands].

[19]If *sluicing* ellipsis is derived by *wh*-movement followed by deletion of all but the moved constituent, then English speakers capable of PE should allow an extracted possessor to serve as the remnant of sluicing, as in (i) below. While I have encountered some disagreement about the judgments on this, most speakers seem to find such sentences unacceptable:

(i)   I know you like someone's cat, but I don't know who*%(se cat).

However, Chung (2013) reports that (ii), which instantiates the right configuration, is acceptable:
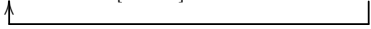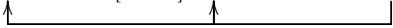
(ii)   They impounded somebody or other's car, but they wouldn't say exactly who(se car).
        (Chung 2013, ex. 66a)

4   Two linearization constraints on English PE

4.1   Cyclic Linearization

As discussed in the introduction, this dissertation argues for the CL theory of phases. To review, under this approach, the completion of the derivation of any given phase results in all of that phase's contents undergoing spell-out and linearization, not just the phase's complement. A necessary consequence of this proposal is the abandonment of the Phase Impenetrability Condition, and thus, a conception of successive-cyclic movement that makes no reference to it. As summarized previously, under CL successive-cyclic movement is predicted from the logic of linearization. By extracting from a phase via its linear edge, in successive-cyclic fashion, the orderings generated by cyclic spell-out are kept coherent for the derivation in question (42):
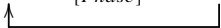
(42)   *Phase exit via linear edge*

   a.   * $[_{ZP}$ α Z $[_{YP_{[Phase]}}$ β Y $[_{XP}$ *t*]]]

   b.   ✓ $[_{ZP}$ α Z $[_{YP_{[Phase]}}$ *t* β Y $[_{XP}$ *t*]]]

In particular, by moving via the linear edge of each phase passed, phase-exiting phrases are determined by PF to precede the rest of the content of each phase in question. This is ultimately consistent with a final representation where the moved material precedes all phases that it has exited.

   Furthermore, we've seen that CL makes a prediction about how to handle phase-exiting movement that does not pass through the linear edge. Any material within a phase crossed-over by such non-successive-cyclic movement must move into the next phase as well, to a position preceding what previously crossed it, thus restoring the original order of those elements. As (43) below shows once again, such additional order-restoring movement keeps the relative order of the relevant constituents consistent, and thus maintains a contradiction-free derivation.

(43)   *Avoiding a linearization problem through order-restoring movement*

   a.   *Illicit non-successive-cyclic phase exit...*
        * $[_{XP}$ β $[_{YP_{[Phase]}}$ α β]]

Chung mentions that such examples are argued to be ungrammatical by Merchant (2001), however. I speculate that this configuration is syntactically licit, but confounded by the complex and often-debated identity relationship between an ellipsis site and its antecedent.

     b.   *...repaired by order-restoring movement*

$$\checkmark\ [_{XP}\ \alpha\ \beta\ [_{YP_{[Phase]}}\ \alpha\ \beta\ ]]$$
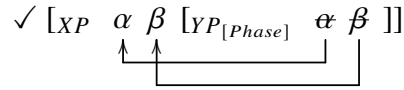
In the next section, we will see that pressure to obey the schemas in (42) and (43) above accurately restricts English PE when combined with the adjacency requirement of [*'s*], which is described next.

4.2   Adjacency of [*'s*] and the possessor

As mentioned previously, several works argue that PE in English is typically prevented by a PF requirement mandating adjacency of the possessor and possessive D (Chomsky 1995; Radford 1997; Gavruseva 2000; Gavruseva and Thornton 2001, a.o.). Indeed, Gavruseva argues that such a requirement is influential in blocking PE in many languages. To describe an English grammar that does not allow PE, it will suffice to state the relevant requirement as follows (44):

(44)   *Possessor-[*'s*] Adjacency (Global version)*
       For any derivation containing [*'s*], [*'s*] must be linearly adjacent[20] to its associated possessor at the final PF representation of that derivation.

This constraint is phrased in such a way that it must be satisfied at the final PF representation generated by the derivation in question. A grammar with such a constraint will never permit separation of possessor and possessum, as is indeed the case for many English speakers.

     However, it is necessary to say something different about the grammar of speakers who do permit PE. I argue that (constrained) PE is an option for such speakers because they are able to satisfy adjacency to [*'s*] more locally. In particular, I argue that such speakers can satisfy this requirement in a phase-bounded way (45):

(45)   *Possessor-[*'s*] Adjacency (Local version)*   [=(8)]
       [*'s*] must be adjacent to its associated possessor at the spell-out of the minimal phase (vP, CP) containing [*'s*].

As we'll see, after satisfaction of this locally-evaluated requirement, subsequent movement operations can break adjacency between the possessor and [*'s*]. Precisely

---

[20]I define adjacency as a relation between two elements α and β, whereby α and β form a linear string with nothing intervening between them. As mentioned in the introduction, this notion of adjacency is not a primitive of CL, which is concerned only with relative order, for which intervening material is irrelevant. I suggest that adjacency of this sort is enforced by the PF requirements of certain elements, intuitively what we term "bound morphemes".

because [*'s*] is not carried along into subsequent phases after successful PE, this adjacency requirement is not applicable to those later phases, and the possessor can move on freely. For more discussion about the nature of this adjacency requirement, see section 8.2 below.

4.3   The importance of spelling-out phase edges

The fact that in the CL theory spell-out targets all material in a given phase is critical for this account of English PE. This system allows phase-level spell-out and the PF adjacency requirement of [*'s*] in (45) to interact with successive-cyclic movement through phase edges. As we'll see, this interaction results in satisfaction of the requirement of [*'s*] only under particular circumstances, as desired.

The needed interaction is not possible for the phase theory proposed by Chomsky (2000, 2001, a.o.), for which phase-level spell-out is limited to phase complements. To see why, consider that in a PE derivation, successive-cyclic A′-movement brings the possessor to the edge of each phase being exited. In order for PE to actually occur, there will necessarily be a point in the derivation where the possessum is stranded in the complement of a phase to whose edge the possessor has extracted. In such a configuration, as schematized in (46), the extracted possessor and the possessum are separated by a spell-out domain (here YP):

(46)   *Possessor and* ['s] *separated by a spell-out domain*
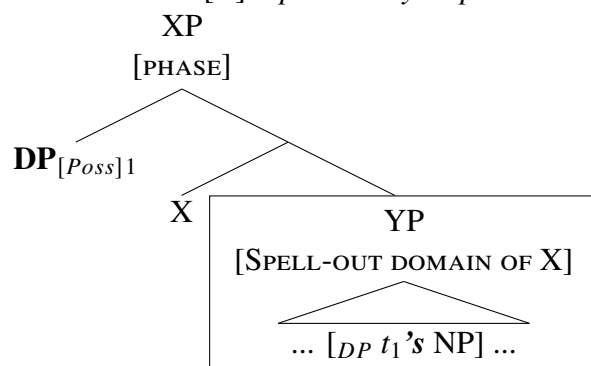
When spell-out applies in (46), the local adjacency requirement of [*'s*] will not be met. This is because the extracted possessor has moved outside of the spell-out domain YP of this phase XP, before spell-out applied to YP. Thus spell-out of YP finds [*'s*] non-adjacent to the possessor, and this derivation fails. This failure is avoided if the possessum is pied-piped along with movement of the possessor to the phase edge. However, in doing this, PE fails to occur. This issue arises at any point where a spell-out domain would separate the possessor and possessum, leading this theory of phases to incorrectly predict a total lack of PE.

116

5 Predicting the facts

This section shows how the concepts explained above predict the details of PE in English, which as we've seen, obeys the following generalization:

(47)   *Generalization about English PE*   [=(6, 17)]
        A possessor must reach the left linear edge of the local CP before extracting
        from the possessum DP.

First I discuss PE from subjects, followed by the more complex case of PE from non-subjects. For the purposes of this analysis I take all vPs to be phasal following Legate (2003), Ko (2014), and references therein.

5.1 PE from subjects

This subsection analyzes PE from subjects. We've seen that English PE is only possible for cross-clausal movement, a fact whose explanation will be clear by the end of this section. Given this constraint, the analysis of a PE sentence must begin with the vP of the embedded clause in which the possessor originates. In the case of PE from subjects, this part of the derivation is straightforward, since the interactions most important to this analysis do not fully emerge until the derivation of the embedded CP.

5.1.1 PE from subjects: The embedded vP level

The facts reported above show PE from both external argument and internal argument subjects of all kinds. (See for instance (9a), (9c), (21a), and (21b) above). To begin the analysis, let's consider what occurs within the embedded vP of a derivation with PE from an external argument. If external arguments originate in spec-vP, as in transitive and unergative contexts, then no intermediate successive-cyclic movement of the possessor is necessary at this stage of the derivation. The in situ external argument and its possessor simply originate at the linear edge of the vP phase, which they will soon exit after respectively undergoing A- and A$'$-movement:

(48)   *Transitive / unergative vP: Subject and its possessor originate at phase edge*
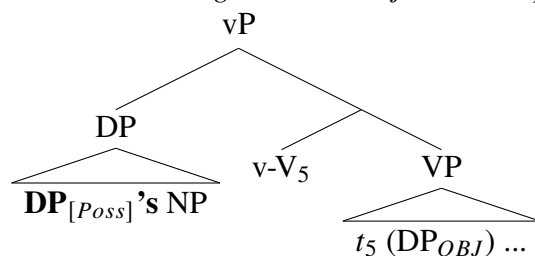
While we can imagine that string-vacuous PE might occur here, moving the possessor to a higher spec-vP above the external argument as in (49) below, this possibility is ruled out by the hypothesis already adopted in the preceding chapters that movement from one specifier to another of the same phrase is banned.

(49)   *No vP-internal PE from transitive / unergative subject*
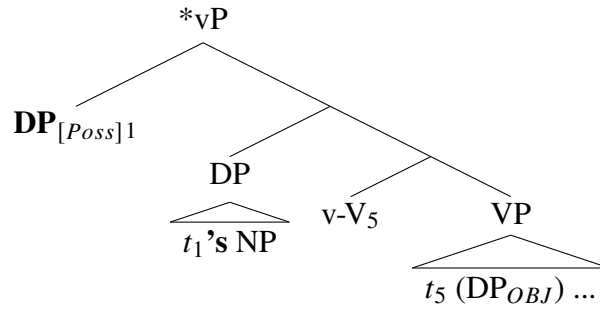
Thus PE at this stage of such a derivation is not only unnecessary as far as CL is concerned, but impossible anyway.

In contrast, consider the vP of a derivation with PE from a subject that is an internal argument, as in unaccusative and passive contexts. Independent of PE, given that such a subject is externally merged as a complement of V, it must move to spec-vP in order to maintain a coherent linearization under CL: since English V moves to v (Larson 1988; Chomsky 1995; Kratzer 1996, a.o.), and because the subject ultimately precedes V from its final position in spec-TP, it is necessary for the internal argument subject to move to spec-vP in order to precede V at the spell-out of vP. That A-movement of subjects indeed passes through vP edges is independently argued by Sauerland (2003) on the basis of reconstruction. Notice that such movement of an internal argument subject automatically brings a possessor contained within it to the linear edge of vP, as we see in (50) below:

(50)   *Subject movement to edge of unaccusative / passive vP*
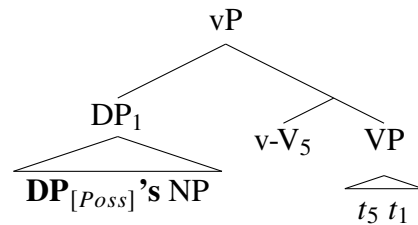
It is also in principle possible for the possessor to extract to the vP edge, with the internal argument subject then moving to a lower specifier of vP below the extracted possessor via tucking-in (Richards 1997, 1999, a.o.) resulting in the structure in (51) below. This string-vacuous possessor extraction satisfies the adjacency requirement

of [*'s*] just as if the possessor had not exited DP.

(51)　　*PE from passive / unaccusative subject with subject tucking-in*

$$
\begin{array}{c}
\text{vP} \\
\text{DP}_{[Poss]2} \quad \\
\text{DP}_1 \quad \text{v-V}_5 \quad \text{VP} \\
t_2 \text{ 's NP} \quad t_5\ t_1
\end{array}
$$
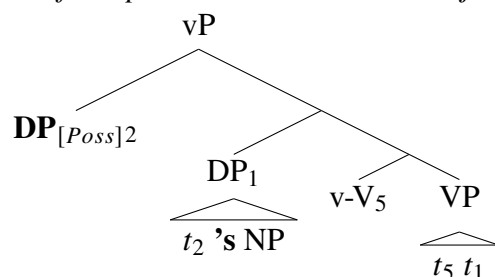
Because the derivation in (50) accomplishes the same thing as (51) but with less movement operations (modulo the different final constituency), we might expect concerns of economy to favor (50).[21] However, nothing of substance for this chapter's account changes if the reality is (51), or if both derivations are available.

　　This concludes the analysis of the embedded vP for derivations with PE from a subject. So far, the adjacency requirement of [*'s*] has not been critical, since the context where its influence is most evident has not yet arisen. This requirement is much more influential on the derivation of the CP phase, as we'll see next.

5.1.2　PE from subjects: The embedded CP level

After the completion of vP, I assume that upon external merge of T, the subject A-moves to spec-TP. When C is merged, the opportunity for PE arrives. In section 2, I showed that at this stage of the derivation, various restrictions hold. In short, as (52) shows again, nothing can intervene between the trace of PE in the possessum DP and the linear edge of the embedded CP:

(52)　　*CP edge restrictions on PE from subjects*
　　　　a.　　Who$_1$ did you say [(***often**) [t_1's friend] (often) has money]?
　　　　b.　　Who$_1$ do they think [(***that**) [t_1's name] (*that) is Mary]?

Before analyzing the unacceptable variants of the examples in (52), let's establish why these examples are acceptable when this problematic material in the left edge of the embedded CP is absent.

　　If no such material is present in the embedded CP edge, then after A-movement of the possessed subject to spec-TP, the possessor it contains is already at the linear

---

[21]Furthermore, the derivation described for (51) may represent an instance of *chain interleaving*, which following Collins (1994), is independently ruled out. If this is so, then on these grounds as well, the derivation in (50) is the preferable hypothesis.

edge of CP. The possessor could string-vacuously extract to spec-CP from here, though such movement is unnecessary (53):[22]

(53)   *Harmless string-vacuous PE from subject*

Linearization: $DP_{[Poss]} < \text{'s} < NP < T < vP$
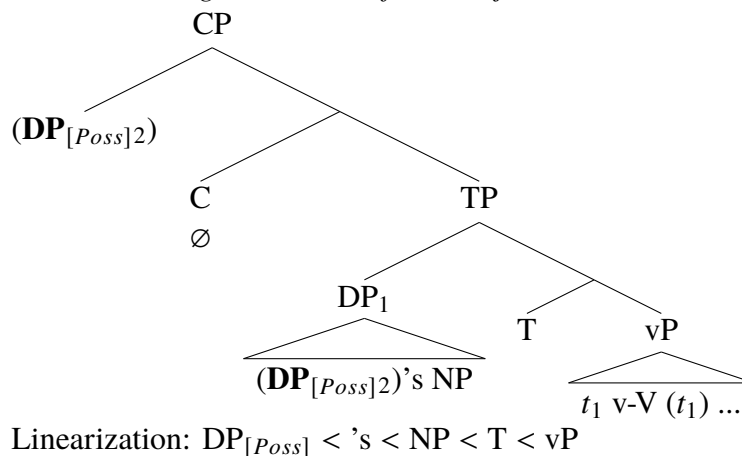
The linear order established at the spell-out of the CP in this situation satisfies the adjacency requirement of [*'s*], which is linearly adjacent to the possessor at PF whether or not the possessor string-vacuously extracts at this point. CL is also satisfied here, since the extracting possessor has reached the linear edge of the embedded CP either way.

Next, the possessor can extract into the matrix vP (and then onward), stranding the possessum and the [*'s*] it contains in the lower CP, as in (54) below. When the matrix vP spells out, [*'s*] is not present within that phase to enforce its adjacency requirement. This is because [*'s*] has been stranded in a lower phase that has already undergone spell-out (as signified by the box around the embedded CP in (54) below), at which point the adjacency requirement of [*'s*] was locally satisfied. As such, extraction of the possessor succeeds:

---

[22]The vPs in the representative diagrams in (53-56) are compressed in order to be compatible with a derivation involving either an external or internal argument subject.

(54)   *Successful PE into matrix vP: Possessum stranded in embedded CP*

```
                    vP
            ╱             ╲
   DP[Poss]2          ╱      ╲
              v-V₆            VP
                        ╱         ╲
                      t₆            ┌────────────────────────┐
                                    │         CP             │
                                    │     ╱      ╲           │
                                    │    C         TP        │
                                    │    ∅      ╱    ╲       │
                                    │        DP₁           │
                                    │       ╱╲    T    vP   │
                                    │   t₂ 's NP      ╱  ╲  │
                                    │             t₁ v-V (t₁) ... │
                                    └────────────────────────┘
```
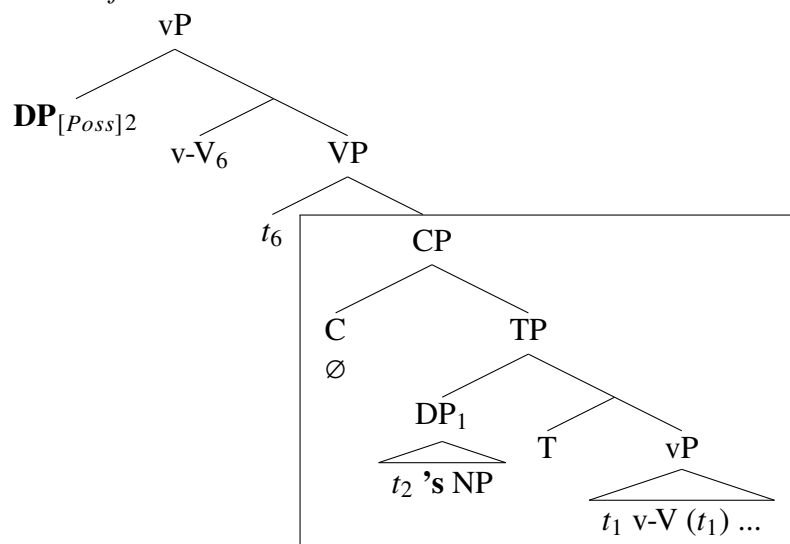
Critical to this analysis is the hypothesis that the adjacency requirement under discussion is a property of the bound morpheme [*'s*] only, not of the possessor.

Next, let's examine a PE derivation where there is problematic material in the edge of the CP exited by PE. Consider (55) below, where the embedded CP contains an adverb in the left periphery. CL motivates the possessor to exit this CP via its linear edge. Therefore the possessor must move to a position within CP to the left of that high adverb, in order to facilitate later movement. Notice that if the possessor moves around the adverb within CP in this way, thereby stranding the possessum subject in spec-TP, then the adverb consequently intervenes between the moved possessor and [*'s*]:

(55)   *Possessor movement around intervening adverb to linear edge of CP*

```
                    CP
            ╱             ╲
  DP[Poss]2          ╱        ╲
            AdvP            ╲
                 ╱      ╲
                C          TP
                        ╱      ╲
                     DP₁          ╲
                    ╱╲    T        vP
                t₂ 's NP        ╱     ╲
                            t₁ v-V (t₁) ...
```
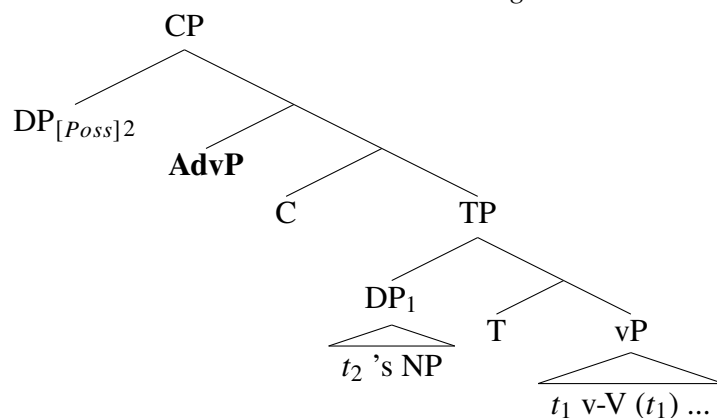
While movement of the possessor from the CP in (55) would be licit under CL,

there is a problem. When spell-out occurs here, PF will find the possessor and [*'s*] non-adjacent in this CP due to the intervening adverb. Therefore this CP will be deviant at PF. However, there is a way to avoid this problem—pied-piping the possessum with the possessor's movement over the adverb to the edge of CP (56):

(56)   *Possessum pied-piping over adverb to the linear edge of CP*
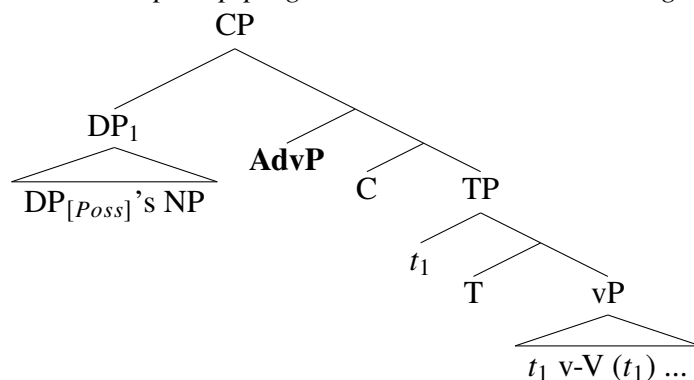
This pied-piping movement satisfies CL as well as the adjacency requirement of [*'s*] within this embedded CP. After (56), the possessor can successfully extract into the matrix vP. In such a derivation we end up with the high adverb to the right of the stranded possessum subject. As we saw in (52a) above, this is indeed the only way to have such an adverb in a CP exited by PE.[23]

Derivations in which instead of a high adverb there is an overt complementizer *that*, such as (52b) above, will be identical in the relevant ways to what we have just seen for the adverb scenario. If CP contains an overt C, the possessor must move to its left, pied-piping the possessum with it to spec-CP in order to maintain adjacency with [*'s*]. The eventual stranding of the possessum in spec-CP will result in deletion of the complementizer due to the Doubly Filled Comp Filter, something we independently know to hold in English. Thus as (52b) showed, an overt C on either side of a subject that PE has exited is impossible.

This concludes the analysis of PE from subjects. Next I consider PE from non-subjects, which is governed by the very same principles.

---

[23]This result could also have been achieved by adjoining the adverb linearly to the right rather than to the left of the subject, but we see here that even if the adverb originates left of the subject, the derivation can converge.

5.2 PE from non-subjects

5.2.1 PE from non-subjects: The embedded vP level

Recall that PE from a non-subject DP requires that DP to be pied-piped to the edge of the local CP in English. Consequently, in contrast to more typical PE in a language like Hungarian, English PE cannot strand a non-subject possessum in its base position. This fact is repeated below:

(57)   *Non-subject exited by PE must be stranded in a CP edge*

    a.   * The person [**who**$_2$ I think [John ate [t_2**'s food**]]] is Mary].

    b.   ✓ The person [**who**$_2$ I think [[t_2**'s food**]$_1$ John ate t_1] is Mary].

To understand why this is so, let's first examine such derivations at the embedded vP level. PE from any non-subject DP will work in essentially the same way.

In (58) below we see a transitive vP in which PE has exited a direct object, stranding it in situ in the complement of V. This derivation would be satisfactory for CL, since the moving possessor has reached the linear edge of this phase. However, spell-out of this structure will find it to be in violation, since the in situ subject (here EA, the external argument) and V, both underlined below, intervene between the moved possessor and [*'s*]:

(58)   *No PE within vP from in situ non-subject*
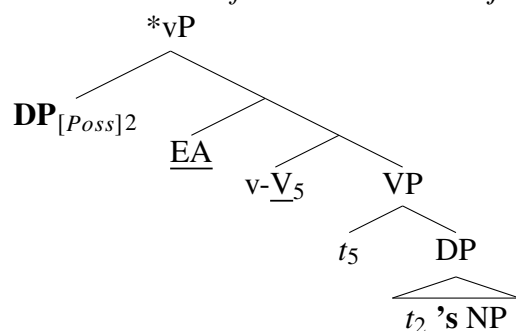
Satisfying CL as well as [*'s*] requires pied-piping the object possessum to the vP edge along with the possessor's movement, instead of stranding it. Consequently, the possessor reaches the phase edge while keeping [*'s*] adjacent, as in (59) below. These considerations make clear why base position stranding of a non-subject exited by PE is impossible. Such stranding violates the requirement of [*'s*] at the spell-out of vP, leaving pied-piping of the non-subject as the only option:

(59)   *Pied-piping of object possessum with moving possessor within vP*
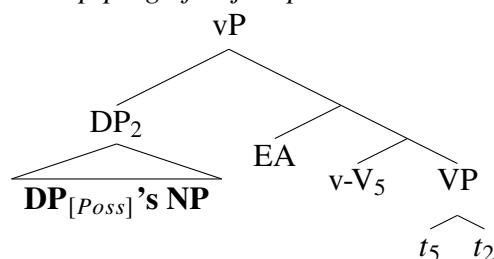
At this point, the non-subject possessum has been pied-piped into the edge of vP, but it cannot remain here, as demonstrated in (60):

(60)   *Non-subject possessum cannot be stranded in spec-vP by PE*
   a.   * The person [**who**$_2$ I think [$_{CP}$ they [$_{vP}$ [t_2**'s food**]$_1$ ate t_1]]] is Mary.
   b.   * [**Which student**]$_2$ did you say [$_{CP}$ she [$_{vP}$ [t_2**'s book**]$_1$ found t_1]]?

As we've seen in previous examples, a non-subject possessum must ultimately be pied-piped to the edge of the embedded CP under PE. This fact is now puzzling, since the analysis so far provides no reason why the pied-piped possessum should not be able to remain in spec-vP, where the adjacency requirement of [*'s*] was met.

To see why the concepts defended in this chapter in fact automatically predict that the non-subject possessum cannot be stranded in spec-vP and thus must move on, we must consider the next phase of the derivation. The explanation for this fact that we are about to see is precisely the same as the account in section 5.1 of the previous chapter for the impossibility of adjunct stranding in spec-vP under *wh*-movement in English.[24]

5.2.2   PE from non-subjects: The embedded CP level

Upon the merge of T, the subject A-moves to spec-TP from its external merge site in spec-vP. This movement carries the subject across the non-subject possessive DP which in (59) above A′-moved to an outer spec-vP, due to pied-piping of the possessum along with the possessor's movement. The result is the structure in (61):

---

[24]Section 5.3 of the previous chapter discussed that there is cross-linguistic variance in the possibility of stranding in a clause-medial position, by assumption the specifier of an intermediate phase such as vP. As mentioned in that section, this result could be taken as evidence for cross-linguistic variation in the set of phases, though this hypothesis raises puzzles about how phasehood might vary that this dissertation will not resolve. Nevertheless, the distribution of both English PE and of adjunct stranding under *wh*-movement corroborate the view that, at least in the English variety under consideration, the relevant phase is simply the vP. We may predict these phenomena to behave differently in some of the varieties of West Ulster English analyzed by Henry (2012), which as we saw in the previous chapter, do appear to permit intermediate stranding in spec-vP.

(61)   *A-movement of subject to spec-TP across moved possessive DP in vP edge*

TP

EA$_1$

T   vP

DP$_2$

t_1 v-V$_5$ VP

DP$_{[Poss]}$'s NP

t_5 t_2
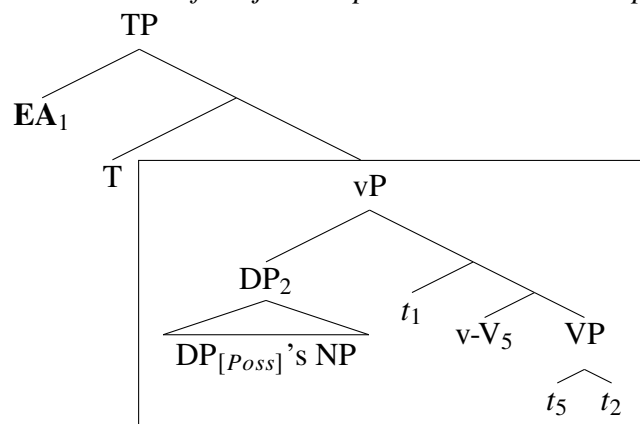
Recall that CL motivates elements exiting a phase to pass through that phase's linear edge. For this reason, A-movement of the subject might in fact pass through a higher spec-vP, above the moved possessive DP, as in (62) below. However, such a derivation requires illegal phrase-bounded specifier-to-specifier movement within vP, and therefore will not be available:

(62)   *No subject A-movement via higher spec-vP above pied-piped possessum*

*TP

EA$_2$

T   vP

t_2

DP$_1$

t_2 v-V$_5$ VP

DP$_{[Poss]}$'s NP

t_5 t_1
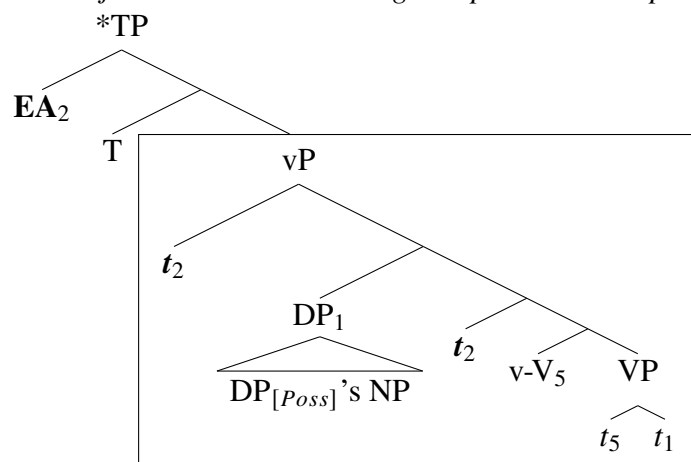
The derivation in (62) would end up problematic for linearization anyway: here the subject would precede the possessor at the spell-out of vP, but the possessor will ultimately extract and move to spec-CP, and thus precede the subject. Since in this situation the relative order of possessor and subject is not consistent across both phases, a linearization contradiction will arise. In contrast, the derivation in (61) avoids this linearization problem, since here the possessor precedes the subject within vP, just as will be the case post-PE. The linearization problem posed by (62) would be avoided by extracting the possessor to a position within vP that is

125

above the subject's derived position here, since this would yield the needed linear order $[DP_{[Poss]} <$ EA $<$'s NP] within the vP. The same order would also have been derived by tucking-in the pied-piped possessive DP into a spec-vP below the subject in its θ-position, before then extracting the possessor to a position in vP above the subject. Both of these possibilities are unavailable because they rely on some degree of specifier-to-specifier movement within vP, however. Consequently, linearization requirements and the impossibility of movement within the vP edge leave the derivation in (61) as the only possibility.

Since in (61) the subject must non-successive-cyclically move across the pied-piped possessive DP in the vP edge, CL makes a prediction about what movement steps will have to follow (61), which turn out to be correct. As discussed earlier, CL makes a general prediction about how to permit phase exits from a non-edge position. In these scenarios, the material crossed over by movement from a non-edge position in the phase must also move into the next phase, to a position preceding what previously crossed it:

(63) *Avoiding a linearization problem through order-restoring movement*

    a. *Illicit non-successive-cyclic phase exit...*

      $* [_{XP} \quad \beta \; [_{YP_{[Phase]}} \quad \alpha \quad \beta \;]]$

    b. *...repaired by order-restoring movement*

      $\checkmark [_{XP} \quad \alpha \quad \beta \; [_{YP_{[Phase]}} \quad \alpha \quad \beta \;]]$

Such additional movement ensures a coherent linearization for the elements involved. Given this prediction, if the A′-moved possessive DP in spec-vP must be non-successive-cyclically crossed by A-movement of the subject as in (61), then we expect that the possessum will be unable to remain in spec-vP. Rather, it must move to a position that precedes the subject within the next phase. This is precisely what is accomplished by continuing to pied-pipe the possessum along with successive-cyclic A′-movement of the possessor to the edge of the embedded CP. As we see in (64) below, this maintains a coherent linearization:

(64)   *Non-subject possessum must be pied-piped to spec-CP*
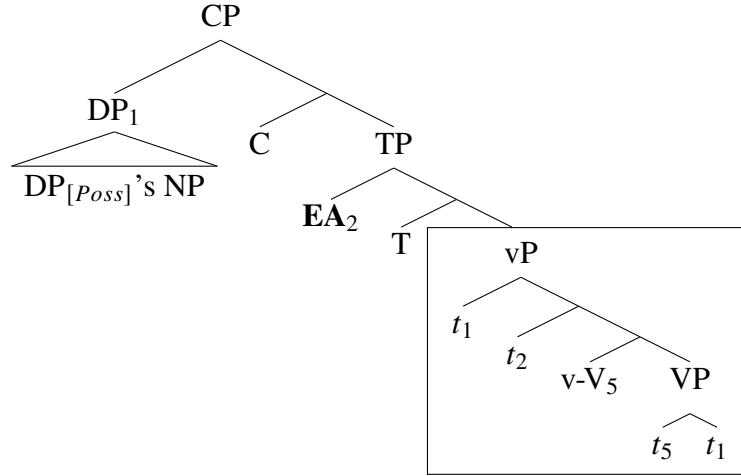
1. Linearization of vP: $DP_{[Poss]} <$'s $< NP < EA < V$ ...
2. Linearization of CP: $DP_{[Poss]} <$'s $< NP < C < EA < T < vP$ ...

The present account thus correctly predicts that non-subject possessum DPs must be pied-piped to the embedded spec-CP when exited by PE. While in principle non-subject possessum stranding in spec-vP should be licit, the interaction of CL with A-movement of the subject forces additional pied-piping to CP.[25]

After pied-piping of the non-subject possessum to the edge of the embedded CP as in (64) above, the possessor can move on freely into the higher clause, as diagrammed in (65) below:

---

[25]Example (15) above showed that expletive associates exited by PE in English must be pied-piped to the edge of the local CP, just as all non-subject DPs must, as repeated in (i):

(i)   a.   * **Who**$_2$ did Mary say [there were [t_2**'s books**] on the table]?
        b.   ? **Who**$_2$ did Mary say [[t_2**'s books**]$_1$ there were t_1 on the table]?

This possessum stranding can be used as a diagnostic for the derivational history of the expletive. Example (ia) above is ambiguous between stranding in situ or in spec-vP, since copular V in English moves to T, unless T is filled by an auxiliary. For this reason (ii) below adds an auxiliary to disambiguate the position of stranding:

(ii)   **Who**$_1$ do you think [$_{CP}$ ($^?t_1$**'s friends**) there have [$_{vP}$ (*t_1**'s friends**) been a lot of stories told to (*t_1's friends)]]?

Here we see that the expletive associate exited by PE cannot remain in a position corresponding to spec-vP, as expected if the expletive moves from vP, and thus prevents the stranded possessum from occupying the vP edge. This follows from a theory in which expletives originate in vP, just as typical subjects do, as discussed in section 5.1.1 of the previous chapter.

(65)   *Successful PE after pied-piping non-subject possessum to CP edge*

[Syntax tree diagram:

vP branches into:
- DP$_{[Poss]3}$
- v-V$_6$
- VP, which branches into:
  - t_6
  - CP, which branches into:
    - DP$_1$ (with t_3's NP)
    - C
    - TP, which branches into:
      - **EA$_2$**
      - T
      - vP, which branches into:
        - t_1
        - t_2
        - v-V$_5$
        - VP (with t_5, t_1)]
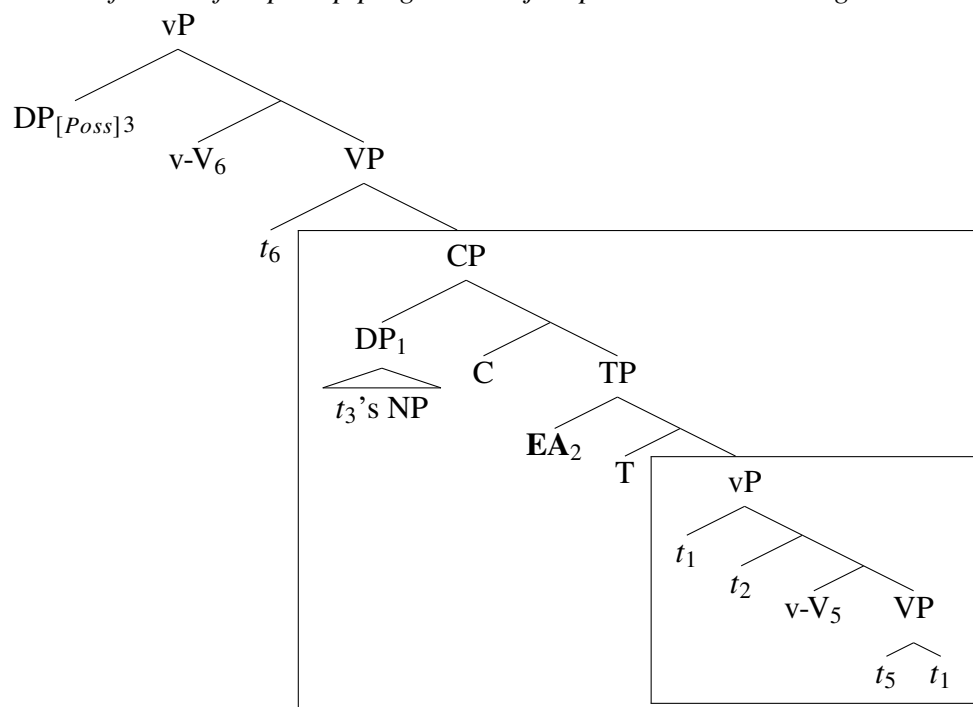
Nothing forces further pied-piping, as the facts have shown. The possessor extracts on into the matrix clause, with [*'s*] remaining stranded below in the lower CP phase.

We have just seen a successful derivation of PE from a non-subject, but this is not the end of the story. Recall that, just as for PE from subjects, PE from non-subjects involves a restriction on the content of the embedded CP. That is, no material may intervene between the trace of PE in the possessum DP and the embedded CP edge:

(66)   *CP edge restriction on PE from non-subjects*

   a.   *No overt C on either side of pied-piped non-subject*
        Who$_2$ did you say [$_{CP}$ (*that) [t_2**'s cat**]$_1$ (*that) John saw t_1]?

   b.   *No adverb left of pied-piped non-subject in CP*
        Who$_2$ did you say [$_{CP}$ (*allegedly) [t_2**'s cat**]$_1$ (allegedly) John stole t_1]?

   c.   *No pied-piped P left of pied-piped non-subject*
        Who$_2$ do you think [$_{CP}$ (***from**) [t_2's house]$_1$ we should leave (**from**) t_1]?

The concepts under discussion yield a familiar explanation for this constraint. To satisfy CL, the possessor must move to a spec-CP above any peripheral material

separating it from the linear edge of CP, before exiting it. In scenarios like those in (66), this would require the possessor to move to a higher spec-CP from within the non-subject possessum that it previously pied-piped into CP. This would thus be extraction from one specifier to another within the same CP, which is independently ruled out. Additionally, such movement over the problematic peripheral material would render the possessor and [*'s*] illegally non-adjacent within this CP at the time it spells-out (see 5.1.2). Since no movement within the CP edge is possible, the only option is for the possessor to non-successive-cyclically cross over any problematic peripheral material in CP when moving from the pied-piped non-subject and into the matrix vP, thus violating CL.[26]

5.3 Why English PE requires cross-clausal movement

Recall that English PE is impossible for A$'$-movement that is not cross-clausal:

(67)   *PE unavailable for clause-internal movement*
     a.   * Who$_1$ will [t_1's article] be published next year?
     b.   * Who$_1$ did they criticize [t_1's article]?

Given that non-subject possessum DPs must be pied-piped to a CP edge prior to PE as just discussed, there is no chance for the possessor to extract from a non-subject if the derivation contains only one clause. In this case, there is no opportunity for the possessor to break away after pied-piping the possessum to the local CP edge, since at this point the derivation ends. Thus more than one clause is needed for PE from non-subjects to occur.

    For similar reasons, PE from a subject possessum cannot become evident in a single clause derivation. In such a scenario, any material between the extracted possessor and [*'s*] that would diagnose the occurrence of that PE violates the adjacency requirement of [*'s*] within CP, as (68) below illustrates:

---

[26]This chapter's account of the CP edge restrictions for PE from both subjects and non-subjects accurately derives the impossibility of any material appearing between a stranded possessum and the linear edge of the local CP. These concerns lead us to expect that material merged linearly adjacent to the stranded possessum, but structurally outside of the CP that contains it, should be legal. In fact, such configurations are judged as unacceptable by many, as shown with a matrix adverb in (i):
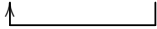
(i)   a.   Remind me who$_2$ you said ($^{\%/??}$yesterday) [$_{CP}$ [t_2's cat] won the contest]?
    b.   Remind me who$_2$ you said ($^{\%/??}$yesterday) [$_{CP}$ [t_2's cat]$_1$ we sent a prize to t_1]?

Examples of this sort likely would require extraposing the embedded CP exited by PE to the right of the matrix adverb to achieve this word order. Given that extraposed CPs in English must have an overt complementizer (Stowell 1981), such movement is not licit here, since a CP exited by PE cannot contain an overt complementizer for the reasons discussed above.

(68)   *Diagnosing clause-internal PE from subject violates adjacency*

    a.   * $[_{CP}$ Who$_2$ **will**(C-T$_1$) $[_{TP}$ $[_{DP}$ t_2's kid] t_1 win the contest]]?

    b.   * Tell me $[_{CP}$ who$_1$ **probably** $[_{TP}$ $[_{DP}$ t_1's pie] was biggest]].

Thus the concepts defended here correctly derive the fact that English PE cannot become apparent unless movement of the possessor crosses a clause boundary.

    Nothing in this account bans string-vacuous or covert possessor extraction from occurring within one clause, however. If variable binding of a pronoun by a quantifier phrase requires the latter to c-command the former ([Reinhart 1983](), and many following), then the fact that a possessor which is a quantifier phrase can bind a pronoun lower in the clause (69) suggests that such covert PE is in fact possible:

(69)   *C-command for variable binding via covert PE*
    [$\varnothing_1$ [[Every boy]$_1$'s mother] likes him$_1$].

5.4 Predictions about multiple embedding

In a configuration with multiple recursively embedded CPs, we expect PE to allow stranding of the possessum in the edge of any CP in the possessor's movement path, provided that the constraints discussed in this section are met. This is correct:

(70)   *PE with multiple embedding: Possessum stranding in any CP edge*[27]

    a.   Who$_1$ do you think $[([t_1$'s cat]) he said $[([t_1$'s cat]) is cute]]?

    b.   Who$_2$ do you think $[([t_2$'s cat]$_1$) he said $[(^?[t_2$'s cat]$_1$) they saw t_1]]?

This analysis also makes several predictions about recursive possession constructions. First of all, nothing prevents extraction of the innermost possessor in such a context, as in (71a), nor of the outermost possessor, as in (71b). Furthermore, given a recursive possession construction in a multiple embedding context like (70) above with multiple intermediate CP edges, we predict the possibility of extracting the outermost possessor to a higher intermediate CP, and then extracting the innermost possessor from this phrase, resulting in the recursive possessive DP having been split into three discontinuous constituents. Such an example is shown in (71c) below, which is acceptable but quite marked, unsurprisingly given its complexity:

---

[27]We see in (70b) that PE from a non-subject that strands the possessum in the lower spec-CP is somewhat degraded. Gary Thoms (p.c.) suggests that the same holds for (70a), and suggests that the distance between the discontinuous elements here must generally be minimized. Gavruseva and Thornton (2001) propose an economy constraint motivating minimization of steps of pied-piping movement, which this contrast appears to contradict. I will leave this puzzle aside for now.

(71)   *PE with recursive possession*

  a. Who_1 did you say [[t_1's brother's friend] is coming with us]?

  b. [Whose brother]$_1$ did you say [[t_1's friend] is coming with us]?

  c. ?? Who_1 did you think [[t_1's brother]$_2$ I said [[t_2's friend] is coming with us]]?

5.5 Why Cyclic Linearization is vital

The CL theory is vital to the account provided in this section, because of its inclusion of phase edges in spell-out domains. This feature of CL allows successive-cyclic movement through phase edges to interact with the adjacency requirement of [*'s*], which is enforced at spell-out. I have shown that CL and this requirement together derive some otherwise puzzling facts about English PE. These facts are intricate, and the construction in which they hold is both far outside the purview of prescriptive English grammar, and does not appear to be very frequent (as perhaps indicated by the fact that it is nearly undocumented). Given this, the complex distribution of English PE is unlikely to be governed by an arbitrary list of memorized restrictions that the learner builds based on their linguistic input. Rather, these details should emerge from more general grammatical principles.

  This chapter has argued that these facts emerge from two main principles, one language-specific and one general. The first is the availability of a phase-bounded adjacency requirement of [*'s*] in colloquial English grammar. The second is CL, which by hypothesis, is automatically possessed by all speakers. Having CL intrinsically, the only piece of information English speakers must learn in order to determine whether their grammar bans PE, or permits it in the restricted form described here, is the point in the derivation when the requirement of [*'s*] may be satisfied.

  An important detail captured by this account is the fact that non-subject possessum DPs must be pied-piped as far as CP under PE. I argued that stranding of non-subject possessum DPs in spec-vP should be possible in principle, though in reality it is not (60). We have seen that under CL, the crossing-over of this position by A-movement of the subject is predicted to require that it be emptied, thus forcing the possessive DP to be pied-piped further. This is an instance of a general prediction of CL that we have seen in action in the previous chapter as well—that if an element in a phase is crossed by something non-successive-cyclically moving from that phase, then that crossed-over material must also move, to a position above what crossed it. In the approach to phases in Chomsky's original work, there is no reason why movement of a lower specifier across a higher one of the same phase should require movement of the higher one as well. This is an automatic consequence of the CL approach, however, which the results of the following chapters will also verify.

6   A PE analysis of *that's*-relatives

This section examines an additional configuration that can be analyzed as involving PE, which I'll refer to as the *that's*-relative (Seppänen and Kjellmer 1995; McDaniel et al. 1998, 2002). This relativization strategy, available to many (though not all) English speakers, can prima facie be described as relativization of a possessor that is accomplished by replacing the typical relative pronoun *who* with *that*:

(72)   *That's-relatives*[28]
      a.   This is the girl [**that's** hat is red].
      b.   There's the house [**that's** door is blue].
      c.   This is the dog [**that's** food the cat ate].
      d.   That is the bike [**that's** wheel the kids broke].

The concepts defended in this chapter lead to an account of this construction with the addition of a proposal from Deal (2016), that TP is a phase in relative clauses.

Deal argues that relativization involves successive-cyclic A′-movement to spec-CP via an intermediate step of movement through an outer specifier of a phasal TP, based on facts about word order, complementizer choice, and case attraction in Nez Perce. Deal argues that this conclusion also explains anti-*that*-trace effects (Bresnan 1972) in English,[29] in the context of the account of *that*-trace phenomena in Pesetsky and Torrego (2001). The concept that TP in relative clauses is a phase sheds new light on the nature of *that's*-relatives, when taken in the context of the analysis of English PE in this chapter.

At first glance, it is conceivable that *that's*-relatives use an instance of *that* which

---

[28]This example shows *that's*-relatives formed with relativization from both subjects (72a-b) and non-subjects (72c-d). McDaniel et al. (2002) report the results of experimental work indicating that examples of both sorts are generally acceptable, with non-subject examples being somewhat less so. I have observed that some speakers accept subject but not non-subject *that's*-relatives. Given the evidence that both are in principle acceptable, I will not dwell on this contrast here.

[29]The anti-*that*-trace effect describes the necessity of the complementizer in subject relatives:

(i)   This is the woman [***(that)** likes cats].

This effect is also present in *it*-clefts:

(ii)   It's Mary ***(that)** likes cats.

While uniting cleft clauses with relative clauses is debated, under Deal's analysis, this fact suggests a TP phase in cleft clauses also. The presence of a TP phase in cleft clauses accurately predicts the availability of a *that's*-relative-like structure in clefts, under the analysis we will see in this section:

(iii)   It's John **that's** family loves dogs.

is a relative pronoun / *wh*-operator. The position occupied by *that* in this construction is not obvious based on the surface string, but if *that* really can be a relative pronoun, then it should be able to take the place of a typical *wh*-operator when relativization involves a pied-piped PP, for instance. This is in fact impossible:[30]

(73) *"That" is not a productive relative pronoun*

    a.   Mary is the student [[about **whose/*that's** book] I wrote a review].
    b.   I saw a man [[on **whose/*that's** face] there were many tattoos].
    c.   That is the person [[with **whom/*that**] I traveled].
    d.   This is the office [[to **which/*that**] I sent a package].

We expect this result if *that* is in fact never a relative pronoun. In the absence of such a possibility, the natural hypothesis is that the *that* of *that's*-relatives is simply a complementizer, as usual, which has somehow come to be adjacent to a possessive [*'s*]. This is the analysis that McDaniel et al. (2002) argue that modern English is indeed converging on. If this is so, it suggests that *that's*-relatives must actually have a gap in possessor position, from which a silent relativizing *wh*-operator moves to spec-CP, as shown below for a relativization of a possessor from a subject:

(74) *That's-relative (with operator movement from a subject possessum)*
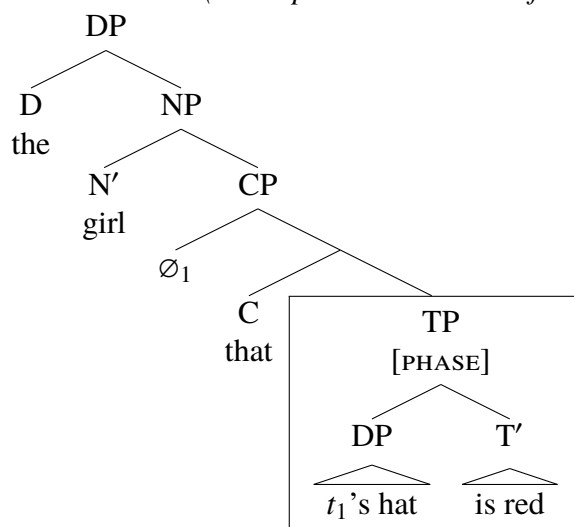
Such an analysis might be regarded as ad hoc under standard assumptions about the impossibility of PE in English, but this is not so for the present chapter.

---

[30]Pied-piping of larger structures, as in appositive relative clauses, shows the same contrast:

(i)   This is Susan, [[pictures of **whom/*that**] always turn out great].

This chapter has shown that English PE can occur with all types of A′-movement, including relativization, provided that such movement strands the possessum at the edge of an embedded CP. Section 5 argued that this is so because CL and the adjacency requirement of [*'s*] together prevent PE from occurring until the possessum can be stranded in a phase edge. As we've seen, independent factors prevent the possessum from occupying the vP edge, thus it always surfaces in an embedded CP edge in the PE examples shown so far:

(75)   *Typical PE in relative clauses with possessum stranding in CP edge*
    a.   Mary is the author [**who**$_1$ they said [$_{CP}$ [t_1**'s new book**] is good]].
    b.   The person [**who**$_2$ I said [$_{CP}$ [t_2 **'s food**]$_1$ John ate t_1] is Mary].

If TP is a phase in relative clauses, then such clauses should allow a second possibility for PE derivations: instead of the possessum following the moving possessor until reaching a CP edge as in (75), the possessum could be stranded in the edge of the phasal TP, below C. Such a derivation produces a *that's*-relative. In the case of a *that's*-relative with a gap inside of a subject possessum, the possessor *wh*-operator simply exits the subject, which remains in the specifier of the phasal TP, as already diagrammed in (74) above. In a *that's*-relative with a gap in a non-subject possessum, this possessum is pied-piped above the subject to an outer specifier at the edge of the TP phase, with the *wh*-operator subsequently extracting to spec-CP:

(76)   *That's-relative (with operator movement from a non-subject possessum)*
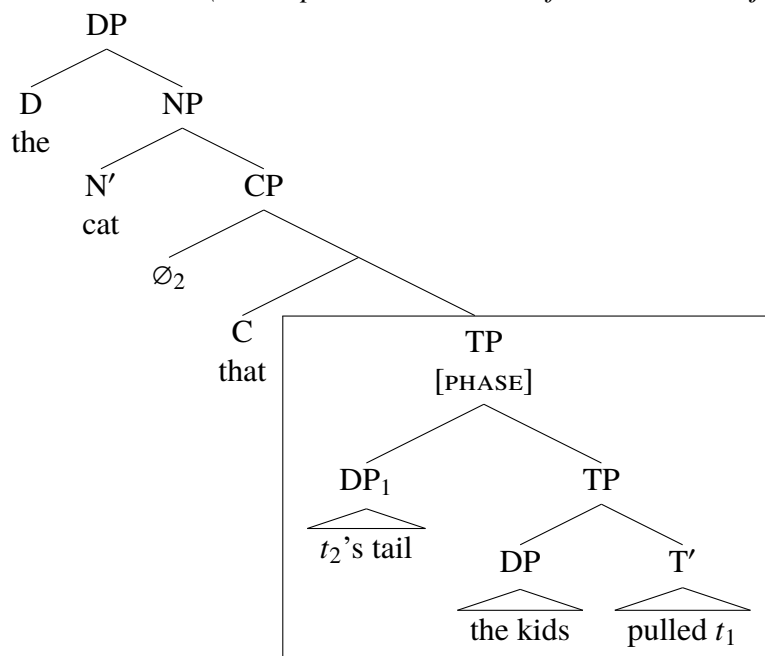
Notice that in *that's*-relatives, the relativizing *wh*-operator that moves from possessor position is null. Since an overt C *that* is present in such contexts, we indeed expect the *wh*-operator to be null, given the Doubly Filled Comp Filter. If in such structures the complementizer is not realized and the possessor operator is overt instead, then the possessor and the possessum will simply appear adjacent as they do in relative clauses of the usual variety, as we see in (77):[31]

(77)   *Alternative resolution of the Doubly Filled Comp Filter*

       a.   *Subject that's-relative structure*

---
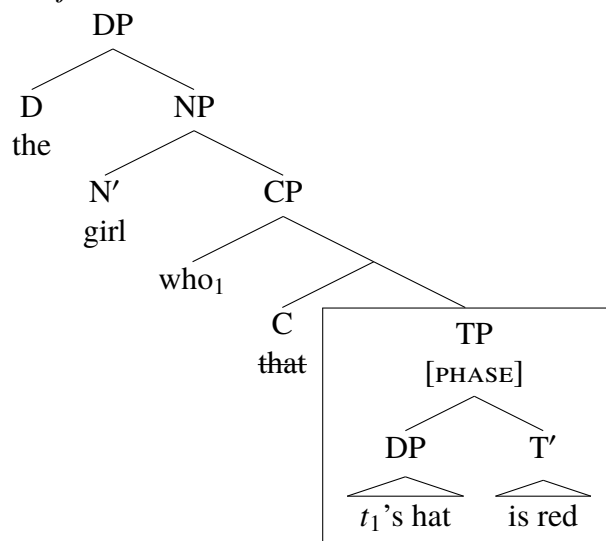
[31]It is possible to have a non-subject relative clause where C and the relativizing *wh*-phrase in its specifier are both silent:

(i)   This is the cake [(which/that) I ate].

However, this possibility does not extend to *that's*-relatives:

(ii)   This is the person [*(who/that)'s friend I like].

But notice that the possibility of a null CP layer also does not extend to subject relatives, a fact partially described by the anti-*that*-trace effect:

(iii)   This is the woman [*(who/that) likes cats].

Subject relatives presumably involve movement of a relativizing operator from spec-TP. Similarly, under the analysis of this section, all *that's* relatives involve movement of an operator from some constituent within TP—either from the subject (74), or from a non-subject pied-piped to an outer spec-TP (76). Thus it could be the case that an empty CP is generally banned for relativization from within TP. If correct, this observation still requires an explanation, however.

b.   *Non-subject that's-relative structure*
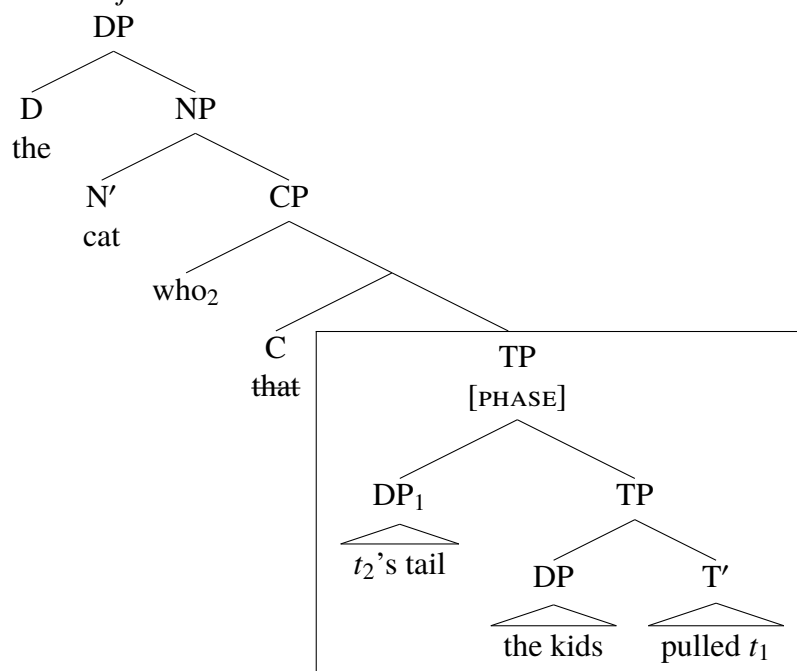
Note that this analysis of *that's*-relatives maintains the phase-bounded adjacency requirement of [*'s*] that I've argued generally constrains English PE. This requirement's interaction with CL forces PE to strand the possessum in the linear edge of a phase crossed by the possessor's movement. Ultimately this is the CP edge in the cases examined earlier in this chapter, since the vP edge is unavailable. If TP in relative clauses is also a phase, then in this context the TP edge provides another position where the possessum can be stranded, as in (74) and (76).[32] [33]

The analysis of this section makes the prediction that the TP edge in relative

---

[32]At this point, we have seen evidence for possessum stranding in the edge of CPs and certain phasal TPs. Den Dikken (2007) argues that constituents instating subject-predicate structures ("small clauses") should also be considered phases, a conclusion that Ko (2011, 2014) defends in a CL context. If such constituents are indeed phases, English PE should allow possessum stranding in their edge. A potential verification of this prediction is instantiated in the examples in (i), but individuals I have asked agree that such examples are not readily acceptable:

(i)   a.   ?? I know who$_1$ you consider [$_{PredP}$ [t_1's family] really annoying].
       b.   ?? Who$_1$ did you find [$_{PredP}$ [t_1's cake] the tastiest] at the baking contest?

[33]An additional puzzle is that while typical relative clauses can be coordinated (ia), the same does not seem to be possible for *that's*-relatives (ib):

(i)   a.   This is the guy [[$_{CP}$ whose cat is big] and [$_{CP}$ whose dog is small]].
       b.   * This is the guy [[$_{CP}$ ∅ that's cat is big] and [$_{CP}$ ∅ that's dog is small]].

clauses should be a licit position for intermediate stranding of other varieties. For instance, if relativizing *wh*-movement in West Ulster English is capable of *all*-stranding, we expect the possibility of *all*-stranding in the TP edge in this dialect of English. I do not know whether this prediction is correct. However, it is likely that an instance of stranding in some language will make it possible to test intermediate stranding in the edge of relative TP.[34] I will leave this as a topic for future work.

7   On the phasehood of DP

Much work in contemporary syntax hypothesizes that DP is a phase (Heck and Zimmermann 2004; Bošković 2005, 2014; Newell 2008; Syed and Simpson 2017, a.o.). This concept is natural in light of the fact that in many languages, extraction from DPs must take advantage of their edge (Szabolcsi 1984; Gavruseva 2000; Coon 2009; Citko 2014, a.o.). However, various works contest that DPs nevertheless do not qualify as phases (Matushansky 2005; Sabbagh 2007; Chomsky et al. 2019; Zyman To appear). While later chapters of this dissertation will require the hypothesis that DPs are phases, the analysis of English PE defended in this chapter is not straightforwardly compatible with this concept.

I have proposed that the adjacency requirement of [*'s*] can be satisfied in its local phase, thus allowing the possessor to separate from [*'s*] provided that [*'s*] remains within a phase where its adjacency requirement was met at spell-out. If DP were a phase in of itself, this requirement of [*'s*] would be immediately satisfied within DP. In this situation, there would be no need to ever pied-pipe under PE, incorrectly predicting the possibility of allowing a non-subject possessum DP exited by PE to remain in situ in VP (78):

This fact is just as puzzling for an account under which the *that* of *that's*-relatives is in fact a relative pronoun, since under such an analysis both of these examples should be completely structurally identical. I do not have a solution for this puzzle.

[34]The strandable adjuncts in English discussed in section 2.2 of chapter 2 are only compatible with interrogative *wh*-phrases, and thus cannot be used to test this prediction about relative clauses.

(i)　　a.　　This is the place [where$_1$ (*exactly) I said t_1 (*exactly) I wanted to go t_1 (*exactly)].
　　　　b.　　This is the amount of flour [which$_1$ (*to the nearest pound) I said t_1 (*to the nearest pound) I ordered t_1 (*to the nearest pound)].

(78)   *Incorrectly predicted if DP is a phase: PE from in situ non-subject*
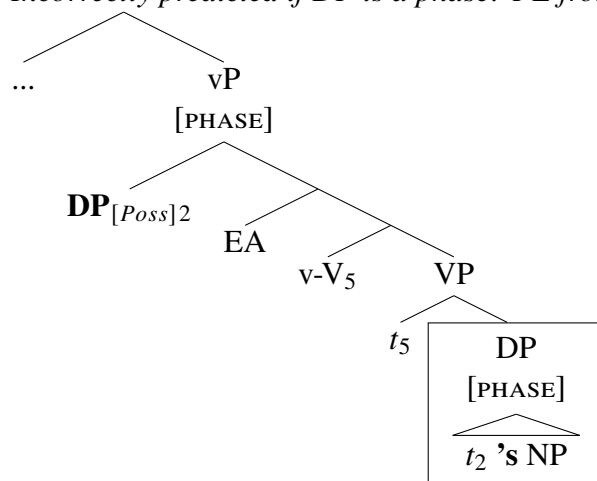
The most straightforward way of interpreting this result would be to conclude that the English DP is not a phase. This result may be also be compatible with a theory in which the English DP is a phase for LF, but not PF (entailing that DP is not an independent domain for linearization), a possibility that arises if a given phase need not necessarily transfer its contents to both interfaces. See Citko (2014) for an overview of works that make use of such a concept.

A phenomenon that may provide an independent diagnostic for DP phasehood is the stranding of adjuncts like *exactly / precisely* under *wh*-movement, which we have already encountered in the previous chapter. As we've seen, such adjuncts can be stranded in their base position, or in an intermediate CP edge:

(79)   Exactly-*stranding*
       What$_1$ (exactly/precisely) did you say t_1 (**exactly/precisely**) that she wants t_1 (exactly/precisely)?

If DP is a phase, and therefore a constituent that movement must exit via the edge of, then such stranding should be possible in the edge of DP. However, as Zyman (To appear) observes, this is not so (80):

(80)   *No* exactly-*stranding in the edge of DP*
       What$_1$ (exactly/precisely) did you write [$_{DP}$ t_1 (***exactly/*precisely**) a book about t_1]?

This chapter's results are compatible with there being a nominal phase boundary below the possessor and [*'s*], though not above them.[35] No problem would be posed

---

[35]This fact is reminiscent of Chomsky's approach to phases, in which edges are exempt from spell-out, since in this case spell-out of DP would exclude D and the possessor. However, as discussed

by the phasehood of NP alone, for instance:

(81) *A compatible proposal: Phasehood of NP but not DP*

```
                    DP
               ╱         ╲
         DP_Poss    ╱        ╲
                  D  ┌─────────────┐
                  's │    NP       │
                     │  [PHASE]    │
                     │    ╱╲       │
                     │  N    ...   │
                     └─────────────┘
```
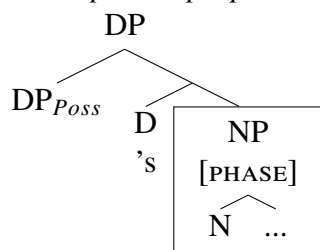
As discussed in chapters 4 and 7, there is precedent for the proposal that NP is a phase.[36] This does not preclude the possibility of DP being a phase as well, however.

We will see that the arguments of chapter 4 depend on the concept that the Russian nominal phrase is a phase. As discussed in that chapter, while some works take the nominal phrases of such article-less Slavic languages to be bare NPs, other works argue that such languages have full DPs at least some of the time. If this is so, then given the results of chapter 4, it would be necessary to posit that DP in Russian is a phase. Furthermore, we will also see that the analysis of extraposition in chapter 6 requires the proposal that DP is a phase in English, and a variety of other languages. Given that the phasehood of DP is thus not easily dispensed with, next I will offer a unification of this chapter's analysis of English PE with the proposal that DP is a phase.[37]

7.1 Towards a reconciliation with the phasehood of DP

When representing linearization information, I have followed Fox and Pesetsky (2005a,b) in using the symbol "<" to encode that one element precedes another. If the elements related by "<" are phrases containing multiple terminals, however, then it is necessary to be more precise about what "<" indicates. Fox & Pesetsky offer the formulation in (82), which aside from a qualification necessary to exclude traces from linearization,[38] states what "<" encodes when relating complex constituents:

---

in section 4.3, such a theory is fundamentally incompatible with this chapter's results.

[36]If the English NP is a phase, we predict that movement from it must pass through its edge. In chapter 7, we will see evidence from parasitic gaps in relative clauses that such movement is possible.

[37]One simple resolution to this conflict would be to take the English DP to be a phase, but to assume that [*'s*] and the possessor are actually adjoined outside the possessum DP. For instance, it could be that [*'s*] is not D, but a functional element whose specifier is the possessor and whose complement is the possessum (which [*'s*] would evidently have to require to have a null D in this case). I am aware of no evidence for such a proposal, however.

[38]Fox and Pesetsky (2005b) later go on to redefine "<" in the context of a theory of movement as multi-dominance. I do not pursue such an approach here.

(82)   *Definition of the relation "<"*
An ordering statement of the form $\alpha < \beta$ is understood by PF as meaning that the last element dominated by α and not dominated by a trace precedes the first element dominated by β and not dominated by a trace.
(Fox & Pesetsky 2005b, ex. 11)

Assume that when a given phase contains another, spell-out triggered by the higher phase does not redundantly re-linearize the embedded one, but rather simply makes reference to its first/last element (modulo traces) in the way (82) states. This allows the lower phase to be ordered with respect to the containing one: if all content of a higher phase precedes the first element of an embedded phase, then it is necessarily the case that all content in the former phase precedes everything in the latter.

Consider this proposal about the linearization of embedded phases in the context of a sentence with PE from an object possessum. It will suffice to examine the vP level of such a derivation, as shown in (83) below. First, here linearization of DP establishes that *who* is the leftmost element in it. Second, *who* moves to the vP edge. Third, spell-out of vP then determines that *who* precedes the subject, that the subject precedes V, and that V precedes the first element of the DP phase. Given that traces of overt movement (or their equivalent in whatever theory of movement one assumes) must be ignored by linearization, the linearly first element of DP is [*'s*], given that the possessor has moved away by the time this vP spells-out.

(83)   *Cyclic spell-out in a vP with PE from an object*
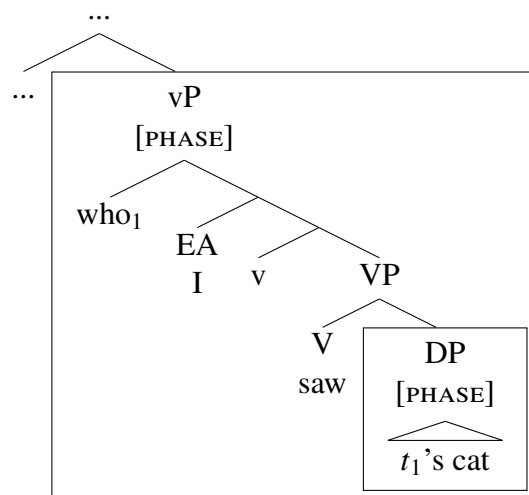
1. Linearization of DP: who < 's < cat
2. Linearization of vP: who < I < saw < **'s**

In this derivation, the adjacency requirement of [*'s*] would be satisfied at the DP

level, since when DP spells-out the possessor is directly adjacent to [*'s*]. However, I suggest that because spell-out of vP makes reference to [*'s*] once more given that at that time it is the leftmost element of DP, the adjacency constraint is re-evaluated. Since at the time vP spells-out [*'s*] is separated from the possessor by the subject and verb, re-evaluation of the adjacency constraint finds this configuration deviant.

This line of reasoning accurately predicts that PE cannot occur in this case, with pied-piping of the possessum being the only option. This is precisely the result argued for in section 5 above, but here we have allowed the necessary result to emerge even with DP as a phase, by allowing the linearization of vP to make reference to the first element of the DP (modulo traces) after PE has occurred.

We have seen that for independent reasons a possessum DP pied-piped into spec-vP must move further, and thus cannot actually separate from the possessor until reaching the edge of an embedded CP. I propose that the adjacency requirement of [*'s*] does not force further movement of the possessum from within CP post-extraction because at that point in the derivation, [*'s*] is separated from the extracted possessor by an additional phase boundary, as shown in (84) below. Here spell-out of the matrix vP post-PE establishes that the matrix V precedes the first element of CP, which is the pied-piped possessum DP. I hypothesize that since spell-out of a given phase only makes reference to whatever the first element of a given embedded phase is, and nothing more, in this situation spell-out of the matrix vP does not make reference to the [*'s*] contained within that DP. It simply sees that this DP is the first element of CP, which is sufficient to correctly establish that this CP will follow the matrix V at the time when these ordering statements are combined and evaluated by PF.[39] Having not directly made reference to [*'s*], its adjacency constraint is not evaluated by spell-out of the matrix vP, and thus the fact that the possessor has moved away causes no violation.

---

[39]This would follow if we specify that the definition of "<" refers specifically to immediate dominance, not dominance in general.

(84)   *Cyclic spell-out of DP, embedded CP, and matrix vP with PE*
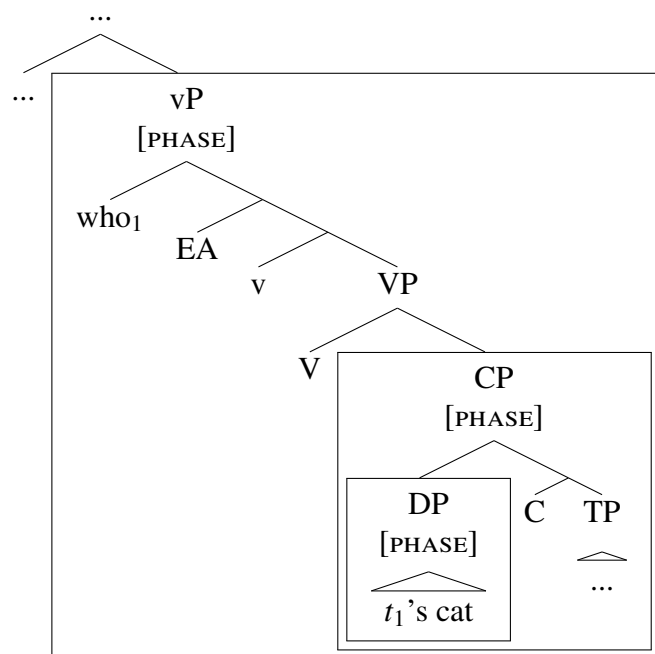
1. Linearization of DP: who < 's < cat
2. Linearization of CP: cat < C < ...
3. Linearization of matrix vP: who < EA < V < DP

If the only relevant elements for linearization are terminal nodes, then spell-out of the matrix vP here might necessarily make reference to ['s], since this is the leftmost overt terminal in CP in (84). However, in this case, the adjacency constraint should be invoked once more, forcing further pied-piping. As the facts have shown, this is not the correct result. Thus I hypothesize that in (84) linearization of the matrix vP can simply make reference to the presence of DP as the first element of CP, without re-linearizing any content within that DP. This allows ['s] to go un-noticed by the linearization of the matrix vP, permitting the possessor to thus move on freely.[40]

---

[40]In essence, the account suggested here predicts that PE that crosses two phase boundaries, as in (84), should not require possessum pied-piping. However, if DPs are phases, this hypothesis predicts the possibility of extracting the possessor of a possessor of a non-subject DP, since such extraction would cross two DP nodes. Such PE is not acceptable in reality, as shown in (i). This is in correspondence with the generalization demonstrated above that English PE cannot strand a non-subject argument in situ:

(i)   a.   * Who$_3$ did you say you saw [$_{DP2}$ [$_{DP1}$ t_3's friend]'s cat] in the yard?
        b.   * That's the guy [who$_3$ we found [$_{DP2}$ [$_{DP1}$ t_3's kid]'s hat] in the mailbox].
        c.   * Let me tell you [who$_3$ I ate [$_{DP2}$ [$_{DP1}$ t_3's child]'s entire birthday cake] last week].

8 Final considerations

Before concluding the chapter, in this section I discuss several additional but less vital considerations for the examination of English PE. First, I defend this chapter's analysis of non-subject displacement under PE as derived by pied-piping along with successive-cyclic A′-movement of the possessor, by showing that such displacement is not the result of independent topic/focus fronting prior to extraction. Second, I offer some final comments on the nature of the adjacency requirement of [*'s*]. Third, I describe a puzzle about PE from infinitives that remains unsolved. Fourth and finally, I comment on the relationship of these PE facts to PE in child speech, as reported by Gavruseva and Thornton (2001).

8.1 Possessum displacement is not topic/focus movement

In section 2, following Gavruseva and Thornton (2001) I adopted the hypothesis that the displacement of non-subjects to the embedded CP edge under PE is caused by pied-piping along with A′-movement of the possessor. Alternatively, we might hypothesize that this displacement is caused by the non-subject possessum undergoing a step of topic/focus movement into the left periphery of the embedded CP prior to PE, since such movement is independently possible for many English speakers. Since it is only vital to this chapter's analysis that the possessum reaches the embedded CP edge, either of these accounts could achieve correct results. Thus this choice does not affect the central findings of this chapter. However, there are reasons to believe that a topic/focus movement analysis is not correct.

For speakers with PE that independently accept topic/focus movement, the semantic/pragmatic effect of such movement is absent in PE examples. For instance, consider (85a) below, where the fronted phrase *John's cake* is associated with a particular information structure: either setting up John's cake as the topic of the clause, or focusing on John's cake in contrast to other alternatives. Speakers who accept (85a) and allow PE report that such information structure is absent from PE sentences with a displaced non-subject like (85b). Rather, such PE examples have a neutral information structure (modulo question-hood), just like their full pied-piping

---

Thus the account suggested above is incomplete. The path to a solution for this problem is not particularly clear. One option might be to modify the structure assumed for possession. Another would be a more nuanced understanding of the circumstances under which DPs count as phases. For instance, if only the outermost DP node in a recursive DP structure counts as a phase, then the extraction in (i) above crosses only one DP phase boundary, and thus the above account ceases to incorrectly predict the acceptability of this configuration. This proposal about phasehood in a recursive DP resembles arguments about the contextuality of phasehood in, for instance, Bobaljik and Wurmbrand (2005); Den Dikken (2007); Despić (2011); Bobaljik and Wurmbrand (2013); Wurmbrand (2013); Bošković (2016), and references therein.

equivalents do (85c).[41] Furthermore, some speakers I have interviewed do not readily accept topic/focus movement in the spoken language (and thus do not accept examples like (85a) below) but nevertheless accept PE examples like (85b). For such speakers, it is especially clear that PE and topic/focus movement are unrelated.

(85)   a.   *Topic / focus movement*
           [John's cake]$_1$, you said they really liked t_1.

        b.   *PE from non-subject*
           Who$_2$ did you say [[t_2's cake]$_1$ they really liked t_1]?

        c.   *Pied-piping equivalent of (91b)*
           [Whose cake]$_1$ did you say [they really liked t_1]?

If the topic/focus movement analysis is incorrect, the relevant sentences with PE from a non-subject should be accepted even with matrix verbs that ban embedded topicalization. Hooper and Thompson (1973) show that anti-factive/non-assertive verbs like *doubt* and *deny* ban embedded topicalization (among other "root transformations" like VP fronting). In fact, such matrix verbs degrade PE, as (86) shows by comparing a deviant sentence with PE from a non-subject (86a) with its pied-piping equivalent (86b).

(86)   *PE is blocked by anti-factive verbs*
        a.   * Who$_2$ do you **deny/doubt** [[t_2's cat]$_1$ they stole t_1]?
        b.   [Whose cat]$_1$ do you **deny/doubt** [they abused t_1]?

If the unavailability of embedded topic/focus movement were the problem in (86a), then such examples should be permitted with a factive matrix verb, like *realize, know*, or *see*, embedding verbs which Hooper and Thompson show allow embedded topicalization. However, PE with such factive embedding verbs also fails:

(87)   *PE is blocked by factive verbs*
        a.   * Who$_2$ did you **know/realize/see** [[t_2's cat]$_1$ they abused t_1]?
        b.   [Whose cat]$_1$ did you **know/realize/see** [they abused t_1]?

Both non-assertive verbs like *doubt* and *deny* in (86) and the factive verbs in (87) are classified as weak island inducers (see Szabolcsi and Lohndal (2017) for a recent overview). Weak islands are constituents classically defined as penetrable to movement of arguments but not adjuncts. This set also includes *wh*-islands caused by *whether*, which interfere with PE as well:

---

[41]The prosody typical of topic/focus movement is also absent from examples like (85b).

(88) *PE is sensitive to the* whether *island*

   a. * Who$_2$ do you wonder [**whether** we said [[t_2's friend]$_1$ we like t_1]]?
   b. [Whose friend]$_1$ do you wonder [**whether** we said we like t_1]?

NPI licensers like *few*/*only* also induce weak islands, and degrade PE (89a vs. 89b). A similar effect occurs with basic negation (89c vs. 89d). While the violations in (89) are not absolute, these contrasts are evident for many speakers.[42]

(89) *PE is degraded by NPI licencers*

   a. [Whose art]$_1$ have **few people/only they** said [t_1 is interesting]?
   b. *??% Who$_1$ have **few people/only they** said [[t_1's art] is interesting]?
   c. That's the author [[whose work]$_1$ I did**n't** say [t_1 is any good]].
   d. ?% That's the author [who$_1$ I did**n't** say [[t_1's work] is any good]].

Overall, the hypothesis that PE (from non-subjects) depends on topic/focus movement fails to predict the actual set of circumstances where English PE is degraded, which fits the set of weak islands. This fact suggests that English possessors are fundamentally not arguments. English PE is in essence an instance of left branch extraction (Ross 1967), a type of extraction that is known to be sensitive to weak islands in several other western European languages. For instance, *combien*-split in French (Obenauer 1984) and *wat voor / wat aan*-split in Dutch (Corver 1990; Beermann 1997; Honcoop 1998) are subject to the same constraint. Thus the weak island sensitivity of English PE falls within the purview of known constraints on extraction of this variety.

8.2   On the adjacency requirement of [*'s*]

Following works proposing that PE in mainstream English is banned because [*'s*] must be adjacent to the possessor at PF, I hypothesized that the constrained nature of English PE emerges if speakers allow this requirement to be enforced within a structurally smaller domain. Namely, I hypothesized that while this constraint is globally enforced in English grammars without PE, those speakers with PE can satisfy this constraint at the minimal phase level. If such requirements can potentially be locally evaluated in this way, we expect to see phenomena in other languages whose distribution is analogous to that of English PE: that is, we expect to find circumstances where something like a bound morpheme can only be stranded after some movement. So far, I have not found additional facts that fit this description. If

---

[42]Examples (89b/d) are best with focus on the intervener. Focus, among other semantic and pragmatic factors, is known to circumvent the (weak) island-hood of negation and related elements, as several works have shown (Kroch 1989; Kuno and Takami 1997, a.o.).

such patterns are not clearly attested, it may be that the English facts studied here are atypical and therefore perhaps unstable. If this is so, we may expect subsequent generations to re-arrange, or fail to acquire, the patterns reported here. Provided that this chapter's arguments for the localized adjacency requirement of ['*s*] are successful, understanding it more deeply should be a topic for further research.

Adjacency between ['*s*] and the possessor it selects can be motivated by the *contiguity theory* of Richards (2016), which hypothesizes that syntactic dependencies like agreement and selection require linear adjacency of the elements involved in many cases, due to prosodic requirements. Importantly, Richards (2016) allows such adjacency to be satisfied derivationally, and later broken. While this concept is reminiscent of this chapter's conclusion that speakers with PE permit extraction of the possessor after earlier satisfaction of adjacency, it does not explain why such eventual breaking of adjacency is available to only some speakers. It is possible that grammars differ in how the relevant adjacency requirement is prioritized. This hypothesis can be implemented using the alignment constraints of Optimality Theoretic phonology (McCarthy and Prince 1993, a.o.). In Optimality Theory, all constraints are in principle violable, depending on how they are ranked relative to each other. For speakers without PE, the hypothesis would thus be that adjacency of the possessor and ['*s*] is enforced by a highly ranked constraint, which is never violated. In contrast, speakers with PE possess a colloquial grammar in which this adjacency constraint is violable, though not utterly, since it must be satisfied for a minimal part of the derivation.[43]

If this adjacency requirement is in principle violable depending on language-internal constraint ranking, we predict a third sort of English grammar—one in which this adjacency constraint is prioritized so low that it asserts no influence. There is preliminary evidence for such a variety. An anonymous reviewer for a previous version of this work points out that in a highly colloquial register, they depart slightly from the judgments reported above in tolerating PE with no pied-piping, stranding an object in situ:

(90)   *PE with no pied-piping*
        (Anonymous reviewer, p.c.)

        a.   %/? So tell me, [which girl]$_1$/who$_1$ didja/'dja meet [t_1's friend] yesterday?

        b.   %/? So tell me, [which girl]$_1$/who$_1$ d'you think he met [t_1's friend] yesterday?

---

[43]Allowing adjacency to be broken only after some minimal satisfaction is also evocative of the Principle of Minimal Compliance of Richards (1998), though it is not obvious that such a principle is relevant to the configurations under consideration here.

Thus it may be the case that the relevant adjacency requirement can be of variable importance from one grammar to another. This is as expected if this is indeed a PF constraint, under the Optimality Theoretic hypothesis that phonological constraints are subject to potentially variable relative ranking.

8.3 A puzzle about infinitives

This chapter's analysis predicts that English PE should be able to strand a possessum in the edge of any phase passed through. Generally this is CP, vP being independently unavailable for stranding. The analysis of *that's*-relatives in the previous section also shows evidence for stranding in the edge of TP in relative clauses, consistent with the proposal of Deal (2016) that such TPs are phases. Beyond the phases already discussed, some works take infinitival clauses to be phases as well (Wurmbrand 2013, 2017, a.o.). If this proposal is correct, it is unsurprising that for many speakers, English PE from a subject possessum can strand it in what appears to be the edge of an infinitival complement:

(91)   *PE from subject of infinitival complement*
    a.   Tell me who$_1$ you believe [$_{TP}$ [t_1's idea] to be the best].
    b.   Who$_1$ do you expect [$_{TP}$ [t_1's horse] to win the race]?
    c.   This is a guy who$_1$ I consider [$_{TP}$ [t_1's work] to be a bit crappy].

However, a puzzling fact I cannot offer a decisive solution to here is that a non-subject possessum cannot also be stranded in the edge of an infinitival complement, as we see in (92). Note that the configurations in (92) are all acceptable provided that the possessum is pied-piped rather than stranded:

(92)   *No PE with stranding of non-subject in edge of infinitive*
    a.   * Tell me who$_2$ you believe [$_{TP}$ [t_2's friend]$_1$ him to dislike t_1].
    b.   * Who$_2$ do you expect [$_{TP}$ [t_2's horse]$_1$ Mary to cheer for t_1]?
    c.   * This is the guy who$_2$ I've found [$_{TP}$ [t_2's jokes]$_1$ my kids to hate t_1].

If the subjects of infinitives in fact must raise to an object position within the higher vP (Postal 1974, a.o.), and if infinitival TPs are indeed phases, then perhaps the problem in (92) is that *wh*-movement of the possessor through the linear edge of the infinitive strands the non-subject possessum in an outer specifier of the phasal TP, which prevents raising of the subject from proceeding via the linear edge of this TP in the way CL would require. A string that would be created by such illicit raising from the non-edge of the infinitival TP is illustrated in (93):

(93)   *Hypothesis: No raising across stranded possessum in outer spec-TP*
        * Tell me who$_3$ you believe **him**$_2$ [$_{TP}$ [t_3's friend]$_1$ t_2 to dislike t_1].

This account accurately predicts that this structure should be acceptable if the non-subject possessum is pied-piped beyond the edge of the infinitival TP to a position above the embedded subject (94). Such movement reverses the linearization contradiction that subject raising from the non-edge in (93) would have caused:

(94)   *PE without non-subject stranding in the edge of the infinitive*
        ? Tell me who you$_3$ think [t_3's friend]$_1$ they believe **him**$_2$ [$_{TP}$ t_1 t_2 to dislike t_1].

This cannot be the end of the story, however. If the subject of infinitival complements truly must raise into the matrix vP, then it is unclear what the status of the subjects in (91) above is. These subjects are remnants of PE, which I have argued must surface at the linear edge of a phase. However, if these subjects raise from the infinitival clauses they originate in, it is unclear whether they can be said to inhabit a phase edge. Since any raising here would be to a post-verbal position in the matrix clause, raising into the vP edge is not a tenable analysis. While it could be the case that these subjects inhabit the edge of some phasal sub-constituent of the vP, I lack independent evidence to motivate this analytical direction.[44] [45]

---

[44]Another possibility is that overt raising of the subject into the matrix clause in these configurations is in fact optional, and that (92) is unacceptable because the material stranded in the outer specifier of TP somehow blocks case assignment to the embedded subject by the matrix v, unless that material moves on (94). This would be predicted by a theory like Bošković (2016) in which a higher specifier of a phase must move away for a lower specifier of that phase to be accessed (a theory I will criticize in the next chapter).

[45]McCloskey (2000) shows that (at least with the embedding verb *expect*) stranded *all* in West Ulster English cannot occupy a position within an infinitive preceding the embedded subject (i), a fact which is analogous to the unacceptable examples in (92) above:

(i)   * <u>Who</u> did you expect <u>all</u> **your mother** to meet at the party?
         (McCloskey 2000, ex. 40b)

McCloskey also shows that stranded *all* can end up in a position following the embedded subject (ii), which is impossible for English PE (iii). I cannot offer an explanation for this difference.

(ii)   <u>Who</u> did you expect **your mother** <u>all</u> to meet at the party?
         (McCloskey 2000, ex. 39b)

(iii)   * <u>Who</u>$_2$ do you expect [$_{TP}$ **her** [t_2's horse]$_1$ to cheer for t_1]?

8.4   PE in child speech and the acquisition process

The distribution of PE in child speech that Gavruseva and Thornton (2001) report is highly similar to that of PE in adult speakers as described here, in a few important ways: child PE both only occurs in long-distance movement contexts, and also overwhelmingly requires a non-subject possessum remnant to end up stranded in an intermediate CP edge. Gavruseva & Thornton argue that children are capable of PE (but adults aren't, they assume) because children have not yet acquired the PF knowledge that would motivate them to always keep the possessor and [*'s*] adjacent. However, given that child and adult PE share the same basic distribution, if the analysis of this chapter is right, it would seem that children also have the localized requirement of adjacency to [*'s*] that I have posited here for adult PE.

   Children at the developmental stage in which PE is frequent also allow extraction of *whose*, as Gavruseva & Thornton show. The possibility of *whose*-movement in combination with the phase-bounded requirement of adjacency to [*'s*] argued for in this chapter makes an accurate prediction about child PE: Children should be able to move *whose* alone to an intermediate spec-CP, and subsequently extract *who*, with the result that [*'s*] will be stranded alone in the intermediate CP edge. Gavruseva & Thornton show that children in fact produce such examples (95).

(95)   *Intermediate stranding of ['s] in child PE*

   a.   **Who**$_2$ do you think [$_{CP}$ [t_2 **'s**]$_1$ Spiderman saved [t_1 **one of these cats**]]?

   b.   **Who**$_2$ do you think [$_{CP}$ [t_2 **'s**]$_1$ Spiderman saved [t_1 **cat**]]?

   c.   **Who**$_2$ do you think [$_{CP}$ [t_2 **'s**]$_1$ Gonzo rubbed down [t_1 **horse**]]?
        (Gavruseva & Thornton, ex. 42)

Thus the existence of such examples in child speech does not violate the hypothesis that adults and children share essentially the same phase-bounded adjacency requirement argued for here. If anything, it confirms this view, given that the lack of *whose*-movement in adult speech is an independent concern.

   If the same PF principle constrains both adult PE and child PE, we can conclude that what differentiates adults with PE from those that do not allow it is that the former group preserves a register with a child-like version of the adjacency constraint, which is abandoned by the latter set. More interesting, however, is that there is no evidence for a developmental stage at which children posit no adjacency requirement: if this were so, slightly younger children than the ones under consideration would be expected to show utterly unconstrained PE. I am aware of no evidence that this is the case. Thus under the analysis of this chapter, it must be that children start out with the locally-defined requirement, before (in many cases) advancing to a purely

global one later on. I will leave further consideration of this topic to future work.

9 Conclusion

This chapter has described and analyzed the complexities of PE in colloquial English, a little-studied possibility for many speakers. I have shown that English PE is subject to a set of intricate but robust limitations, which can be described as follows:

(96) *Generalization about English PE*
      A possessor must reach the left linear edge of the local CP before extracting from the possessum DP.

I have argued that this phenomenon's distribution serves as evidence for the CL theory of phases, which constrains English PE via its interaction with a phase-level version of an independently proposed PF requirement of [*'s*]. The ban on phrase-bounded specifier-to-specifier movement was also important for the analysis of this chapter, as it was for the previous. Along the way, this analysis has engaged with many topics relating to the nature and location of phases. These concepts will also feature in the next chapter, which continues to explore sub-extraction from nominal phrases, with a focus on Russian.

Chapter 4

Limiting discontinuity in Russian

1   Introduction

In the previous two chapters, I have argued that CL and certain locality constraints on movement together explain several word-order limitations on sub-extraction. In this chapter, I examine how the concepts defended so far play out in Russian—a language ideal for testing the predictions of the framework developed here, given its great richness in word-order-manipulating movement processes.

This chapter focuses on a variety of sub-extraction termed left branch extraction (LBE). LBE refers to extraction of an element that originates in the left periphery of the nominal domain, a process that is productive for many languages in the Slavic family, and beyond (Ross 1967; Corver 1990, 2007, a.o.). In Russian, LBE signals a contrastive topic / focus interpretation for the displaced element (Pereltsvaig 2008), and provided that an appropriate context is salient, LBE of any left branch constituent (adjective, quantifier, demonstrative, possessor, and so on) is in principle possible. Some preliminary examples of Russian LBE are provided in (1):[1]

(1)   *Some examples of LBE in Russian*

    a.   **Kakuju**$_1$ ty   uvidel [$_{NP}$ t_1 košku]?
           what   you saw   cat
           'What cat did you see?'

    b.   **Ètu**$_1$ ty   uvidel [$_{NP}$ t_1 košku].
           this   you saw   cat
           'You saw THIS cat.'

---
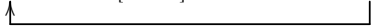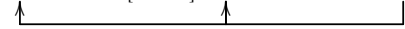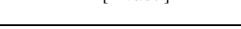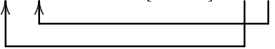
[1]In Russian, all nominal phrases bear person / number / gender / case features, which are referenced by their modifiers and by verbal agreement. I will generally omit this information, however, unless it is useful to mention. Glossing these features will sometimes be useful for clarifying the structure intended for sentences involving a great deal of discontinuity, for instance.

c. **Miluju**$_1$ ty   uvidel [$_{NP}$ t_1 košku].
cute   you saw   cat
'You saw a CUTE cat.'

In this chapter, I integrate and extend the findings of two collaborative projects on Russian (Antonenko and Davis 2020; Bondarenko and Davis To appear), which when taken together reveal some general constraints on Russian sub-extraction that align directly with the concepts defended so far in this dissertation. In particular, I will argue that the facts presented here reveal order preservation effects in the nominal domain as well as the vP, and thus indicate the phasehood of these constituents, while at the same time supporting the correctness of the CL approach. As in previous chapters, the interaction of CL with independent syntactic constraints on when movement can occur will also be vital to predicting the attested patterns.[2]

1.1   Main results

In the preceding chapters, we have seen numerous manifestations of two general limitations on cross-phasal movement predicted by CL—that movement from a phase must either pass through its linear edge (2), or must involve additional order-restoring movement in the case of a non-successive-cyclic phase exit (3). Both of these scenarios yield contradiction-free linearization instructions, which we expect to be required if what governs cross-phasal movement is a need for phase-by-phase spell-out to culminate in a coherent PF representation.

(2)   *Phase exit via linear edge*
  a.   * [$_{ZP}$ α Z [$_{YP_{[Phase]}}$ β Y [$_{XP}$ t]]]
  b.   ✓ [$_{ZP}$ α Z [$_{YP_{[Phase]}}$ t β Y [$_{XP}$ t]]]

(3)   a.   *Illicit non-successive-cyclic phase exit...*
        * [$_{ZP}$ β Z [$_{YP_{[Phase]}}$ α t Y ...]]
  b.   *...repaired by additional order-restoring movement*
        ✓ [$_{ZP}$ α β Z [$_{YP_{[Phase]}}$ t t Y ...]]

In this chapter, we will see several circumstances in Russian for which the scenarios schematized in (2) and (3) clearly assert their influence.

---

[2]The facts reported in this chapter are the aggregate of judgments provided by my collaborators Andrei Antonenko and Tanya Bondarenko, as well as by five other native Russian speakers interviewed during this research.

In the first part of this chapter, we will see evidence for these concepts from constraints on LBE from nominal phrases with multiple left branches, expanding on facts observed in Antonenko and Davis (2020). In brief, while LBE is unrestricted for nominal phrases with only one left branch, LBE of a left branch that is preceded by another within the same nominal is typically banned, as previewed in (4). Here we see that LBE of an adjective across a demonstrative is unacceptable, since demonstratives originate preceding adjectives in Russian:

(4)   *Adjective LBE cannot cross co-occurring demonstrative*
      * **Bol'šogo**$_2$ ja videl [ètogo t_2 kota].
         big   I saw this   cat
      'I saw this BIG cat.'

However, if a left branch crossed over by LBE from the non-edge also extracts, the result is acceptable, as (5) shows. Importantly, the final order of the two left branches must match the relative order that they had in their original positions. Thus as we see in (5a), order-preserving LBE of both a demonstrative and adjective is acceptable, while order-reversing LBE of these elements is not (5b):

(5)   *Non-edge LBE repaired by order-restoring second LBE*

      a.   ✓ Ètogo$_1$ **bol'šogo**$_2$ ja videl [t_1 t_2 kota].
             this   big   I saw   cat
           'I saw this big cat.'

      b.   * **Bol'šogo**$_2$ ètogo$_1$ ja videl [t_1 t_2 kota].
             big   this   I saw   cat
           'I saw THIS BIG cat.'

As we will see, patterns like this are pervasive for Russian LBE, as described by the following empirical generalization:

(6)   *LBE order preservation generalization*
      LBE from a Russian nominal phrase containing multiple left branches must result in a relative linear ordering that matches the original order that those left branches had prior to any extraction.

I argue that this result is correctly predicted by hypothesizing that the Russian nominal phrase is a phase, and thus under CL, a domain to which order preservation applies. Familiar principles of locality will also be central to understanding this generalization. Similar patterns in Serbo-Croatian have recently been examined under very different assumptions by Bošković (2016), with which I will compare the analysis presented here, ultimately arguing in favor of the CL approach.

In the second part of this chapter, I discuss the interaction of scrambling and LBE from subjects in Russian, building from work in Bondarenko and Davis (To appear).[3] While LBE from Russian subjects is generally possible, such extraction fails when combined with scrambling of other material that originates in the same vP. This fact, involving a scrambled object, is previewed in (7):

(7) *No object scrambling + LBE from subject*
     * **Každaja**$_2$ <u>kota</u>$_1$ [t_2 **devočka**] pogladila t_1.
     every.NOM cat.ACC **girl.NOM** stroked

     'Every girl stroked a cat.'

Importantly, if the scrambled object in a structure like (7) moves further, to a position above the extracted element, then the resulting configuration is acceptable (8):

(8) *Further scrambling ameliorates LBE from subject*
     <u>Kota</u>$_2$ **každaja**$_1$ včera večerom t_2 [t_1 **devočka**] pogladila t_2.
     cat.ACC every.NOM yesterday evening   girl.NOM stroked

     'Every girl stroked the cat.'

Building from the analysis in Ko (2007, 2011, 2014) for similar effects in Korean and Japanese, Bondarenko and Davis (To appear) argue that the pattern in (7-8) emerges from two concepts—the phasehood of vP, and the ban on phrase-bounded specifier-to-specifier movement. As I discuss in this chapter, these concepts predict the following generalization for Russian, which applies not just to LBE but to sub-extraction from subjects in general:

(9) *Scrambling and sub-extraction from subject*
     Sub-extraction from a Russian subject is banned when other material originating in vP scrambles between the subject and the sub-extracted constituent.

However, we will see that this generalization must be revised, because certain elements are exempt from it. In particular, strings like (7) are permitted for LBE of an adjective or possessor. I will suggest following Bondarenko and Davis (To appear) that this is the result of externally merging these elements in alternative ways which create the appearance of exceptional sub-extraction. We will also see that verb position can improve configurations like (7), which I will suggest following Bondarenko and Davis (To appear) is due to a mechanism of phase extension via head movement (Den Dikken 2007; Gallego 2010; Alexiadou et al. 2014).

In the majority of this chapter, I will make the assumption that Russian LBE simply involves direct extraction of a left branch from the nominal phrase, as has been

---

[3]This topic has also received some preliminary attention in Pesetsky (2006).

argued for LBE in many other languages (Ross 1967; Borsley and Jaworska 1998; Uriagereka 1988; Corver 1990, 1992; Wiland 2010; Stjepanović 2010; Bošković 2005, 2016, a.o.). This analysis is illustrated in (10) below, where XP represents the extracted left branch:

(10)   *Direct extraction analysis of LBE*
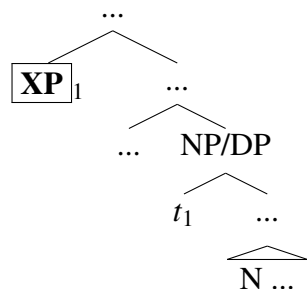
This is not the only analysis of LBE in the literature. I will not consider the alternatives in detail in this chapter, though the appendix briefly discusses them.

1.2   Chapter contents

The concepts needed to analyze the Russian facts previewed above have already been introduced in previous chapters. Therefore in what follows I will proceed directly to a fuller illustration of the facts and their analysis. In section 2, I provide the evidence from extraction in the presence of multiple left branches that the Russian nominal domain is a phase, building on observations from Antonenko and Davis (2020). In section 3, I summarize the findings of Bondarenko and Davis (To appear) about how order preservation in the vP restricts the availability of LBE from subjects. Section 4 concludes, and is followed by an appendix which discusses some remaining facts about extraction from Russian PPs, and its relation to attested analyses of LBE.

2   Linearization within the Russian nominal phrase

In this chapter, we will see that LBE from Russian nominal phrases with multiple left branches is limited in precisely the way we expect if the Russian nominal phrase is a phase, and CL is the correct theory of phases in general. In the previous section, I remained neutral on the precise makeup of the Russian nominal. Several recent works have argued that Slavic languages with no overt articles like Russian are best analyzed as altogether lacking D (Bošković 2005, 2016; Despić 2011, 2013, a.o.). However, there is also a great deal of work arguing that such Slavic languages, and Russian in particular, may indeed have a D projection at least some of the time,

despite never expressing it overtly (Franks 1995; Engelhardt and Trugman 1986; Progovac 1998; Franks and Pereltsvaig 2004; Pereltsvaig 2006, 2007; Bailyn 2012; Pesetsky 2013, a.o.). The Russian patterns explored in this section (and in this chapter generally) do not directly provide evidence for the presence or absence of D in this language. For this reason I will remain neutral on this issue by simply speaking in terms of "nominal phrases". However, in order to be appropriately thorough, this section will provide an analysis that is consistent with the presence of D in Russian, as well as one consistent with its absence.

2.1   The data

Here we will see evidence that there is a tight correlation between the original order of left branch constituents, and the possible orders that LBE can yield in a given context. In the introduction, we saw some preliminary evidence for this fact based on the interaction of demonstrative and adjective LBE. To begin this section, let's consider these elements again, but in greater detail.

As mentioned above, in Russian a demonstrative precedes an adjective occurring in the same nominal phrase, as we see in (11):

(11)   *Demonstrative precedes adjective in Russian*
        Ja videl [ètogo **bol'šogo** kota / * **bol'šogo** ètogo kota].
        I saw   this   big   cat   big   this cat
        'I saw this big cat.'

When these elements co-occur, LBE of the demonstrative alone is possible (12):

(12)   *Demonstrative LBE*
        Ètogo$_1$ ja videl [t_1 **bol'šogo** kota].
        this   I saw   big   cat
        'I saw THIS big cat.'

However, in this context LBE of the adjective alone, which would need to cross the demonstrative, is unacceptable (13):

(13)   *Adjective LBE cannot cross preceding demonstrative*
        * **Bol'šogo**$_2$ ja videl [ètogo t_2 kota].
          big   I saw   this   cat
        'I saw this BIG cat.'

Importantly, when an adjective is the only left branch, it is fully capable of undergoing LBE (14). For this reason, the fact that the presence of the demonstrative removes the possibility of adjective LBE is a puzzle.

(14)   *Adjective LBE possible in absence of demonstrative*
        Bol'šogo$_2$ ja videl [t_2 kota].
        big   I   saw   cat
        'I saw a BIG cat.'

However, as previewed earlier, there is a way for adjective LBE in the presence of a demonstrative to succeed. That is, the co-occurring demonstrative must also move, and land in a position above the extracted adjective, as we see in (15). Example (15a) shows this fact for multiple LBE to adjacent landing sites, while example (15b) shows that the two left branches may be extracted to non-adjacent positions, provided that this constraint on their ordering is obeyed:

(15)   *Adjective LBE permitted by demonstrative LBE to preceding position*

        a.   Ètogo$_1$ **bol'šogo**$_2$ ja videl [t_1 t_2 kota].
              this   big   I saw   cat
              'I saw THIS BIG cat.'

        b.   Ètogo$_1$ ja dumaju **bol'šogo**$_2$ ty   videl [t_1 t_2 kota].
              this   I think big   you saw   cat
              'I think you saw THIS BIG cat.'

Example (16) below shows concretely that movement of the demonstrative to a position below the extracted adjective is not acceptable, regardless of whether the two extracted left branches land in adjacent positions (16a) or not (16b):

(16)   *Adjective LBE not permitted with demonstrative LBE to a lower position*

        a.   * **Bol'šogo**$_2$ ètogo$_1$ ja videl [t_1 t_2 kota].
                big   this   I saw   cat
              'I saw THIS BIG cat.'

        b.   * **Bol'šogo**$_2$ ja dumaju ètogo$_1$ ty   videl [t_1 t_2 kota].
                big   I think this   you saw   cat
              'I think you saw THIS BIG cat.'

In sum, demonstrative and adjective LBE can occur simultaneously, provided that their post-movement order matches the order that these elements had prior to extraction. This is a manifestation of the general requirement that Russian LBE must be order-preserving, which the following facts provide additional support for.

The correspondence between original and post-LBE order also holds, for instance, in contexts with a quantifier and adjective. As we see in (17a-b) below, prior to any LBE, quantifiers precede adjectives in Russian. While LBE of the quantifier by itself is acceptable in this situation (15c), LBE of the adjective is not (17d). However, if both the quantifier and adjective undergo LBE in a way that

preserves their original order, then the result is acceptable (17e), in contrast to a similar configuration where these two elements reverse their order post-LBE (17f):

(17)   *LBE from a nominal phrase with a quantifier and adjective*

    a.   Petja pročital [**každuju** novuju stat'ju].
          Petja read   every   new   article
          'Petja read every new article.'

    b.   * Petja pročital [novuju **každuju** stat'ju].
            Petja read   new   every   article
          'Petja read every new article.'

    c.   **Každuju**$_1$ Petja pročital [t_1 novuju stat'ju].
          every   Petja read   new article
          'Petja read EVERY new article.'

    d.   * Novuju$_2$ Petja pročital [**každuju** t_2 stat'ju].
          new   Petja read   every   article
          'Petja read every NEW article.'

    e.   **Každuju**$_1$ novuju$_2$ Petja pročital [t_1 t_2 stat'ju].
          every   new   Petja read   article
          'Petja read EVERY NEW article.'

    f.   * Novuju$_2$ **každuju**$_1$ Petja pročital [t_1 t_2 stat'ju].
          new   every   Petja read   article
          'Petja read EVERY NEW article.'

Contrasts of the same sort hold for contexts with a co-occurring demonstrative or quantifier and numeral, as shown in (18) below. The demonstrative or quantifier must precede the co-occurring numeral (18a vs. 18b). In this situation, LBE of the demonstrative or quantifier is licit (18c), while extraction of the numeral is not (18d), unless both the demonstrative or quantifier and numeral undergo LBE with an order-preserving result (18e vs. 18f):[4]

---

[4]The numeral *pjat'* ("five") assigns genitive plural morphology to N, though I have not glossed this in (18). It is possible for this numeral to precede a demonstrative if the demonstrative also carries genitive plural morphology, as in (i). Such a sentence communicates that the books are all of the same kind, for instance, all written by Tolstoy. This meaning is absent from typical examples with a demonstrative and numeral.

(i)   Ja pročitala [pjat' ètix   knig].
       I read   five those.GEN.PL books.GEN.PL
       'I read those five books (of that type).'

(18)   *LBE from a nominal phrase with a demonstrative/quantifer and numeral*

      a.   Ja pročitala [**èti/každye** pjat' knig].

            I   read   those/every $\overline{\text{five}}$ books

            'I read those five books / every set of five books.'

      b.   * Ja pročitala [pjat' **èti/každye** knig].

             I   read $\overline{\text{five}}$ those/every books

            'I read those five books / every set of five books.'

      c.   **Èti/každye**$_1$ ja pročitala [t_1 pjat' knig].

            those/every I read $\overline{\text{five}}$ books

            'I read THOSE five books / EVERY set of five books.'

      d.   * Pjat'$_2$ ja pročitala [**èti/každye** t_2 knig].

            $\overline{\text{five}}$ I read   those/every   books

            'I read those FIVE books / every set of FIVE books.'

      e.   ? **Èti/každye**$_1$ pjat'$_2$ ja pročitala [t_1 t_2 knig].

            those/every $\overline{\text{five}}$ I read   books

            'I read THOSE FIVE books / EVERY set of FIVE books.'

      f.   * Pjat'$_2$ **èti/každye**$_1$ ja pročitala [t_1 t_2 knig].

            $\overline{\text{five}}$ those/every I read   books

            'I read THOSE FIVE books / EVERY set of FIVE books.'

So far we have seen contexts where the original order of left branches is fixed, with the word order derivable by LBE being correspondingly limited. As expected, when two left branches can be freely ordered within the nominal phrase, the possibilities for LBE become correspondingly free. For instance, while in an unmarked context numerals precede adjectives (19a), adjectives can also precede numerals provided that the adjective is contrastively focused (19b).[5] When these left branches co-occur, both the numeral and adjective can undergo LBE freely (19c-d). Further, if they both undergo LBE, they can be ordered freely post-movement (19e-f):[6]

I will leave configurations of this sort aside, though the expectation is that any LBE in this context will be order-preserving, as we see throughout this section.

    [5]This is opposed to the above configurations combining, for instance, a demonstrative with an adjective: under no reading is placing an adjective in a position preceding a demonstrative acceptable.

    [6]Given the CL-based explanation I will provide for the facts shown in this section, we expect examples like (19d) where an adjective ultimately precedes a numeral after LBE to have been derived based on an original ordering like that in (19b), which I noted involves focus on the adjective. Consequently, we also expect the adjective to be focused in examples like (19d) and (19f), though the fact that the adjective has undergone LBE in these examples means that it will be focused regardless. Thus the interpretations of (19d) and (19f) do not directly allow us to know that these examples were derived from (19b), though this is what the theory I argue for predicts.

(19)   *LBE from a nominal phrase with an adjective and numeral: Free ordering*

     a.   Ja vstretila [**pjat'** krasivyx   mal'čikov].
          I   met   five   handsome boys

          'I met five handsome boys.'

     b.   Ja vstretila [krasivyx   **pjat'** mal'čikov].
          I   met   handsome five   boys

          'I met five HANDSOME boys.'

     c.   **Pjat'**$_1$ ja vstretila [(t_1) krasivyx   (t_1) mal'čikov].
          five   I   met   handsome   boys

          'I met FIVE handsome boys.'

     d.   Krasivyx$_2$ ja vstretila [(t_2) **pjat'** (t_2) mal'čikov].
          handsome I   met   five   boys

          'I met five HANDSOME boys.'

     e.   **Pjat'**$_1$ krasivyx$_2$ ja vstretila [($t_1 t_2$)/($t_2 t_1$) mal'čikov].
          five   handsome I   met   boys

          'I met FIVE HANDSOME boys.'

     f.   Krasivyx$_2$ **pjat'**$_1$ ja vstretila [($t_1 t_2$)/($t_2 t_1$) mal'čikov].
          handsome five   I   met   boys

          'I met FIVE HANDSOME boys.'

To summarize, we have now seen a variety of facts that can all be described by the following generalization:

(20)   *Russian LBE order preservation generalization*
      LBE from a Russian nominal phrase containing multiple left branches must result in a relative linear ordering that matches the original order that those left branches had prior to any extraction.[7]

Next I will go on to provide an analysis of this generalization.

2.2   Predicting the generalization

Next I will provide an account for the generalization in (20) above. This account will require the hypothesis that the Russian nominal phrase is a phase. As mentioned

---

[7]The evidence for this generalization that I have provided above involves nominal phrases containing two left branches. We expect the same to hold for nominal phrases with three or more left branches as well. So far, I have been unable to identify a Russian nominal phrase with more than two left branches that does not allow some degree of pre-extraction word order variability. For this reason, I cannot test this prediction for more complex nominal phrases here.

in the introduction, there is debate about whether the nominal phrase in article-less Slavic languages like Russian is a bare NP, or a full DP. Since the facts examined here do not directly contribute in one way or another to this debate, I will provide analyses compatible with both the bare NP hypothesis and the DP hypothesis.

2.2.1 Analysis under the bare NP hypothesis

Under the bare NP hypothesis, a CL account of the generalization demonstrated above requires the assumption that NP in Russian is a phase. The concept that NP is a phase does have precedent in previous literature. Bošković (2016), who focuses on LBE in Serbo-Croatian, proposes that an NP that is not sister to D is a phase. Furthermore, several works on morpho-phonology have independently argued that NP is a phase (Marvin 2003; Newell 2008; Embick and Marantz 2008; Embick 2010, a.o.). In specific, such works argue that lexical projections like NP, VP, and AdjP actually consist of category-less lexical roots selected by phasal categorizing heads, respectively n^0, v^0, and a^0. Work in this vein argues that the phasal status of categorizing heads is revealed by the way that category changing morphemes bound the application of various morpho-phonological processes like stress assignment and contextual allomorphy.

If the Russian NP is a phase, then upon its completion it will undergo spell-out, which linearizes its contents. Under CL, any later extraction from this constituent will need to be appropriately order-preserving. The facts examined above show that LBE from the Russian nominal phrase is indeed order-preserving, as expected.[8] CL alone does not predict this generalization, however: independent locality constraints on movement are also vital, as we will see next.

Consider once more LBE from a Russian nominal phrase in which a demonstrative and adjective originate. As we've seen, prior to any movement, the demonstrative precedes the adjective (21):

(21)   *NP containing demonstrative and adjective*
       [$_{NP_{[Phase]}}$ DemP AdjP N]

Notice that the relative order of the demonstrative and adjective is not modified by the demonstrative extracting alone, which is indeed possible in this context (22):

(22)   *LBE of demonstrative alone preserves relative order with adjective*
       DemP   S   V   [$_{NP_{[Phase]}}$ *t*   AdjP   N]

---

[8]In chapter 7, I report new evidence that successive cyclic movement through the NP edge is possible, which is consistent with the proposal that NP is a phase. Since this evidence demonstrates only the possibility of such movement, this is not in itself evidence for the phasehood of NP, however.

If instead the adjective is extracted alone, it must cross over the demonstrative on the way out of this NP. In this case, the original order for the demonstrative and adjective established within NP is reversed by LBE. For this reason CL predicts this derivation to result in unacceptability, as we have seen to be the case in reality (23):

(23)   *Adjective cannot extract across demonstrative*
        * AdjP S V [$_{NP_{[Phase]}}$ DemP t N]

Furthermore, as expected given CL, we have seen that LBE of an adjective in the presence of a demonstrative succeeds if the demonstrative moves also, to a position preceding the adjective (24):

(24)   *Order-restoring demonstrative LBE rescues adjective LBE*
        ✓ DemP AdjP S V [$_{NP_{[Phase]}}$ t t N]

This understanding generalizes to predict all of the constraints on LBE we have seen in this section: any co-occurring left branches, regardless of their category, will need to have the same relative order both before and after any LBE.

However, to make this account complete, another potential derivation must be considered. We indeed expect adjective LBE across a demonstrative in the way shown in (23) to result in an ordering problem if the demonstrative does not also move. This is because in (23) the adjective makes a single step of movement, which simultaneously takes the adjective out of NP while also carrying it across the demonstrative. However, if the adjective were to successive cyclically pass through the NP edge before exiting this constituent, then we would incorrectly expect adjective LBE across an in situ demonstrative to succeed, as in (25). Since the movement step that extracts the adjective from NP does not cross the demonstrative in this situation, movement of the demonstrative as well is not necessary for a coherent linearization in this hypothetical scenario:

(25)   *Hypothetical movement of adjective over demonstrative prior to LBE*
        AdjP S V [$_{NP_{[Phase]}}$ t DemP t N]

Since LBE of an adjective from beneath an un-moved demonstrative is banned in reality, this derivation must be ruled out.

If Russian nominal phrases are bare NPs, the their left branches must all be specifiers or adjuncts to NP. If this is so, then the impossibility of re-arranging the left branches of a given NP prior to LBE arises directly from the ban on phrase-bounded edge-internal movement discussed in the previous chapters (Ko 2007, 2011,

2014). This constraint will accurately prevent movement of an adjective around a co-occurring demonstrative within the same NP, for instance, as we see in (26):

(26)   *No NP-internal left branch re-arranging*

                          *NP
                    ╱   ╲
              AdjP$_1$ ╲
                        DemP   ╲
                               ╱   ╲
                           t_1 N   ...
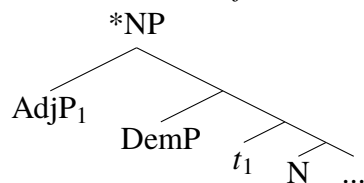
Given this locality constraint, all left branches must remain in their origination positions until the spell-out and linearization of the containing NP, after which CL will require any order derived by LBE to match the order previously established within that NP.[9] [10] These considerations appropriately predict only LBE of the form described by the generalization in (20) above.

2.2.2   Analysis under the DP hypothesis

Next I will provide an analysis of the generalization under consideration that assumes Russian to have full DPs. That DPs are phases is assumed in a great deal of contemporary syntactic work, and explicitly argued for by many (Heck and Zimmermann 2004; Bošković 2005, 2014, 2016; Newell 2008; Syed and Simpson 2017, a.o.).[11] The diagrams in (21-25) above that I have discussed in the context of the NP hypothesis could also be used to illustrate how the relevant generalization emerges under the DP hypothesis, provided that we re-label these diagrams to mark the local phase as being DP rather than NP. However, what changes under the DP hypothesis is that we cannot necessarily invoke the ban on edge-internal movement in order to rule out a successive cyclic LBE derivation like that in (25) above.

   To see this, consider again a context with a co-occurring demonstrative and adjective, assuming that Russian has full DPs. Adjectives are widely argued to be adjuncts of NP, which are interpreted with it via Predicate Modification due to the

---

[9]I have assumed demonstratives and adjectives to be phrases here, given that they appear to be capable of phrasal movement, which differs from the highly local movement typically available to heads (Travis 1984, a.o.). However, it remains possible in principle that these constituents (as well as other left branches) are in fact heads within an extended nominal projection. This analysis would be straightforwardly compatible with the alternative LBE derivations discussed in the appendix, for which LBE does not actually involve extraction of a left branch.

[10]If left branches cannot be re-arranged within the nominal phrase they are externally merged in, we must ask what allows the word order variance for a numeral combined with an adjective shown in (19) above. This fact will be captured if numerals and adjectives can externally merge in either order, as proposed by Bošković (2016) in reference to a similar issue in Serbo-Croatian.

[11]The phasehood of DP will also be important for the account of extraposition in chapter 6.

two both having the semantic type <e,t> (Heim and Kratzer 1998). Additionally, if Russian has DP, we can capture the fact that Russian demonstratives precede adjectives (11) by assuming that demonstratives merge within DP, as in (27):[12]

(27)   *Hypothetical structure for DP with demonstrative and adjective*

```
              DP
           [Phase]
          /      \
      DemP        \
              D      NP
                   /    \
               AdjP   N
                         ...
```
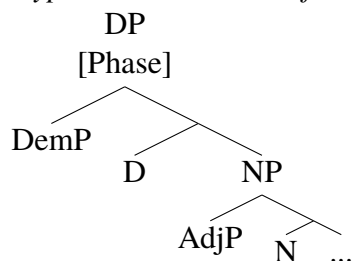
The ban on phrase-bounded edge-internal movement would not prevent the adjective from moving from NP to the edge of DP, as in (28):

(28)   *Hypothetical movement of adjective above demonstrative*

```
              DP
           [Phase]
          /      \
    AdjP_1        \
           DemP    \
               D     NP
                    /   \
                  t_1   N
                          ...
```
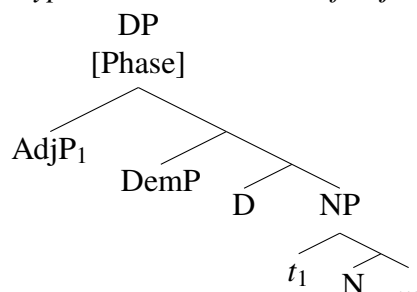
Given that Russian demonstratives must precede co-occurring adjectives, an adjective clearly cannot move to the DP edge and form a terminal landing site there. However, if this position is at least available as an intermediate landing site, then we predict the possibility of successive cyclic adjective LBE via spec-DP. Since such movement will bring the adjective to a position preceding the demonstrative before it exits DP, the possibility of such movement incorrectly predicts the acceptability of adjective LBE from beneath an un-extracted demonstrative.

     While the ban on phrase-bounded edge-internal movement is not applicable here, at least some versions of anti-locality are (Bošković 1997; Ishii 1999; Grohmann 2003; Abels 2003, 2012, a.o.). In particular, Bošković (2005), Brillman and Hirsch (2016), and Erlewine (2016) argue for an anti-locality constraint banning movement from the edge of a given phrase XP to the edge of a constituent YP that immediately dominates XP. This constraint, schematized in (29) below, rules out the hypothetical

---

[12]Alternatively, if adjectives denote functions of type <e,t> and demonstratives are type <<e,t>,e>, then demonstratives will always need to merge above adjectives, as Bošković (2016) correctly notes.

undesirable adjective movement in (28) above.

(29)   *No movement from edge of XP to edge of immediately dominating YP*

$$*\text{YP}$$

(tree diagram: *YP dominating ZP$_1$ and Y, XP; XP dominating t_1 and X, ...)
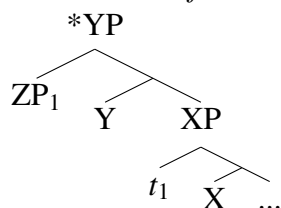
This concept thus accurately predicts that adjective LBE from DP will have to non-successive-cyclically cross a co-occurring demonstrative, forcing additional order-restoring LBE of that demonstrative as well, as in (15) above.[13]

These considerations generalize to prevent movement of any left branch from NP to DP from ever being available as a means of changing the pre-LBE word order of a given set of left branches. Consequently, we predict that any left branches will remain in the order in which they were externally merged until the linearization of the containing DP, after which any LBE from that DP will necessarily be order preserving. This concludes the analysis of the generalization in (20) above.[14]

In the remainder of this section, I will address how the account just provided compares to that in Bošković (2016), who examines similar patterns in Serbo-Croatian, but under very different assumptions.

2.3   Against the phase theory in Bošković (2016)

Bošković (2016) observes patterns of extraction in Serbo-Croatian that are analogous to those shown for Russian above. For instance, in (30) below, we see facts indicating that only the outermost of multiple modifiers in a given nominal phrase is able to undergo LBE in Serbo-Croatian:

---

[13]Additionally, if DP were present, but both the adjective and demonstrative were externally merged in NP, then the same considerations prevent one of those left branches from moving to spec-DP prior to undergoing LBE.

[14]The version of anti-locality just discussed will not be appropriately restrictive if we were to assume an even more articulated structure for the Russian nominal phrase. While the locality principle posited to ensure the rigidity of left branches pre-LBE will necessarily have to change in correspondence with the theory of nominal phrases assumed, this is less important than the simple empirical fact that the pre-LBE order of left branches is indeed constrained: given this fact, CL makes testable predictions about the configurations LBE can form, which I have argued are correct.

(30)   *Adjective order correlates with extractability in Serbo-Croatian*

      a.   neozbiljni mašinski   inžinjer
             not.serious mechanical engineer
             (Bošković 2016 ex. 28a)

      b.   ? Neozbiljnog$_1$ je on otpustio [t_1 mašinskog inžinjera].
              not.serious   he is fired   mechanical engineer
             'He fired a mechanical engineer who was not serious.'
             (Bošković 2016 ex. 27b)

      c.   ?* mašinski   neozbiljni inžinjer
               mechanical not.serious engineer
             (Bošković 2016 ex. 28b)

      d.   * Mašinskog$_2$ je on otpustio [neozbiljnog t_2 inžinjera].
              mechanical he is fired   not.serious   engineer
             'He fired a mechanical engineer who was not serious.'
             (Bošković 2016 ex. 27a)

Bošković also reports patterns in which LBE of the outermost element is necessary to permit LBE of the inner one, as we have seen in multiple scenarios above for Russian. To account for such facts, Bošković assumes that the Serbo-Croatian nominal phrase is a phasal NP, and argues for a derivationally-evaluated version of phase theory that predicts patterns of this variety.

In particular, Bošković proposes that when a phase has multiple edge constituents, only the highest is accessible for extraction. However, he also argues that when the highest edge constituent moves away, the next highest edge constituent then retroactively becomes accessible as well. Thus for Bošković, when a demonstrative and adjective co-occur, for instance, the former is always mobile due to being the initially outermost edge constituent of NP, with the latter derivationally becoming the highest if the former moves away. This proposal captures the basic pattern that when a demonstrative and adjective co-occur, the latter cannot undergo LBE unless the former does also. Bošković's modified phase theory fails to explain several points, however, which I'll argue that a CL approach is better equipped to handle.

An important component of Bošković's proposal is that the Phase Impenetrability Condition (Chomsky 2000, 2001, a.o.) can, in essence, be counter-cyclically re-evaluated. At a conceptual level, it is unclear whether this approach can be usefully related to Chomsky's original hypothesis that Phase Impenetrability and the exceptional nature of phase edges stem from the properties of spell-out. For Chomsky, phase edges are uniquely accessible because spell-out makes only the complement of the phase it applies to inaccessible for later syntactic operations. Uniting this perspective with Bošković's proposal would require stating something

like the following: spell-out applies to all but the highest edge constituent in a phase, but when the highest edge constituent moves, spell-out of the next highest edge constituent is reversed, and that constituent hence becomes accessible. Such a reconciliation certainly appears forced. Alternatively, one might posit that spell-out and Phase Impenetrability effects are unrelated, though in this case we take a step backward in reducing Phase Impenetrability to a stipulation or a description, rather than a generalization with a foundation in independent principles. On the other hand, under the CL approach, we circumvent questions about the nature of Phase Impenetrability by simply abandoning it, and instead capturing the edge-sensitivity effects characteristic of phases in a way that connects to testable predictions about the laws of spell-out and linearization.

The advantage of the CL approach is not only conceptual, but also empirical. Unlike the proposal in Bošković (2016), CL automatically predicts the fact that LBE of multiple elements from the same nominal phrase must preserve the original relative order of those elements in Russian. As repeated in (31) below, in a context with LBE of a demonstrative and an adjective, for instance, adjective LBE is not simply contingent on the demonstrative also moving. Rather, it is vital that the demonstrative's concomitant extraction brings it to a position preceding the adjective, thus maintaining the original order of these elements:

(31)   *Extracted adjective must follow extracted demonstrative in Russian*

   a. $\grave{\text{E}}$togo$_1$ **bol'šogo**$_2$ ja videl [t_1 t_2 kota].
     this   big   I   saw   cat
     'I saw ᴛʜɪꜱ ʙɪɢ cat.'

   b.   * **Bol'šogo**$_2$ ètogo$_1$ ja videl [t_1 t_2 kota].
      big   this   I   saw   cat
     'I saw ᴛʜɪꜱ ʙɪɢ cat.'

Bošković (2016) observes analogous patterns in Serbo-Croatian. For instance, as we see in (32) below, LBE of both a co-occurring demonstrative and adjective requires the former to precede the latter after extraction, just as in Russian:

(32)   *Demonstrative + adjective LBE is order-preserving in Serbo-Croatian*

   a.   Onu$_1$ staru$_2$ prodaje [t_1 t_2 kuću].
     that   old   sells   house
     'He/she is selling that old house'
     (Bošković 2016, ex. 64)

   b.   * Staru$_2$ onu$_1$ prodaje [t_1 t_2 kuću].
      old   that   sells   house
     (Bošković 2016, ex. 65)

Bošković argues that this occurs because a single head attracts both moving elements in these contexts, which thus forces their movements to form crossing paths, following the theory of multiple specifiers in Richards (1997):[15]

(33)   *Claim of Bošković 2016: Multiple LBE is triggered by one head*
       $[_{XP}$ DemP $[_{XP}$ AdjP $[_{X'}$ X S V $[_{NP_{[Phase]}}$ t t N $]]]]$
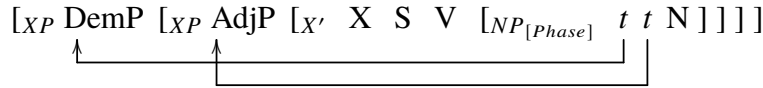
Prima facie, it is not obvious that this syntactic derivation is the only conceivable one for such sentences. Given the fact that word order in Russian (as in many Slavic languages) is extremely free (Bailyn 2012), it would be surprising if there were indeed only one landing site for the movements in a context like this. It is likely that there are more. This concern is especially salient when we consider the contrast repeated in (34) below. Here the two elements undergoing LBE clearly target separate positions, but we see that they nevertheless must move in order-preserving fashion. This fact does not emerge from the system provided by Bošković (2016), but is an automatic prediction of the CL-based account of this chapter.

(34)   *Multiple LBE to non-adjacent positions is also order-preserving*

       a.   Ètogo$_1$ ja dumaju **bol'šogo**$_2$ ty videl $[t_1\ t_2$ kota].
           <u>this</u>   I think   big   you saw   cat
           'I think you saw THIS BIG cat.'

       b.   * **Bol'šogo**$_2$ ja dumaju ètogo$_1$ ty videl $[t_1\ t_2$ kota].
              big   I think   <u>this</u>   you saw   cat
           'I think you saw THIS BIG cat.'

Bošković (2016) notes that the facts he focuses on are evocative of a CL approach, but argues that further facts about binding favor his analysis. In brief, Bošković argues that if binding dependencies are sensitive to phases, his approach accurately predicts that in Serbo-Croatian, an anaphoric possessor cannot be licensed unless it occupies the outermost edge of the containing NP, as shown in (35) below. We see in (35a) that an anaphoric possessor preceded by an adjective within the same nominal phrase is unacceptable, though this configuration is ameliorated by LBE of the adjective as in (35b). For Bošković, such movement of the adjective makes the anaphor count as the outermost edge constituent of the nominal phrase, making it accessible for binding by the subject:

---

[15]We will hear a great deal more about theories of multiple specifier formation in the next chapter.

(35)   *Anaphoric possessor in Serbo-Croatian must be at the edge of NP*

     a.   * **Marija** je prodala [omiljenu **svoju** knjigu].
          Marija is sold   favorite   self's book
          (Bošković (2016) ex. 69)

     b.   Omiljenu$_i$ je **Marija** prodala [t_i **svoju** knjigu].
          favorite   is Marija sold   self's book
          (Bošković (2016) ex. 68)

Bošković cites Zanon (2015) for the proposal that the same account is applicable to Russian. Zanon shows that a variety of elements in fact can precede anaphoric possessors in Russian. However, she argues that in such cases some form of covert movement, either QR or focus movement, extracts the material separating the anaphor from the edge of the nominal phrase. Correspondingly, she argues that for this reason non-quantificational material preceding an anaphoric possessor within the same NP must be focused. I argue that this understanding is incorrect.

     When an adjective and possessor are combined in Russian, the unmarked word order is for the possessor to precede the adjective, even if non-anaphoric (36a). While the adjective can be placed in front of the possessor, this is a marked word order signaling that the adjective is focused, though this focus occurs even with a non-anaphoric possessor (36b). Thus focus in such a configuration is independent, and has nothing to do with anaphor licensing:

(36)   *An adjective preceding a (non-anaphoric) possessor is focused in Russian*

     a.   Ona polgadila [**moju bol'šuju** sobaku].
          she stroked   my   big   dog
          'She stroked my big dog.'

     b.   Ona polgadila [**bol'šuju moju** sobaku], a   ne malen'kuju.
          she stroked   big   my   dog   but not small
          'She stroked my BIG dog, but not my small one.'

     Numerals are also able to precede possessors, as in (37) below. This structure has a partative interpretation, but does not involve focus. Example (37a) shows this for a 1st person possessor, while (37b) shows the same with an anaphoric possessor co-indexed with the matrix subject. There is no difference in interpretation between (37a) and (37b) aside from the difference in the possessor's referent:

(37)   *Numeral preceding (anaphoric) possessor in Russian: No focus*

     a.   Ona polgadila [**pjat' moix** sobak].
          she stroked   five   my   dogs
          'She stroked five of my dogs.'

b.   Ona polgadila [**pjat' svoix** sobak].
     she stroked   five   self's dogs
     'She stroked five of her own dogs.'

Furthermore, Antonenko and Davis (2020) show that it is possible for a Russian anaphoric possessor to be licensed by the subject even when occupying an embedded NP, as in (38). The only plausible way that movement could be involved in the licensing of an anaphor in this context would be if the anaphor itself covertly moves, and thus enters a higher phase in which it is accessible for binding.[16]

(38)   *Anaphor in an embedded nominal phrase*
      Katja pročitala [knigu [**svojego** druga]].
      Katja read   book   self's   friend
      'Katja read her friend's book.'

However, since overt extraction of a possessor originating in this position is unacceptable, as (39) shows, a covert movement analysis of (38) is dubious:

(39)   *No extraction of embedded possessor*
      a.   **Katja** pročitala [knigu [**mojego** druga]].
         Katja read   book   my   friend
         'Katja read my friend's book.'
      b.   * **Mojego**$_1$ **Katja** pročitala [knigu [t_1 druga]].
         my   Katja read   book   friend

Overall, the Russian facts do not verify Bošković's prediction that an anaphor can only be licensed when at the edge of the nominal phrase. Therefore this phenomenon does not provide a basis for favoring his proposals about phases over the CL theory. Since the claims about binding in Bošković (2016) are accompanied by little supporting data, it is possible that they are not ultimately correct for Serbo-Croatian either, though further empirical research is required.

2.4   Interim summary

In this section, I have argued that the following generalization emerges naturally from the CL theory, given the hypothesis that the Russian nominal phrase is a phase:

---

[16]The fact that A′-movement can bring an anaphor into a position in which it is licensed has been observed by at least Barss (1986) and Nissenbaum (2000), for instance:

(i)   a.   *Mary$_1$ told me [$_{CP}$ that we should keep [this picture of herself$_1$]].
      b.   Mary$_1$ told me [$_{CP}$ [which picture of herself$_1$]$_2$ we should keep t_2].

(40)   *Russian LBE order preservation generalization*
         LBE from a Russian nominal phrase containing multiple left branches must
         result in a relative linear ordering that matches the original order that those
         left branches had prior to any extraction.

CL alone is not sufficient here, however: if Russian nominal phrases are NPs, then the
ban on phrase-bounded edge-internal movement is also critical, whereas if Russian
has DPs, the version of anti-locality proposed in Bošković (2005) and Erlewine
(2016) serves to appropriately constrain the derivation. Either way, these Russian
facts, like many others examined in previous chapters, are accurately predicted by
combining CL with independently supported limitations on syntactic movement.
I have also argued here that the CL approach is superior both conceptually and
empirically (at least for Russian) to the proposal advanced by Bošković (2016) in
regard to similar facts from Serbo-Croatian. In the next section, I will set nominal-
internal issues aside in order to focus on order preservation in the Russian vP.

3   Linearization and scrambling in the Russian vP

In this section, I summarize the findings of Bondarenko and Davis (To appear) about
the interaction between scrambling and LBE from subjects in Russian. Following
work by Ko (2007, 2011, 2014) on similar patterns in Korean and Japanese, the
facts we will see in this section reveal evidence for the phasehood of the vP, as well
as the ban on phrase-bounded edge-internal movement.

3.1   The core data

Russian typically permits LBE from subjects, as exemplified in (41):

(41)   *LBE from subject*
         Èta/každaja$_1$ včera   [t_1 **devočka**] pogladila kota.
         this/every   yesterday **girl**   stroked   cat
         'This/Every girl stroked the cat yesterday.'

Another productive movement process in Russian is scrambling. Most relevant to
this section is object scrambling, exemplified in (42):

(42)   *Object scrambling*
         Kota$_1$ èta devočka pogladila t_1.
         cat   this girl   stroked
         'This girl stroked the cat.'

However, it is not possible to both extract from a subject, and scramble the object to a position between the subject and extracted constituent, as in (43):

(43)   *No object scrambling + LBE from subject*
         * **Èta**$_2$ kota$_1$ [t_2 **devočka**] pogladila t_1.
         this.NOM cat.ACC   **girl.NOM** stroked
         'This girl stroked the cat.'

This restriction is surprising, since both of the movement processes needed to create this structure are independently available. Example (43) above shows this fact for LBE of a demonstrative, but the same limitation holds for LBE of a quantifier (44a) or numeral (44b), for instance:

(44)   a.   *No object scrambling + quantifier LBE from subject*
            * **Každaja**$_2$ kota$_1$ [t_2 **devočka**] pogladila t_1.
            every.NOM cat.ACC   **girl.NOM** stroked
            'Every girl stroked a cat.'

       b.   *No object scrambling + numeral LBE from subject*
            * **Tri**$_2$ košku$_1$ [t_2 **mal'čika**] uvideli t_1.
            three.NOM cat.ACC   **boy.NOM** saw
            'Three boys saw a cat.'

As addressed later on, examples like (43-44) are acceptable for many speakers when they involve LBE of an adjective or possessor. Consequently, the facts discussed in this section can be described by the following empirical generalization:

(45)   *Scrambling and sub-extraction from subject*
         Sub-extraction from a Russian subject (excluding some adjectives and possessors) is banned when other material originating in vP scrambles to a position between the subject and the sub-extracted constituent.

Here I will first provide a means of understanding the unacceptable configurations just shown, before discussing later on why adjectives and possessors are exempt.[17]

---

[17]A reviewer of an earlier version of this material points out another exception to the basic pattern analyzed in this section: when the scrambled object is pronominal, it does not block LBE from the subject. This fact cannot be attributed to pronouns being phonologically "light", since scrambling of an equally light proper name does block such LBE.

(i)   Každaja$_1$/Èta$_1$ **ego**$_2$/***L'va**$_2$ [t_1 devočka] pocelovala t_2.
         every/this   him/Lev   girl   kissed.
         'Every girl kissed him/Lev'.

A similar constraint is known to hold for numeral quantifier stranding in Korean and Japanese (Saito 1985; Miyagawa 1989; Ko 2007, 2014). As (46) below shows with Korean data, while scrambling of an object that strands a numeral quantifier below the subject is possible (46a), similar scrambling from subject position (which Ko shows is typically permitted) combined with scrambling of an object to an intervening position, is unacceptable (46b). Since Korean and Japanese are unrelated to Russian, this limitation on sub-extraction must not be a language-particular oddity, but rather a fact that emerges from more general properties of natural language.

(46)   a.   *Numeral quantifier stranding by object scrambling*

Maykcwu-lul$_1$ John-i   [t_1 **sey-pyeng**] masiessta.
beer-ACC   John-NOM [3-bottle   ] drank

'John drank three bottles of beer.'
(Ko 2014, pg. 31, ex. 1b)

b.   *No stranding by subject scrambling combined with object scrambling*

* Haksayng-tul-i$_1$ **maykcwu-lul**$_2$ [t_1 sey-myeng] t_2 masiessta.
 students-PL-NOM beer-ACC   3-people   drank

'Three students drank beer.'
(Ko 2014, pg. 32, ex. 7)

Ko (2007, 2014) argues that in Korean and Japanese this restriction arises from two concepts that feature throughout this dissertation—CL and the ban on phrase-bounded edge-internal movement. This section will show how such an account extends to Russian, though a few additional Russian-specific puzzles will also arise.

3.2   Analysis

Assume that vP, in whose edge external arguments originate, is a phase in Russian. If vP is a phase, then given CL, the final relative ordering of any material originating in vP must have been established by the time the vP was linearized. Otherwise, the derivation will ultimately suffer from a linearization contradiction. Given this expectation, deriving the unacceptable sentences in (43-44) would require scrambling the object over the subject within vP, and then sub-extracting a constituent from the subject and placing it in a vP-internal position preceding the scrambled object. As shown in the schema in (47) below, these operations derive the correct word order:

Bondarenko and Davis (To appear) suggest that this is related to the fact that (unfocused) pronominal objects in Russian prefer to shift to a pre-verbal position, which is perhaps analogous to facts about pronouns in English particle verb constructions, e.g. *I cleaned it up / *I cleaned up it*. If pronoun displacement in these contexts is the result of a post-syntactic PF operation rather than actual syntactic movement, it is conceivable that the concepts discussed in this section will not be applicable to it.

(47)   *Object scrambling over subject followed by LBE from subject in vP*
        $[_{vP}$ XP O $[_{Subj}$ t NP] v V t_{o}]
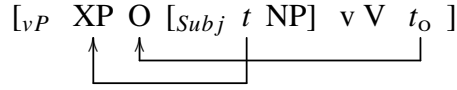
If the Russian subject A-moves to spec-TP as Bailyn (2012) argues, after the forma-
tion of (47), the (remnant) subject will move from vP, thus forcing order-preserving
additional movement of the extracted XP and the scrambled object to vP-external
positions. If the vP structure in (47) is available, then nothing rules out such a deriva-
tion, which would yield strings like (43-44). However, if we can identify a problem
with either of the vP-internal movement steps that are necessary to form (47), then
we will have a reason why sentences with this word order are unacceptable.

    The first movement step, where the object scrambles over the subject to the edge
of vP, does not violate any attested locality condition:

(48)   *No problem for object scrambling over subject*
        ✓ $[_{vP}$ O $[_{Subj}$ XP NP] v V t_{o}]
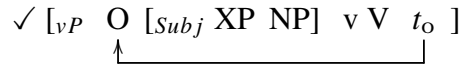
However, there is a problem with the second movement step, in which a constituent
undergoes LBE from the subject to a position above the scrambled object within vP.
Such LBE violates the ban on movement from (within) one specifier to another of
the same phrase, given that the in situ external argument originates in spec-vP:

(49)   *Illegal movement from subject within spec-vP*
        * $[_{vP}$ XP O $[_{Subj}$ t NP] v V t_{o}]
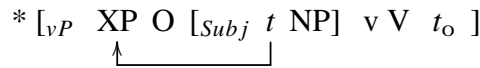
Given this restriction, we have an explanation for the unacceptability of scrambling
an object over the subject, and then sub-extracting from the subject. Since the second
of these movements cannot occur within the vP, the word order characteristic of these
sentences cannot be derived in vP. Thus if the word order in question were derived
later on, by non-successive-cyclically extracting an element from the subject across
the scrambled object as in (50) below, there will be a linearization contradiction:
this movement reverses the relative order of the extracted element and of the object
that was derived in vP, and is thus illicit.

(50)   *LBE from subject directly across scrambled object*
        $[_{CP}$ XP ... $[_{vP}$ O $[_{Subj}$ t NP] v V t_{o}]]
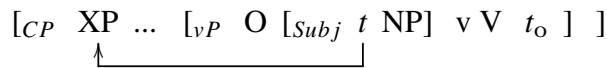
    While a derivation like (50) would be illicit at PF if no other movement occurs,
recall that CL makes a prediction about how to rescue such configurations. That
is, non-successive-cyclic movement from a phase is predicted to be licit provided

that it is accompanied by additional order-restoring movement of any crossed-over material. This prediction leads us to expect that the unacceptable Russian sentences we are primarily concerned with should be repaired, if the scrambled object moves far enough to end up to the left of the constituent extracted from the subject. This is predicted because such movement restores the order of the extracted material and the object that was established in vP, and so should avoid a linearization contradiction. This prediction is correct, as (51-53) below show. The (a) versions of these examples show familiar unacceptable configurations, which are repaired in the (b) sentences by further scrambling of the object. The (b) sentences also include an adverb to ensure that the intended LBE from subject is evident. The (c) variants of these examples serve to show that such configurations with additional scrambling (as signaled by the presence of the adverb) are indeed unacceptable if the scrambled object ultimately follows the sub-extracted element, instead of preceding it:

(51)   a.   *No scrambling to position between subject and extracted element*

  * **Každaja**$_2$ <u>kota</u>$_1$ [t_2 **devočka**] pogladila t_1.
  every.NOM cat.ACC   girl.NOM stroked

  'Every girl stroked a cat.'

  b.   *Amelioration by further scrambling above extracted element*

  <u>Kota</u>$_1$ **každaja**$_2$ včera   večerom [t_2 **devočka**] pogladila t_1.
  cat.ACC every.NOM yesterday evening   girl.NOM stroked

  'Every girl stroked the cat yesterday.'

  c.   *No amelioration for further scrambling below extracted element*

  * **Každaja**$_2$ <u>kota</u>$_1$ včera   večerom [t_2 **devočka**] pogladila t_1.
  every   cat   yesterday evening   girl   stroked

  'Every girl stroked the cat yesterday.'

(52)   a.   *No scrambling to position between subject and extracted element*

  * **Pjat'**$_2$ <u>kota</u>$_1$ [t_2 **mal'čikov**] pogladili t_1.
  five   cat.ACC   boys.GEN.PL stroked

  'Five boys stroked a cat.'

  b.   *Amelioration by further scrambling above extracted element*

  <u>Kota</u>$_1$ **pjat'**$_2$ včera   [t_2 **mal'čikov**] pogladili t_1.
  cat.ACC five   yesterday   boys.GEN.PL stroked

  'Five boys stroked a cat yesterday.'

177

    c.   *No amelioration for further scrambling below extracted element*

        * **Pjat'**$_2$ <u>kota</u>$_1$ včera   [t_2 **mal'čikov**] pogladili t_1.
        five   cat.ACC yesterday   boys.GEN.PL stroked

        'Five boys stroked a cat yesterday.'

(53)   a.   *No scrambling to position between subject and extracted element*

        * **Ètot**$_2$ <u>mašinu</u>$_1$ [t_2 **rebënok**] slomal t_1.
        this.NOM car.ACC   child.NOM broke

        'This child broke a car.'

    b.   *Amelioration by further scrambling above extracted element*

        <u>Mašinu</u>$_1$ **ètot**$_2$ včera   [t_2 **rebënok**] slomal t_1.
        car.ACC this.NOM yesterday   child.NOM broke

        'This child broke a car yesterday.'

    c.   *No amelioration for further scrambling below extracted element*

        * **Ètot**$_2$ <u>mašinu</u>$_1$ včera   [t_2 **rebënok**] slomal t_1.
        this.NOM car.ACC yesterday   child   broke

        'This child broke a car yesterday.'

These facts make clear that the requirements of linearization are the only reason for a conflict between extraction from the subject and scrambling of an object. Both of these processes are independently available in Russian syntax, and can in principle freely combine, provided that the result satisfies order preservation. Also important is that, as (51-53) show, the order that must be preserved is the one that was established upon the completion of the vP (and not for instance the original order of the relevant elements). That is, in these examples the scrambled object must ultimately precede the element extracted from the subject, since this reflects the order established in the vP after object scrambling within that phase occurred.

3.3   Predictions for non-LBE extraction

Russian allows a nominal phrase to be inverted and split in such a way that its specifiers/adjuncts are stranded in a lower position in the sentence (Lyutikova 2012). As expected, object scrambling prevents the formation of a split inversion of a subject, unless the scrambled object moves high enough to precede the extracted segment of the subject. Two such patterns are shown in (54) and (55) below.[18]

---

[18]By hypothesis such inverting extractions must involve successive cyclic movement through the edge of the nominal phrase. I will refrain from making a precise proposal about what segment of

(54)   a.   *Inverted split with demonstrative stranding*

       Mal'čik$_1$ (včera)   [**ètot** t_1] pogladil <u>sobaku</u>.
       boy.NOM (yesterday) **this.NOM**   stroked   dog

       'This boy stroked the dog (yesterday).'

   b.   *Interference of object scrambling*

       * **Mal'čik**$_1$ <u>sobaku</u>$_2$ [**ètot** t_1] pogladil t_2.
       boy.NOM dog.ACC **this.NOM**   stroked

       'This boy stroked the dog.'

   c.   *Amelioration by object scrambling above extracted element*

       <u>Sobaku</u>$_2$ **mal'čik**$_1$ včera   [**ètot** t_1] pogladil t_2.
       dog.ACC **boy.NOM** yesterday **this.NOM**   stroked

       'This boy stroked the dog yesterday.'

   d.   *No amelioration for lower scrambling*

       * **Mal'čik**$_1$ <u>sobaku</u>$_2$ včera   [**ètot** t_1] pogladil t_2.
       boy.NOM dog.ACC yesterday **this.NOM**   stroked

       'This boy stroked the dog yesterday.'

(55)   a.   *Inverted split with quantifier stranding*

       Mal'čik$_1$ (včera)   [**každyj** t_1] <u>košku</u> uvidel.
       boy.NOM (yesterday) **every.NOM**   cat   saw

       'Every boy saw the cat (yesterday).'

   b.   *Interference of object scrambling*

       * **Mal'čik**$_1$ <u>košku</u>$_2$ [**každyj** t_1] uvidel t_2.
       boy.NOM cat.ACC **every.NOM**   saw

       'Every boy saw the cat.'

   c.   *Amelioration by object scrambling above extracted element*

       <u>Košku</u>$_2$ **mal'čik**$_1$ včera   [**každyj** t_1] uvidel t_2.
       cat.ACC **boy.NOM** yesterday **every.NOM**   saw

       'Every boy saw the cat yesterday.'

---

the nominal phrase is moved in the process of forming such inversions, since this would not add anything substantial to the analysis at hand.

d. *No amelioration for lower scrambling*

* **Mal'čik**$_1$ košku$_2$ včera [**každyj** t_1] uvidel t_2.
boy.NOM cat.ACC yesterday **every.NOM** saw

'Every boy saw the cat yesterday.'

The explanation for examples (51-53) above, involving LBE from the subject, applies directly to the examples with split inversion in (54-55) as well. Nothing substantial differentiates these configurations, aside from the choice of element that is sub-extracted from the subject. This explanation also extends to the pattern of judgments shown in (56) for extraction of a post-nominal genitive complement from a subject:

(56) a. *Extraction of complement from subject*

Xudožnika$_2$ (včera) [**dočka** t_2] pogladila kota.
painter.GEN (yesterday) **daughter.NOM** stroked cat

'Daughter of a painter stroked the cat.'

b. *Interference of object scrambling*

* **Xudožnika**$_2$ kota$_1$ [**dočka** t_2] pogladila t_1.
painter.GEN cat.ACC **daughter.NOM** stroked

'Daughter of a painter stroked the cat.'

c. *Amelioration by object scrambling above extracted element*

Kota$_1$ **xudožnika**$_2$ včera [**dočka** t_2] pogladila t_1.
cat.ACC **painter.GEN** yesterday **daughter.NOM** stroked

'Daughter of a painter stroked the cat yesterday.'

d. *No amelioration for lower scrambling*

* **Xudožnika**$_2$ kota$_1$ včera [**dočka** t_2] pogladila t_1.
painter.GEN cat.ACC yesterday **daughter.NOM** stroked

'Daughter of a painter stroked the cat yesterday.'

3.4 Predictions for other material originating in vP

So far we have seen configurations where object scrambling interferes with extraction from the subject. The relevant restriction and its amelioration should hold not only for scrambling of objects, but for scrambling of any phrase whatsoever that originates in vP. This is because any material that is introduced in and thus linearized in vP will be linearized with respect to the subject, and thus can become a source of linearization contradictions with respect to material extracted from the subject.

As expected, scrambling of an oblique argument, for instance, also interacts with extraction from a subject in precisely the same way that object scrambling does (57):

(57) a. *Initial sentence*
Ètot **student** predstavil <u>učitel'nice</u> Mašu.
this.NOM student.NOM introduced teacher.DAT Masha

'This student introduced Masha to the teacher.'

  b. *Dative scrambling*

<u>Učitel'nice</u>$_1$ **ètot** **student** predstavil t_1 Mašu.
teacher.DAT **this.NOM student.NOM** introduced Masha

'This student introduced Masha to the teacher.'

  c. *No dative scrambling + LBE from subject*
* **Ètot**$_2$ <u>učitel'nice</u>$_1$ [t_2 **student**] predstavil t_1 Mašu.
this.NOM teacher.DAT **student.NOM** introduced Masha

'This student introduced Masha to the teacher.'

  d. *Amelioration by further scrambling above extracted element*
<u>Učitel'nice</u>$_1$ **ètot**$_2$ včera [t_2 **student**] predstavil t_1
teacher.DAT **this.NOM** yesterday **student.NOM** introduced
Mašu.
Masha

'This student introduced Masha to the teacher yesterday.'

  e. *No amelioration by further scrambling below extracted element*
* **Ètot**$_2$ <u>učitel'nice</u>$_1$ včera [t_2 **student**] predstavil t_1 Mašu.
this teacher.DAT yesterday **student** introduced Masha.ACC

'This student introduced Masha to the teacher yesterday.'

The same applies for scrambling of a VP-level PP (58):

(58) a. *Initial sentence*
Pjat' **mal'čikov** prinesli [<u>v</u> <u>klass</u>] pivo.
five.NOM boys brought in classroom beer

'Five boys brought beer into the classroom.'

  b. *PP scrambling*
[<u>V</u> <u>klass</u>]$_1$ **pjat'** **mal'čikov** prinesli t_1 pivo.
in classroom **five.NOM boys** brought beer

'Five boys brought beer into the classroom.'

  c.   *No PP scrambling + LBE from subject*
     * **Pjat'**$_2$ [v klass]$_1$ [t_2 **mal'čikov**] prinesli t_1 pivo.
     five.NOM in classroom   **boys**   brought   beer

     'Five boys brought beer into the classroom.'

  d.   *Amelioration by further scrambling above extracted element*
     [V klass]$_1$ **pjat'**$_2$ včera   [t_2 **mal'čikov**] prinesli t_1 pivo.
     in classroom **five.NOM** yesterday   **boys**   brought   beer

     'Five boys brought beer into the classroom yesterday.'

  e.   *No amelioration by further scrambling below extracted element*
     * **Pjat'**$_2$ [v klass]$_1$ včera   [t_2 **mal'čikov**] prinesli t_1 pivo.
     five.NOM in classroom yesterday   **boys**   brought   beer

     'Five boys brought beer into the classroom yesterday.'

Scrambling of a low adverb behaves the same (59):

(59)   a.   *Initial sentence*
        Každaja   **devočka** vyčistila jaščik polnostju.
        every.NOM girl.NOM cleaned drawer completely

        'Every girl cleaned a drawer completely.'

     b.   *Low adverb scrambling*
        Polnostju$_1$ **každaja**   **devočka** vyčistila jaščik t_1.
        completely **every.NOM girl.NOM** cleaned drawer

        'Every girl cleaned a drawer completely.'

     c.   *No low adverb scrambling + LBE from subject*
        * **Každaja**$_2$ polnostju$_1$ [t_2 **devočka**] vyčistila jaščik t_1.
        every.NOM completely   **girl.NOM** cleaned drawer

        'Every girl cleaned a drawer completely.'

     d.   *Amelioration by further scrambling above extracted element*
        Polnostju$_1$ **každaja**$_2$ včera   [t_2 **devočka**] vyčistila jaščik t_1.
        completely **every.NOM** yesterday   **girl.NOM** cleaned drawer

        'Every girl cleaned a drawer completely yesterday.'

     e.   *No amelioration by further scrambling below extracted element*
        * **Každaja**$_2$ polnostju$_1$ včera   [t_2 **devočka**] vyčistila jaščik t_1.
        every.NOM completely yesterday   **girl.NOM** cleaned drawer

        'Every girl cleaned a drawer completely yesterday.'

Finally, we see in (60) that adverbs which plausibly originate outside of the vP do not interfere with LBE from the subject. This is expected, since an adverb that originates outside of vP cannot have an effect on the linearization of the vP, in which the relevant interactions that yield the patterns shown above occur. Thus such vP-external material cannot be the basis for a linearization contradiction with respect to the contents of the subject.

(60)   *High adverbs do not interrupt extraction from subject*

      a.   **Každaja**$_1$ [včera   večerom] [t_1 **devočka**] vyčistila jaščik.
            every   yesterday evening   **girl**   cleaned drawer
            'Every girl cleaned a drawer yesterday evening.'

      b.   **Ètot**$_1$ [po vsej vidimosti] [t_1 **student**] predstavil učitel'nice Mašu.
            this at all sight   **student** introduced teacher   Masha
            'Apparently, this student introduced Masha to the teacher.'

      c.   **Pjat'**$_1$ verojatno [t_1 **mal'čikov**] prinesli v klass   pivo.
            five   probably   **boys**   brought in classroom beer
            'Probably, five boys brought beer into the classroom.'

This concludes the core analysis of this section. We have seen here that the hypothesis that the Russian vP is a phase, combined with the ban on edge-internal movement, accurately predicts certain interactions between extraction from subjects, and scrambling of other material originating in vP. In what follows, I will address a few additional puzzles about the restriction examined here.[19]

---

[19]If the syntax of unergative predicates is like that of transitive ones, aside from the absence of an object, then we expect the restriction examined here to emerge in unergative contexts, though presumably not in unaccusative ones. In Bondarenko and Davis (To appear), we note that this prediction is generally not verified. Both unergative and unaccusative predicates generally allow material originating in vP to appear in between the subject and an element extracted from it:

(i)   a.   **Pjat'**$_{j1}$ [s   dereva] [t_1 **jablok**] upalo.
            five   from tree   apples fell
            'Five apples fell from the tree.'

     b.   **Pjat'**$_j$ [v klasse] [t_j **devoček**] tancevali.
            five in class   girls   danced
            'Five girls danced in the class.'

Sometimes there is indeed a contrast, as we see in (ii), where quantifier LBE in (iia) from an unergative subject is degraded relative to quantifier LBE from an unaccusative subject in (iib):

3.5 Further complications

3.5.1 The behavior of adjectives and possessors

As mentioned earlier, the constraint that this section has focused on does not hold consistently for LBE of certain elements—adjectives and possessors. As (61) below demonstrates, LBE of these elements from the subject is not interrupted by scrambling of the object to an intervening position:

(61)   *Exceptions to the constraint: Adjectives and possessors*

    a.   *Object scrambling doesn't block LBE of adjective from subject*

       Vesëlaja$_2$ tort$_1$ [t_2 **devočka**] ela t_1.
       happy.NOM cake.ACC   **girl.NOM** ate

       'The happy girl ate cake.'

    b.   *Object scrambling doesn't block LBE of possessor from subject*

       $^{\%}$ **Naš/Vasin**$_2$ knigu$_1$ [t_2 **syn**]   pročital t_1.
       our.NOM/Vasja's.NOM book.ACC   **son.NOM** read

       'Our/Vasja's son read the book.'

Bondarenko and Davis (To appear) observe that configurations like this are quite productive for adjectives, though somewhat less so for possessors. Initially, we might wonder whether these configurations are necessarily derived by movement. Displacement of these elements does appear to be island-bounded, as (62) shows using a relative clause island, suggesting that movement is indeed involved:

(62)   *Adjective / possessor LBE is island-bounded*

    a.   * **Čërnuju**$_2$ devočka, [kotoraja uvidela [t_2 **košku**]], ela tort.
       black.ACC girl   who   saw   cat.ACC ate cake
       'The girl who saw a black cat ate cake.'

(ii)   a.   * **Každaja**$_2$ [so   skakalkoj]$_1$ [t_2 **devočka**] prygala t_1.
         every   with jumping.rope   **girl**   jumped
         'Every girl jumped with a jumping rope.'
    b.   **Každoe**$_2$ [na zemlju]$_1$ [t_2 **jabloko**] upalo t_1.
         every   on ground   **apple**   fell
         'Every apple fell on the ground.'

One possibility is that the argument structure of the predicates we have tested is flexible enough to frequently circumvent the restriction analyzed here. Another is that the elements we have attempted to use as interveners are not all strictly generated vP-internally. Yet another possibility is that the syntax of unergative predicates in Russian is generally closer to that of unaccusatives than to that of transitives. This issue must be left aside for future work.

b. * **Našu**$_2$ / **Vasinu**$_2$ devočka, [kotoraja uvidela **[t_2 košku]**], ela
   our.ACC / **Vasja's.ACC** girl who saw cat.ACC ate
   tort.
   cake
   'The girl who saw a our /Vasja's cat ate cake.'

While the displaced adjective / possessor in strings like (61) thus presumably reaches its left peripheral position via movement, the possibility of these strings will be predicted if these elements can participate in a derivation in which they are externally merged outside of the subject that they are apparently construed with. In such a derivation, the adjective or possessor would not in fact sub-extract, and thus the constraints investigated above would not be applicable.

As Bondarenko and Davis (To appear) discuss, there is tentative evidence for such an analysis. Adjectives, at least, are capable of forming predications like that in (63), where the right dislocated adjective *vesëlaja* ("happy") can be translated as approximating "while being happy", rather than plain adjectival modification. This might be analyzed as some variety of secondary predication, but most importantly, here we see an adjectival structure which is plausibly base generated external to the subject, presumably modifying the predicate instead:

(63) *Subject external adjective*

Maša otpravilas' domoj, **vesëlaja** (**i vsem dovol'naja**).
Masha went home happy and all satisfied

'Masha went home, (while being) happy (and satisfied with everything).'

If a sentence like (63) indeed involves the adjective being externally merged external to the subject, then by forming such a configuration and then fronting the adjective, a sentence like (61a) would be derived.[20] There is another construction that is also

---

[20]The reading for the adjective in a structure like (63) is presumably most sensible with stage-level rather than individual-level adjectives. Consider, for example, the following contrast:

(i) a. While being in a good mood, John took a walk in the park.
    b. * While being blue-eyed, John took a walk in the park.

At least some Russian speakers judge the displaced adjective in sentences like (61a) to be interpreted in such a manner, and furthermore, find that stage-level adjectives are most acceptable in this context. The examples in (ii-iii) below illustrate this contrast:

(ii) *Stage-level adjective displacement + object scrambling*

a. **Vesëlaja**$_2$ tort$_1$ [t_2 **devočka**] ela t_1.
   happy.NOM cake.ACC girl.NOM ate
   'A happy girl ate cake.'

185

relevant to this hypothetical analysis. Both adjectives and possessors can be what Graschchenkov (2016, 2018) analyzes as "shifted attributes", illustrated in (64):

(64) *Shifted attributes*

Maša   včera   razbila vazu, **doroguščuju, maminu / moju**.
Masha yesterday broke   vase **very.expensive mom's / my**

'Mary broke a vase yesterday, a very expensive one, mom's / mine.'

Graschchenkov (2016) argues that such shifted attributes originate in a predicational phrase within the clause, and not in a nominal phrase. Fronting of the "shifted" constituent in such a structure to clause initial position would also successfully derive sentences of the form in (61) above.

---

b. **Novyj**$_2$ Petju$_1$ [t_2 **direktor**] uvolil t_1.
new.NOM Petja.ACC   director.NOM fired
'A new director fired Petja.'

(iii) *Individual-level adjective displacement + object scrambling*

a. * **Goluboglazyj**$_2$ stixotvorenije$_1$ [t_2 **mal'čik**] pročital t_1.
blue-eyed.NOM poem.ACC   boy.NOM read
'A blue-eyed boy read a poem.'

b. * **Vysokaja**$_2$ zabor$_1$ [t_2 **devočka**] perelezla t_1.
tall.NOM   fence.ACC   girl.NOM climbed.over
'A tall girl climbed over a fence.'

This contrast is expected if these sentences do not involve sub-extraction (unlike what the traces here suggest), but rather, fronting of an adjective introduced in a predication like that in (63) above. The hypothesis that it is possible to generate an adjective external to a nominal phrase also accurately predicts the possibility of strings resembling violations of the generalization illustrated in section 2 of this chapter. We see two examples of this sort in (iv), which show fronted adjectives modifying subjects containing demonstratives:

(iv) a. **Vesëlaja**, tort   **èta**   devočka ela.
happy.NOM cake.ACC this.NOM girl.NOM ate
'Being happy, this girl ate cake.'

b. **Vysokaja**, zabor **èta**   devočka perelezla.
tall.NOM   fence.ACC this.NOM girl.NOM climbed.over
'Being tall, this girl climbed over a fence.'

While extracting an adjective across a demonstrative merged in the same nominal phrase should be unacceptable, we expect no violation here if the relevant adjectives were in fact introduced external to the nominal phrase. Importantly, the adjectives in these examples are most naturally paraphrased with an interpretation like that of the adjective in (63), rather than straightforward adjectival modification.

Unfortunately, the judgments reported in this footnote are not stable across speakers, though the analysis just described appears promising for at least some. Clearly additional research is needed.

Another hypothesis about the derivation of the problematic configuration exemplified in (61) becomes available if it is conceivable that adjectives and pre-nominal possessors in Russian are both adjuncts. Considering adjectives to be adjuncts is uncontroversial. If Russian pre-nominal possessors are also adjuncts (Lyutikova 2014), and adjuncts are capable of externally merging late to a phrase that has already been constructed and moved (Lebeaux 1988, a.o.), then we can analyze the sentences in (61) as involving object scrambling combined with late merge of an adjective or possessor to a position created by covert movement of the subject over the object, as in the simplified derivation diagrammed in (65):

(65)   a.   *Object scrambling + covert movement of subject*
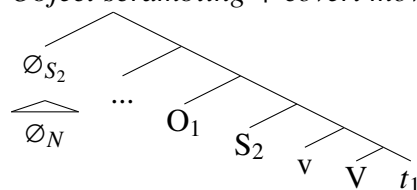
        b.   *Late merge of adjunct to covertly moved subject*
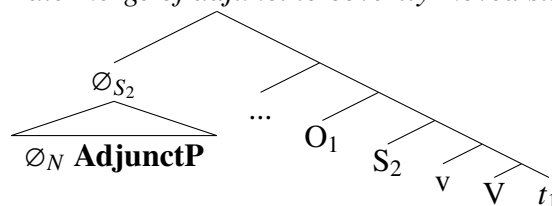
The derivation in (65) is similar to what Fox and Nissenbaum (1999) propose for extraposition, though with the seemingly displaced constituent appearing leftward rather than rightward. Chapter 6 discusses this analysis of extraposition in detail.[21]

We have now seen two potential methods of deriving the exceptional sentences in (61). Both of these analyses are variants of a proposal that adjectives and possessors can be externally merged in a distinct way that allows them to become seemingly sub-extracted, despite not actually sub-extracting from a nominal phrase. Neither

---

[21]Russian pre-nominal possessors capable of LBE for the most part bear agreement morphology (received under concord with N) just as adjectives and most other left branches do. However, third person pronominal possessors in Russian do not have such agreement morphology, and Bondarenko and Davis (To appear) show that scrambling of vP-internal material consistently blocks LBE of these possessors from the subject. We see this in (i) with scrambling of an object:

(i)   [**Ego/eë/ix**]$_1$ včera   večerom (*<u>goršok$_2$</u>) [t_1 **koška**]   razbila (<u>goršok$_2$</u>)
        his/her/their yesterday evening pot.ACC   **cat.NOM** broke pot.ACC
        'Yesterday evening his / her / their cat broke the pot.'

This may suggest that these possessors are syntactically distinct from others, though it is difficult to formulate a precise hypothesis about what the relevant distinction is based on this fact alone.

of these proposals is fully decisive, however, since the empirical landscape here is quite complex. I will leave deeper investigation of this topic to other work.

3.5.2 The ameliorating effect of VS order

All the sentences we have seen in this chapter use SV word order. Russian also permits VS orders, however, which for some speakers interestingly differ by rescuing the unacceptable configuration analyzed earlier. In particular, as (66) below shows, for at least some speakers object scrambling and extraction from the subject can co-occur when the verb precedes the subject. While we have seen that the restriction in focus here can be repaired by additional scrambling to a position above the extracted element, in (66), this does not occur. Instead, use of a different verb position alone improves the configuration:

(66) *VS order permits object scrambling + extraction from subject*

a. *OVS order*

% **Každaja/Èta$_2$** kota$_1$ pogladila [t_2 **devočka**] t_1.
    every.NOM/this.NOM cat.ACC stroked **girl.NOM**

'Every/this girl stroked the cat.'

b. *VOS order*

% **Každaja/Èta$_2$** pogladila kota$_1$ [t_2 **devočka**] t_1.
    every.NOM/this.NOM stroked cat.ACC **girl.NOM**

'Every/this girl stroked the cat.'

Several works have argued that if a phase head moves, it extends phasehood up to the head moved to (Den Dikken 2007; Gallego 2010; Alexiadou et al. 2014). Following such works, Bondarenko and Davis (To appear) hypothesize that in sentences like (66), V moves to a head above vP, carrying v into a higher head, which thereby inherits its phasehood, as in (67).

(67) *Head movement enlarges the phase (here to a hypothetical XP)*
[$\boxed{XP[Phase]}$ X-v-V [$_{vP}$ S v-V [$_{VP}$ V ...]]]

The constraint on LBE from subjects analyzed in this chapter is only predicted to hold when vP is a phase, due to the ban on phrase-bounded specifier to specifier movement. But if phase extension makes the relevant phase a constituent larger than vP as in (67), then this ban will cease to be applicable. Thus in a structure like (67) it should be possible to scramble an object to spec-XP or spec-vP, and then extract

an element from the subject and into spec-XP—the edge of the local phase. Such movements successfully derive the word orders of (66a) and (66b), as shown below:

(68)   *Object between subject and extracted element within expanded phase*
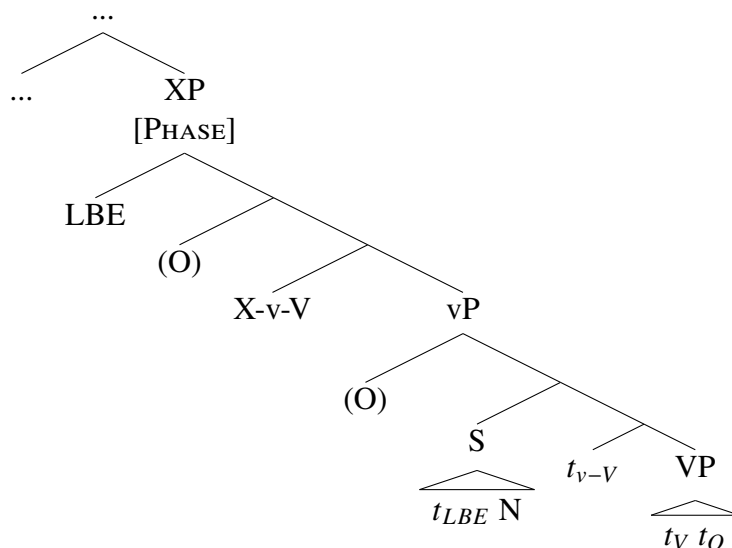
Under this analysis, it is necessary to propose that Russian V typically raises no higher than v in the syntax, but enlarges the local phase when it does so. Consistent with this proposal is Bailyn (2012), who represents the Russian V in v, and Bailyn (1995), for which Russian V moves somewhat, but not as far as T. In contrast, Gribanova (2013, 2017) argues that Russian V always moves as far as an Aspect head above vP for the purposes of morpho-phonological unification. However, Gribanova and Harizanov (2018) raise the possibility of such movement being the result of post-syntactic morphological amalgamation rather than head movement beyond vP in the syntax itself. Given that we do not expect post-syntactic morpho-phonological verb displacement to retroactively expand the local phase beyond vP, this understanding is compatible with the analysis proposed here.

Additional facts that are of interest for this proposal emerge from an examination of the Russian imperfective future tense, which uniquely uses an auxiliary copula. In such a context, it is possible for the auxiliary to precede the subject, in which case LBE from the subject is possible:

(69)   *LBE from subject preceded by auxiliary*

   Každaja$_1$/Èta$_1$ **budet**   [t_1 devočka] **gladit'**   kota?
   every.NOM/this.NOM will.be   girl.NOM to.stroke cat.ACC

   'Will every/this girl stroke a cat?'

In this context, scrambling of an object in between the subject and the extracted constituent is unacceptable:

(70)   *LBE from subject preceded by auxiliary blocked by object scrambling*

    * Každaja$_1$/Èta$_1$ <u>kota$_2$</u>   **budet** [t_1 devočka] **gladit'** t_2?
    every.NOM/this.NOM cat.ACC will.be   girl.NOM to.stroke

    'Will every/this girl stroke a cat?'

However, if both the auxiliary and the main verb precede the subject, scrambling the object to an intervening position becomes acceptable:

(71)   *Subject after auxiliary and verb permits LBE + object scrambling*

    Každaja$_1$/Èta$_1$ <u>kota$_2$</u>   **budet gladit'**   [t_1 devočka] t_2?
    every.NOM/this.NOM cat.ACC will.be to.stroke   girl.NOM

    'Will every/this girl stroke a cat?'

These facts are consistent with the proposal that it is the raising of v (presumably preceded by V to v movement) that importantly extends the phase: only this movement, as opposed to movement of an auxiliary, will carry v to a higher position and thus enlarge the local phase.[22]

4   Conclusion

In this chapter I have argued, expanding on the findings of Antonenko and Davis (2020) and Bondarenko and Davis (To appear), that sub-extraction in Russian obeys two generalizations, repeated below:

(72)   *LBE order preservation generalization*
    LBE from a Russian nominal phrase containing multiple left branches must result in a relative linear ordering that matches the original order that those left branches had prior to any extraction.

---

[22]Even if phase extension is an incorrect analysis of these facts, this concept is generally useful in a CL framework. Without something along these lines, CL predicts that V should never be able to precede an external argument. This is because there is no position within vP where V can precede an external argument before spell-out applies, assuming that such arguments originate in the vP edge. Of course, many languages do allow V to precede an external argument, as we see here for Russian. As touched on at the end of chapter 2, it may be that the verbal domain is more complex than generally assumed in this dissertation. A more fine-grained understanding of this part of the clause will likely be necessary to predict the full range of cross-linguistic variation we see here. See Fox and Pesetsky (2005a,b) for additional discussion of the relevance of this topic to CL.

(73)   *Sub-extraction from subject blocking generalization*
Sub-extraction from a Russian subject (excluding some adjectives and possessors) is banned when other material originating in vP scrambles to a position between the subject and the sub-extracted constituent.

I argued that the first of these generalizations emerges from the phasehood of the Russian nominal phrase, while the second stems from the phasehood of the vP, building on the analysis of similar restrictions in Korean and Japanese from Ko (2007, 2011, 2014). Importantly, both of these results emerge from the interaction of CL with certain independent locality constraints on movement in syntax. These findings are thus thematically parallel to the results of the previous two chapters. Regarding the generalization in (73), the exceptional nature of adjectives and possessors is not decisively solved, but hopefully the discussion provided here appropriately sets the stage for further inquiry.

In the following appendix, I discuss some remaining facts about extraction from PP in Russian that are relevant to the concerns of this chapter, though more complex than can be usefully analyzed in detail here.

5   Appendix: On non-constituent extraction from PP

In chapter 2, we saw that given the complement-to-specifier anti-locality of Abels (2003, 2012), preposition stranding is predicted to be impossible in a language in which PP is a phase: since the complement of PP cannot pass successive cyclically through its edge, the only option is for the entire PP to be pied-piped. We have seen that in Russian P-stranding is indeed banned (except for the more complex adpositions discussed by Podobryaev 2009), as (74) shows once again. I will take this fact as a piece of evidence that the Russian PP is a phase:

(74)   *No preposition stranding in Russian*

    a.   [**O**   **čem**]$_1$ ty govoriš t_1?
          About what you talk?
          'About what are you talking?'

    b.   * **Čem**$_1$ ty   govoriš **o** t_1?
           What you talk   about?
          'About what are you talking?'
          (Podobryaev 2009, ex. 1)

The concept that PP is a phase in Russian helps make sense of a further fact not yet shown in this chapter. A well-known characteristic of LBE in Slavic, reported for

Russian by Pereltsvaig (2008), for instance, is that LBE from prepositional phrases must displace P as well, as in (75):

(75)   *LBE from PP must also displace P*

      a.   **Protiv**$_1$ **sovetskoj**$_2$ on vystupal   [t_1 t_2 **vlasti**].
           against Soviet   he demonstrated   regime
           'It is against the Soviet regime that he demonstrated.'
           (Pereltsvaig 2008, ex. 4a)

      b.   (Tanya Bondarenko, p.c.)

           * **Sovetskoj**$_2$ on vystupal   [**protiv** t_2 **vlasti**].
           Soviet   he demonstrated against   regime

           'It is against the Soviet regime that he demonstrated.'

This limitation is quite strict. The following examples demonstrate this fact for a variety of other prepositions:

(76)   *More instances of P-displacing LBE*

      a.   * **Novyj**$_2$ my pošli [**v** t_2 magazin].
           new   we went to   shop
           'We went to a NEW shop.'

      b.   **V**$_1$ **novyj**$_2$ my pošli [t_1 t_2 magazin].
           to new   we went   shop
           'We went to a NEW shop.'

      c.   * **Kotorym**$_2$ ty igral [**s** t_2 ščenkom]?
           which   you played with   puppy
           'With WHICH puppy did you play?'

      d.   **S**$_1$ **kotorym**$_2$ ty igral [t_1 t_2 ščenkom]?
           with which   you played   puppy
           'With WHICH puppy did you play?'

      e.   * **Malen'kom**$_2$ ja jezdila [**na** t_2 dinozavrike].
           small   I rode   on   dinosaur
           'I rode on a SMALL dinosaur'

      f.   **Na**$_1$ **malen'kom**$_2$ ja jezdila [t_1 t_2 dinozavrike].
           on small   I rode   dinosaur
           'I rode on a SMALL dinosaur'

If the Russian PP is indeed a phase, this fact can be analyzed as an order preservation effect. Since a preposition necessarily precedes any left branches of the nominal

phrase in its complement, any LBE from beneath P will incur an ordering contradiction unless the P can somehow be displaced as well (assuming that concerns of locality prevent successive cyclic LBE through the PP edge).

A difficulty for this analysis is that P does not form a constituent with a left branch of its sister:

(77)   *P and left branch of its sister do not form a constituent*

     PP

   P   NP/DP

    XP   N   ...
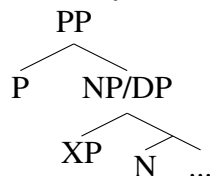
Since targeting a non-constituent for movement is presumably impossible, this raises a question about the syntactic status of the extraction in examples like (76). An additional relevant fact is that P in Russian is capable of independent extraction, as in (78) below. Pereltsvaig (2008) also notes that examples of this form are attested.

(78)   *Long distance P movement in Russian*
   Protiv$_1$ on vystupal   [t_1 **sovetskoj vlasti**].
   against he demonstrated   Soviet   regime
   'It is against the Soviet regime that he demonstrated.'
   (Tanya Bondarenko, p.c.)

Given the possibility of such movement, it could be that in examples like (76) above, order preservation forces independent extraction of P to co-occur with LBE from beneath P, in order to maintain an ordering in which P precedes the content of its sister. However, since heads are not generally capable of non-local movement (Travis 1984), the possibility of examples like (78) is in itself a puzzle.

The possibility of the non-constituent LBE in (76) and the head-only extraction in (78) is predicted by approaches to LBE under which this phenomenon does not actually involve extraction of the seemingly displaced material. For instance, under a *distributed deletion* approach to phenomena like LBE (Faneslow and Ćavar 2002; Bošković 2001, 2015; Fanselow and Féry 2013; Bondarenko and Davis 2019, 2020, a.o.), LBE is the result of movement of an entire nominal or prepositional phrase. The moved phrase only appears discontinuous because of a PF mechanism that causes everything in the moved phrase to be pronounced in the tail of the movement chain, aside from the seemingly extracted material in question. A general schema for distributed deletion is provided in (79):

(79)   *Distributed deletion analysis of LBE*

```
                        ...
              ┌──────────┴──────────┐
        NP/DP/PP₁                   ...
        ┌─────┴─────┐          ┌─────┴─────┐
      X(P)    ...    ...   NP/DP/PP₁
             ◸N̶...◹       ┌────┴────┐
                        X̶(̶P̶)̶    ...
                               ◸N ...◹
```
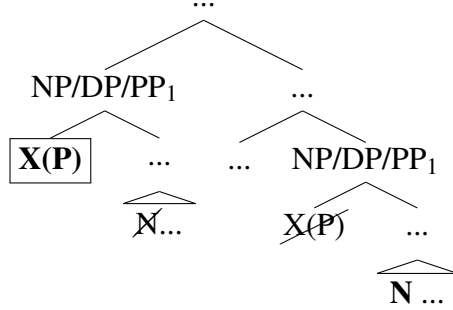
Under this analysis, we do not expect LBE to be limited to phrases, or even to syntactic constituents. Pereltsvaig (2008) argues that for Russian, this prediction is indeed correct.[23] Another account that makes a similar prediction (though different in several details, as Pereltsvaig discusses) is the *remnant movement* approach (Franks and Progovac 1994; Starke 2001; Kayne 2002; Bašic 2008, 2009; Abels 2003, 2012, a.o.). For this analysis, LBE is derived by movement of a remnant phrase that has been previously evacuated by everything but the relevant left branch material, as diagrammed in (80):[24]

(80)   *Remnant movement analysis of LBE*

```
                    ...
          ┌──────────┴──────────┐
      NP/DP/PP₂                 ...
      ┌─────┴─────┐        ┌─────┴─────┐
    X(P)    t₁   ...      ...
                     ┌─────┴─────┐
                   YP₁          ...  t₂
                  ◸N ...◹
```

---
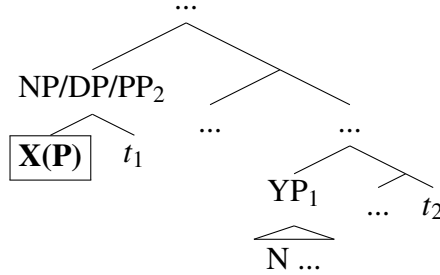
[23]Bondarenko and Davis (2019, 2020) argue in support of Pereltsvaig's conclusion, using the interaction of LBE and parasitic gaps in Russian as the primary diagnostic. Given that LBE in Russian signals contrastive topic or focus, Bondarenko and Davis (2020) suggest that the distributed deletion rule for Russian can be defined as follows:

(i)   When a moved NP/DP/PP contains an element bearing a feature F_{CTF} encoding contrastive topic/focus, realize only the element bearing F_{CTF} in the highest copy of the movement chain, and all other nodes in the lowest copy of the movement chain of that NP/DP/PP.

Defined in this way, distributed deletion is essentially the result of applying the typical PF rule for overt movement to some nodes in a given moved phrase, and the PF rule for covert movement to others. For further discussion of the in principle possibility of distributed deletion, see Nunes (2004).

[24]See Murphy (2020a,b) for recent criticism of the remnant movement approach to LBE.

Since for this analysis the material seemingly displaced by LBE is not actually extracted, there is no reason for LBE to be limited to only phrases, or even constituents.

In this chapter, I assumed at the outset that Russian LBE involves straightforward extraction. If Russian LBE is in fact a product of either distributed deletion or remnant movement, then the primary empirical results of this chapter should still emerge as desired. This is because there is no reason why CL would only be applicable to actual extraction: CL will equally constrain all word order changing movement processes, regardless of their syntactic identity.

Since the generalization in (73) above about LBE from subjects emerges from the interaction of CL with the ban on phrase-bounded edge-internal movement, and since all analyses of LBE depend on movement, (73) will be predicted regardless of which analysis of LBE is assumed (though the fact that adjectives and possessors are exempt from this generalization is puzzling under any analysis of LBE).

The generalization in (72) about LBE from phrases with multiple left branches is straightforwardly predicted by a distributed deletion approach. Order preservation will reject the output of a distributed deletion operation that creates a discontinuous NP/DP/PP whose overt nodes are distributed in a way that contradicts the initial linearization established for the phrase in question prior to its movement. The emergence of (72) is less obvious under a remnant movement analysis of LBE, however. Under such an approach, extraction of certain sub-parts of NP/DP/PP prior to movement of the resulting remnant must be assumed. However, depending on which sub-parts of NP/DP/PP are posited to be mobile, it is conceivable that re-arrangements within the phrase in question prior to its spell-out might be ruled in that would allow the formation of remnants that would, after movement, result in violations of (72). Whether or not this is the case depends on what structure for the NP/DP/PP is assumed, however: it is likely that any excessive re-arrangement prior to movement of the remnant would violate at least one of the locality principles discussed in this chapter. Since establishing this concretely would require detailed analysis of several largely theory-internal logical possibilities, I will not investigate this topic further in this chapter.

I will conclude this appendix by mentioning a final fact about extraction from PP in Russian. We have seen in this chapter (section 3.3) that Russian can form inverted split noun phrases, where the N head appears to be extracted, stranding any left branches it may have had. Split inversion from within PP, just like LBE, requires displacement of P as well for most speakers I have encountered:

(81)   *Obligatory P-fronting with split inversion*

    a.   * **Magazin**$_1$ ja sxodil [**v** dorogoj t_1]

        store   I   went   to expensive

        'I went to an expensive store'

b. $\mathbf{V}_2$ **magazin**$_1$ ja sxodil [t_2 dorogoj t_1]
     to   store   I   went   expensive
     'I went to an expensive store'

c.   * **Dinozavrike**$_1$ ja ezdila [**na** malen'kom t_1]
      dinosaur   I   rode   on   small
     'I rode on a small dinosaur'

d.   **Na**$_2$ **dinozavrike**$_1$ ja ezdila [t_2 malen'kom t_1]
     on   dinosaur   I   rode   small
     'I rode on a small dinosaur'

If the movement step that forms this inversion were able to pass through the edge of the PP phase, then P should be able to be stranded by such displacement. However, if both the nominal phrase and PP are phases in Russian, then the movement step that derives this inversion would need to proceed from the edge of the nominal phrase to the edge of PP, and then onward. Such a derivation will be banned by the form of anti-locality diagrammed in (29) that I adopted from Bošković (2005) and other works cited above. For this reason, the constituent being extracted by split inversion cannot precede P before the spell-out of PP. Consequently, order preservation will require split inversion from a PP to carry along P as well. Since the details of the relevant derivation depend heavily on the analysis that is assumed for this inversion, I will not explore this topic in further detail here.[25]

---

[25]Yet another interesting fact is that when a preposition participates in sub-extraction in Russian, the preposition can be simultaneously pronounced in situ as well:

(i)   a.   U nego lebed' **na kladbišče** tam   **na Novodevičjem** byl.
            to him   swan   on cemetery   there on Novodevichie   was
            'He had a swan there at the Novodevichie cemetery.'
            (Pereltsvaig 2008, ex. 34g)
       b.   **Iz**   **čaški**   ja pila   **iz**   **krasnoj**.
            from cup.F.GEN I   drank from red.F.SG.GEN
            'As for cups, I drank from a red one.'
            (Goncharov 2015: p. 734)

Preposition doubling is also attested with *wh*-movement in informal English (Fitzgibbons 2012):

(ii)   a.   **For** who did you buy this **for**? (Fitzgibbons 2012, ex. 3b)
       b.   **In** which box did you put it **in**?
       c.   Tell me **with** which person you were speaking **with**.

This possibility is interesting for the theory of linearization, since it is conceivable that in the context of CL, such doubling might be predicted to incur an ordering contradiction. This likely depends on precisely how the process of linearization and its evaluation is formalized. I will not investigate this topic further here, but I note the possibility of P-doubling as an intriguing fact for future analysis.

Chapter 5

Overlapping paths and parasitic gaps

1 Introduction

In the previous chapters, we have seen a variety of circumstances where multiple instances of movement interleave and interact. Many of these dealt with extraction from a larger constituent that was previously pied-piped, but we have also seen important interactions between independent movement paths that are not in a sub-extraction relationship. This chapter focuses on certain patterns of the latter variety.

Here I argue that the concepts defended throughout this dissertation shed light on the properties of sentences where multiple instances of (A′-)movement overlap, as well as several intertwined issues about the formation of multiple specifier configurations, the pronunciation of movement chains, and the interpretation of parasitic gaps (PGs). PGs play a particularly important role in this chapter. Here I build from the theory of PG licensing in Nissenbaum (2000), which generalizes to a prediction about how PG interpretation relates to the structure of derivations with overlapping movement paths. I argue that this prediction ultimately provides independent verification of the expectations that CL makes about such derivations. This general result will serve as the basis for exploring the properties of overlapping movement in a variety of contexts.

In principle, there are two possible outcomes for a derivation in which two phrases' movement paths overlap. Either their paths will cross, resulting in no change in their relative order (1a), or their paths will nest, reversing their order (1b):

(1)   *Two possible results for overlapping movement*

      a.   *Crossing paths (Order-preserving)*

         α β ... α β
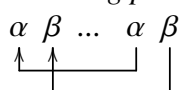
    b.   *Nesting paths (Order-reversing)*

        β α ... α β
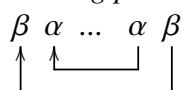
While it is conceivable that only one of these scenarios might be attested in natural language, in reality both are, as the following examples illustrate:

(2)   *Examples of crossing paths*

    a.   *Bulgarian multiple wh-movement*

       Kogo$_1$ kakvo$_2$ e   pital Ivan t_1 t_2?
       who.NOM what.ACC AUX asked Ivan

       'What did Ivan ask who?'
       (Richards 1997 p. 277 ex. 75a)

    b.   *Multiple wh-scrambling in Japanese*

       Dare-ni$_1$ nani-o$_2$ John-ga   [Takana-sensee-ga t_1 t_2 yomaseta
       who-DAT what-ACC John-NOM [Takana-teacher-NOM   read-CAUS
       to   ] itta no?
       that] said Q

       'Who did John say professor Tanaka made read what?'
       (Richards 1997 p. 63 ex. 12b)

(3)   *Examples of nesting paths*

    a.   *Multiple topicalization in Mandarin*

       [Shengfan]$_2$, [lubian   xiao gou]$_1$, wo wei guo t_1 t_2.
       leftover-rice streetside small dog   1SG feed PERF

       'Leftovers, I have fed them to small dogs on the street side'
       (Chen 2017, ex. 27a)

    b.   *Topic/focus movement plus wh-movement in English*
       [This book]$_2$, who$_1$ should we talk to t_1 about t_2?

The simplest way to achieve a configuration with overlapping A′-movement paths in English in a surface-evident way is to combine topic/focus fronting with *wh*-movement, as in (3b) above. Here we see that the landing site of the former movement is above that of the latter. We will see other overlapping movement configurations in English later in this chapter, but for the moment, consider what the derivational history of (3b) might look like.

    Under phase theory, in (3b) the two phrases undergoing A′-movement must have formed intermediate positions in the vP edge before reaching the clause periphery.

Since here there are two moving phrases, there are two conceivable structures for the vP that are plausible for this derivation, illustrated in the simplified diagrams in (4) below: either the topic/focus phrase could form an outer spec-vP, with the *wh*-phrase forming an inner one (4a), or the reverse might occur (4b).

(4)   *Two plausible specifier orders for multiple A′-movement from vP*

a.

```
                    ...
                   /    \
          this book₂     ...
                        /    \
                    who₁      ...
                            /    \
                         ...      vP
                                 /    \
                        this book₂     who₁
                                        /    \
                                     v-V      VP
                                             /    \
                                    ... who₁ ... this book₂
```

b.

```
                    ...
                   /    \
          this book₂     ...
                        /    \
                    who₁      ...
                            /    \
                         ...      vP
                                 /    \
                             who₁      
                                  this book₂
                                        /    \
                                     v-V      VP
                                             /    \
                                    ... who₁ ... this book₂
```
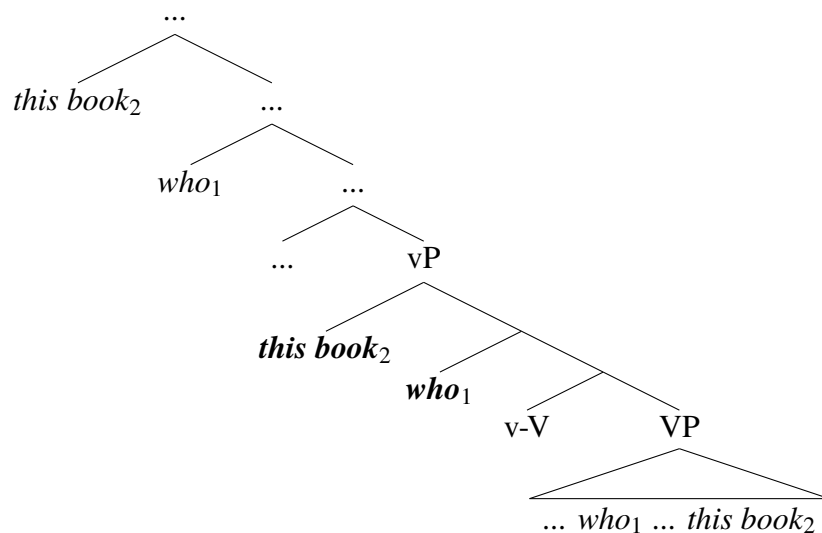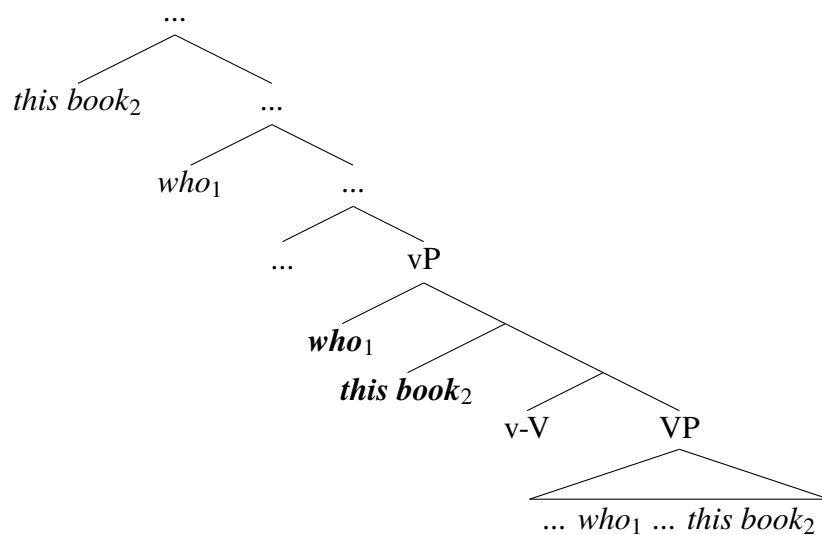
In principle, either specifier order should be sufficient to allow further movement to the left periphery in the sentence under consideration. The CL approach to phases, however, makes a more constrained prediction. Recall that for CL, movement from a phase must be order-preserving: the relative order that any two (or more) elements have at the end of the construction of a given phase must be preserved after any

further movement beyond that phase, or else a linearization contradiction will arise. Under this framework, we expect two phrases whose successive-cyclic movement paths overlap to have the same relative order both in their final landing sites, and in their intermediate positions in the edge of any phase they both pass through. These concerns predict that only the derivation in (4a) above should be available for the sentence in (3b), since this is the only option that maintains a consistent phase-by-phase linearization. Here I will argue that this prediction is correct, and is an instance of a more general pattern, described next.

1.1   Main results of this chapter

The main goal of this chapter is to argue that overlapping movement is constrained by the concerns of order preservation in the way just described. The generalization that I defend here, which is a direct prediction of CL, is stated as follows:

(5)   *Overlapping Chain Generalization*
      If two A′-moving phrases XP_1 and XP_2 form specifiers of an intermediate phase YP such that XP_1 c-commands XP_2, then XP_1 c-commands XP_2 in all subsequent positions these phrases occupy.

The derivation in (4a) above which I will argue is the correct analysis of sentences like (3b) fits (5), since here the moved phrases *this book* and *who* have the same relative order both in their final landing sites, and in their intermediate landing sites as specifiers of the vP. By showing that overlapping successive cyclic movement is indeed generally order-preserving in this way, this chapter amasses further evidence in favor of the CL approach to phase spell-out and linearization.

The key evidence for (5) comes from the behavior of PGs. PGs have been used in much previous work to reveal the nature and distribution of (A′) movement (Engdahl 1983; Nissenbaum 1998, 2000; Legate 2003; Overfelt 2015; Branan 2017; Erlewine and Kotek 2018; Fox and Nissenbaum 2018, a.o.), and will be central here as well, as a diagnostic for the derivational history of overlapping movement configurations. In particular, this chapter will make use of a prediction from Nissenbaum (2000) about PG licensing in multiple movement scenarios, stated in (6) below:

(6)   *Multiple Specifier Single Parasitic Gap Prediction*
      In multiple specifier constructions in which XP_1 is the highest specifier of vP and XP_2 is a low specifier of vP, only XP_1 can license a single parasitic gap in an adjunct to vP.[1]

---

[1]This statement of the prediction is taken from Fox and Nissenbaum (2018), who consider it in the context of certain sentences involving pied-piping which are beyond this chapter's scope.

The framework for PGs from which (6) arises will be discussed in the next section. If the generalization about the ordering of overlapping movements in (5) above is combined with the prediction about PGs in (6), then we expect that in a sentence with overlapping A′-movement chains, only the moved phrase with the highest final landing site should be able to license a PG in an adjunct to a vP crossed by those movements. I show that this is precisely what we find. This result is previewed in (7) below, which modifies (3b) to include a sentential adjunct with a PG:

(7)   *A PG is licensed only by the higher of two overlapped moved phases*
        [This book]$_2$, who$_1$ should we talk to t_1 about t_2...

    a.   ✓ ...before commenting on PG$_2$?

    b.   * ...before arranging a meeting with PG$_1$?

While this result clarifies the derivation of overlapping movement paths, it raises further questions about the formation of multiple specifier configurations. For instance, as discussed in the next section, the derivation I attribute to sentences like (3b) and (7) involves a violation of the hypothesis of Richards (1997, 1999) that all but the first movement to a given head will "tuck-in" to a lower specifier of that head, yielding crossing paths. To resolve this tension, I will argue for a less constrained theory of multiple specifier formation. First, I will propose that tucking-in is not always mandated by multiple movement to one head, but rather is only required for multiple movements to one head that are triggered by the same feature:[2]

(8)   *Proposal #1: One attracting feature for multiple goals → tucking-in*
        $[_{XP}\ \alpha_{[F1]}\ \beta_{[F1]}\ X_{[uF1]}\ ...\ t\ t\ ]$

Second, I will propose that when a head bears multiple movement-triggering features, the goals respectively attracted by those features are free to form crossing (9a) or nesting (9b) paths.[3] I will argue that the hypothesis of free ordering in such contexts is vital to capturing the full range of multiple movement configurations.

---

[2]I follow Chomsky (2000) and a great deal of following work in taking unvalued features (prefixed with "u" in (8-9) below) to be those which probe their c-command domain for a goal bearing the relevant feature, as a prerequisite for establishing an agreement or movement dependency.

[3]The concept that the presence of multiple movement triggering features on one head removes the need to tuck-in has precedent in several works (McGinnis 1998; Rackowski 2002; Doggett 2004; Rackowski and Richards 2005), many of which assume that tucking-in in fact cannot occur in this situation. Here I maintain that tucking-in is possible in this case, but not required, a concept which will receive empirical support from Bulgarian and Romanian in section 4.

(9)   *Proposal #2: Attraction by separate features on one head → variability*

a. $[_{XP}\ \alpha_{[F1]}\ \beta_{[F2]}\ \ X_{[uF1,uF2]}\ \ ...\ \ t\ t\ ]$

b. $[_{XP}\ \beta_{[F2]}\ \alpha_{[F1]}\ \ X_{[uF1,uF2]}\ \ ...\ \ t\ t\ ]$
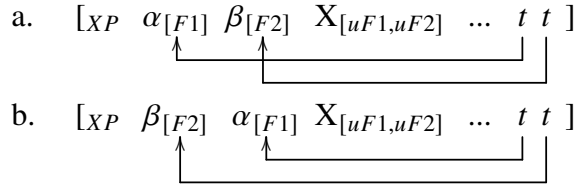
A phenomenon closely related to the above concerns is the *path containment condition* (PCC, Pesetsky 1982, a.o.), which describes the fact that in various languages, English among them, overlapping (overt) A′-movement generally requires the formation of nested dependencies. This is evident in English scenarios like the one in (3b) above, in which the original order of moved phrases must be reversed post-movement, as (10) shows:

(10)   *Forced order reversal for overlapping movement paths*

a.   [This book]$_2$, who$_1$ should we talk to t_1 about t_2?   [2 1 ... 1 2]

b.   * [This student]$_1$, what$_2$ should we talk to t_1 about t_2?   [1 2 ... 1 2]

Richards (1997) hypothesizes that PCC effects are correlated with the possibility of there being only one overt specifier of CP in a given language, and hence, the absence of the PCC in languages where multiple overt specifiers of CP are permitted. Extending this hypothesis, I argue that the PCC emerges from the CL theory when combined with the proposal that in languages like English, inner specifiers formed by A′-movement are sometimes covert (Richards 1997; Nissenbaum 2000, a.o.). I will argue that this covertness results in a linearization problem in certain configurations, which is avoided by running the derivation in such a way that results in order-reversal, as described by the PCC. I will show that this understanding, when combined with the theory of multiple specifier formation described by (8-9), accurately predicts the absence of the PCC in multiple *wh*-movement languages like Bulgarian and Romanian, and fits into a general theory of crossing versus nesting paths.

1.2   Chapter contents
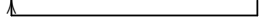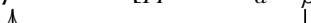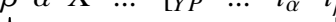
Next, section 2 provides background on the approach to multiple specifier formation in Richards (1997, 1999), and its relationship to the account of PGs in Nissenbaum (2000), leading to a justification for the prediction in (6). Section 3 goes on to show that when (6) is considered in light of the PG asymmetry in contexts like (7), we arrive at evidence for the correctness of the generalization in (5), and thus a verification of CL. Section 4 provides a linearization-based account of the PCC, as a prelude to a revised formalization of the syntactic mechanisms that underlie crossing versus nesting paths, illustrated in section 5. Section 6 goes on to show that

the theory developed here accurately predicts how overlapping movements interact with stranding in phase edges. Section 7 concludes, and is followed by an appendix that discusses a remaining puzzle about multiple *wh*-questions.

2   Background: Multiple specifiers and parasitic gaps

Much of this chapter is concerned with examining the multiple specifier structures that are created when multiple phrases successive cyclically A′-move from a given phase. The possibility of a phrase having multiple specifiers has been considered since at least Chomsky (1995). See also, for instance, Koizumi (1995); Ura (1996); Zwart (1997); Doron and Heycock (1999); Chomsky (2000), and others. As mentioned above, Richards (1997, 1999) offers a hypothesis about how the formation of such configurations is constrained. Richards argues that when two phrases are targeted for movement by one head, the closest phrase is attracted first (11a), while the lower one which is attracted second "tucks-in" to a specifier below the one formed by the first instance of movement (11b). The non-initial movement must tuck-in rather than move to a position above the phrase that moved first (11c) because of an economy constraint dictating that movement paths must be as short as possible.

(11)   *Non-initial instances of movement to one head must tuck-in*

       a. $[_{XP}\ \alpha\ \mathrm{X}\ ...\ [_{YP}\ ...\ t_\alpha\ \beta\ ]]$

       b.   ✓ $[_{XP}\ \alpha\ \beta\ \mathrm{X}\ ...\ [_{YP}\ ...\ t_\alpha\ t_\beta\ ]]$

       c.   * $[_{XP}\ \beta\ \alpha\ \mathrm{X}\ ...\ [_{YP}\ ...\ t_\alpha\ t_\beta\ ]]$

Richards argues that such a derivation is instantiated by sentences like those in (2) above, where multiple phrases move to adjacent positions in order-preserving fashion. For Richards, in these sentences the final linear order of moved phrases reflects the order in which the relevant attracting head targeted them, since the pressure to tuck-in ensures that the lower of the two moved phrases, which moves second, ends up in a position below the one that moved first. Richards' theory accurately predicts that the sentences in (2) are unacceptable if the moved phrases reverse their order, as (12) below shows:

(12)   *Unacceptable non-tucking-in variants of the sentences in (2)*

      a.   *Bulgarian multiple wh-movement*

          * Kakvo$_2$ kogo$_1$ e   pital   Ivan t_1 t_2?
           what.ACC who.NOM AUX asked Ivan

          'What did Ivan ask who?'
          (Richards 1997, p. 277 ex. 75b)

      b.   *Multiple wh-scrambling in Japanese*

          * Nani-o$_2$ dare-ni$_1$ John-ga   [Takana-sensee-ga t_1 t_2
           what-ACC who-DAT John-NOM [Takana-teacher-NOM
          yomaseta to   ] itta   no?
          read-CAUS that] said Q

          'Who did John say professor Tanaka made read what?'
          (Richards 1997, p. 63 ex, 12c)

While tucking-in is evidently a requirement for the contexts in (2/12), in this chapter we will encounter evidence that a tucking-in strategy is not always required for movement of multiple phrases to the same position. For this reason, as previewed in the introduction, I will argue for an alternative theory of multiple specifier formation. The relevant evidence will come from the behavior of PGs in a variety of English sentences like (3b) above, repeated in (13) below, in which multiple phrases overtly A′-move through the vP on their way to the clause periphery. As discussed above, the PCC applies to such sentences in English. While the PCC will be addressed in detail later in this chapter, for the meantime it will be useful to set aside the fact that the PCC holds and simply consider what the tucking-in theory of Richards (1997, 1999) predicts about how such a sentence was derived.

(13)   *Topic/focus movement + wh-movement*
       [This book]$_2$, who$_1$ should we talk to t_1 about t_2?

Under Richards' hypothesis about tucking-in, the vP-level derivation of the sentence in (13) should proceed as follows. First, the originally higher A′-moved phrase *who*$_1$ should form an initial A′-specifier of vP:[4]

---

    [4]I generally diagram A′-movement of non-subjects as passing through a position in vP above the subject in situ, as we see in (14-17). This order of constituents has independently been assumed by a variety of works which take tucking-in below thematic specifiers to be banned or at least dis-preferred (McGinnis 1998; Nissenbaum 2000; Chomsky 2001; Rackowski 2002; Rackowski and Richards 2005), and is also predicted by CL, since a non-subject A′-moving from vP must precede the subject at the spell-out of vP just as it will in its final landing site in CP. However, this chapter predicts tucking-in below the subject in situ in certain cases, such as when the subject participates in

(14)   *Originally higher moved phrase forms higher spec-vP*
         [$_{vP}$ *who*$_1$ S v-V [$_{VP}$ t_1 *this book*$_2$]]

Second, the initially lower phrase *this book*$_2$ should tuck-in to a spec-vP below *who*$_1$:

(15)   *Originally lower moved phrase forms lower spec-vP*
         [$_{vP}$ *who*$_1$ *this book*$_2$ S v-V [$_{VP}$ t_1 t_2]]

From their intermediate positions in the edge of the vP phase, these phrases will later continue on to their respective final landing sites in the clause's periphery:

(16)   *Movement to left periphery after tucking-in at vP*

         [*this book*$_2$ *who*$_1$ S T [$_{vP}$ *who*$_1$ *this book*$_2$ t_S v-V [$_{VP}$ t_1 t_2]]]

We have just seen the derivation that Richards' theory of tucking-in predicts. In contrast, this theory excludes a derivation in which the initially lower A′-moving phrase forms a specifier of vP above the phrase that moved first, resulting in these phrases' paths nesting within vP prior to their later movement onward:

(17)   *Multiple A′-movement from vP without tucking-in (=nesting paths in vP)*

         [*this book*$_2$ *who*$_1$ S T [$_{vP}$ *this book*$_2$ *who*$_1$ t_S v-V [$_{VP}$ t_1 t_2]]]

Building from the prediction about PGs introduced in (6) above, I will argue that this non-tucking-in derivation is in fact the correct one. Notice that this derivation fits what (5) states, as CL predicts: here *this book* precedes *who* both in these phrases' final landing sites, and in their intermediate vP landing sites. Next, I will overview the components of Nissenbaum (2000) that justify the prediction in (6) and thus will allow the evidence for this conclusion to become apparent.

2.1   Parasitic gaps, successive-cyclicity, and specifier order

PGs have been mentioned in a few circumstances in the preceding chapters. To review, a PG is a gap whose interpretation depends on A′-movement external to and structurally crossing over the constituent that contains it, among other requirements that we will see in detail shortly. That a given gap is indeed "parasitic" in the relevant

---

the formation of a multiple *wh*-question, as discussed in section 5.3 below.

way is clearest when in an island, since in such a context, it is clear that the gap was not simply derived via typical extraction of some variety.[5] In (18) below, for instance, we see PGs in the object position of sentential adjuncts that are licensed by *wh*-movement from a position external to those adjuncts.

(18) *PG in adjunct island*

    a.   Who$_1$ did you forget about t_1 [because I didn't mention PG$_1$]?

    b.   Tell me what$_1$ I should reread t_1 [before giving you comments on PG$_1$].

Since such adjuncts are generally islands, as (19) shows, it is clear that the PGs shown in (18) do not involve movement from the adjuncts:[6]

(19) *Adjunct island*

    a.   *?? Tell me what$_1$ you ate fried chicken for lunch [before giving them comments on t_1].

    b.   *?? Who$_1$ do you think I'm unreasonable [because I don't like t_1 at all]?

PGs in sentential adjuncts will be the focus of this chapter, since these are the most straightforward PG-hosting constituents that interact with movement within the clause in the relevant way. PGs in subjects will be briefly addressed as well.

    The licensing of PGs depends on several requirements. For instance, the PG-containing island must be adjoined to a structural position crossed by the licensing movement chain (20a), and not below the gap of the licensing movement (20b) or above the landing site of the licensing phrase (20c), as observed by at least Kayne (1983) and Longobardi (1984):

(20) *PG-containing island must be within the movement path of the licenser*

    a.   *Island within licensing movement path*
            ✓ Who$_1$ did you tell t_1 about our idea [in order to impress PG$_1$]?

    b.   *Island below licensing movement path*
            * Who$_1$ will you tell t_1 [that I was a criminal [before I met PG$_1$]]?

---

[5]It is of course possible for a PG to appear in a non-island, as in *Who did Mary send a friend of __ a cool picture of __?*, though in such cases it is difficult to be certain which is the "true" gap, and which is the parasitic one. For this reason, I constrain this study to PGs in islands.

[6]We can also tell that the PGs in (18) are not actually omitted objects, since *mention* in (18a) is obligatorily transitive, and the preposition *on* in (18b), like all English prepositions, is as well:

(i)   a.   * Did you mention __ yesterday?

        b.   * Please give me comments on __ soon.

c. *Island above licensing movement path*
    * I [explained [[which cake]$_1$ I have a recipe for t_1] [while eating a piece of PG$_1$]].

Additionally, this property must hold of the smallest island containing the PG. Separating the PG-containing island from the licensing movement chain with a second island results in unacceptability (Kayne 1983; Chomsky 1986):

(21) *PG licensing across multiple islands fails*[7]

a. *Relative clause island plus adjunct island*
    * Tell me who$_1$ you talked to t_1 [after meeting a person [who likes PG$_1$]].

b. *Subject island plus adjunct island*
    * Durian is a fruit [which$_1$ I ate t_1 for the first time [after [a fan of PG$_1$] visited me]].

c. *Adjunct island in adjunct island*
    * Guess who$_1$ I ironically ran into t_1 [after taking the other hallway [because I wanted to avoid PG$_1$]].

Nissenbaum (2000) argues that these facts are captured by a theory in which PGs are formed by movement of a null operator (OP) to the edge of the containing island (Chomsky 1986, a.o.),[8] as in (22) below, in combination with the proposal that for semantic reasons, the island adjoins to a landing site of (successive cyclic) A′-movement. The theory along these lines that Nissenbaum develops, which I describe in detail next, makes predictions about the behavior of PGs in multiple specifier contexts that will be central to this chapter's concerns.

(22) *PG as trace of null operator*
       Who$_1$ did you forget about t_1 [OP after talking to t_{OP}(=PG)]?

Nissenbaum argues that a sentence like (22) is formed and interpreted in the following way. Within the sentential adjunct island, operator movement to its edge triggers the semantic rule of Predicate Abstraction (Heim and Kratzer 1998), which changes the island into a predicate of type <e,t>, as in (23):[9]

---

[7]As discussed in chapter 7, PG licensing is possible in a relative clause of a subject, both of which are generally regarded as islands. This is a challenging puzzle that will need further study.

[8]This is in contrast to "shared antecedent" theories of PGs, under which a PG is not the trace of a null operator, but simply a typical gap. See Kayne (1983) for an example of work in this direction, as well as Nissenbaum (2000) and Nissenbaum and Schwarz (2011) for criticisms of it.

[9]In this chapter I adopt from Nissenbaum the simplifying assumption that vPs and vP modifiers (such as the sentential adjuncts under consideration) are type t, modulo A′-movement within them

(23)   *Null operator movement inside adjunct forms a derived predicate*

            ...

      ...   AdjunctP
                  <e,t>

     OP_1 λ_1 after talking to t_{OP_1}(=PG)
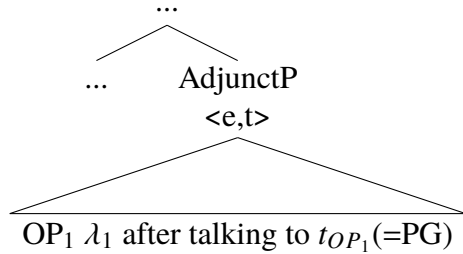
This operator movement within the adjunct is the first ingredient for PG licensing. The second ingredient is successive cyclic movement of the licensing phrase *who* in (22) through the edge of vP, which is independently necessary for phase-theoretic reasons. This intermediate step of A′-movement triggers an application of Predicate Abstraction there as well, creating an <e,t> node in the vP edge. This is diagrammed in the partial structure in (24) below. Here we see that the resulting <e,t> node is sister to the type e trace left behind in the vP edge by successive cyclic movement of *who* from this domain:

(24)   *Successive-cyclic A′-movement creates an <e,t> node in vP*

            ...

      ...   vP
                 t

     t_{wh_2} <e,t>
      e

        λ_2 you forgot about t_{wh_2}
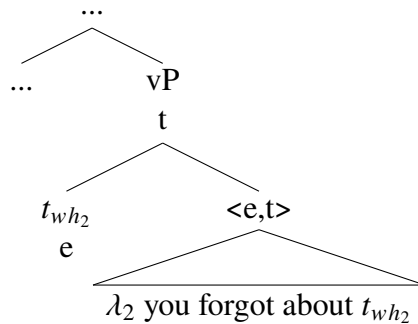
The PG-containing adjunct island in (23) above is type <e,t>, and as we've just seen in (24), a node of the same type exists in the vP edge after successive-cyclic A′-movement from it. Thus this adjunct can be externally merged as the sister of the derived <e,t> position in the vP, resulting in a structure that is successfully interpreted via Predicate Modification (Heim and Kratzer 1998), as in (25):

---

triggering Predicate Abstraction. As Nissenbaum (2000, p. 47) notes, this is a simplification because it ignores the presence of temporal and event arguments, but enriching the semantic type of these constituents does not make any important difference for the account of PGs. In particular, given a more general version of the rule of Predicate Modification that allows constituents of the same semantic type to be combined (which is independently required to handle adjuncts containing multiple PGs, as we'll see shortly), the account described here functions the same regardless of whether the type of vPs and their adjuncts is t, or something more complex.

(25)   *Predicate Modification of vP with PG-containing island*

```
                    ...
              ╱           ╲
           ...              vP
                            t
                        ╱        ╲
                   t_wh₂        ┌──────┐
                    e          │ <e,t> │
                               └──────┘
                              ╱         ╲
                       <e,t>              AdjunctP
                       ╱   ╲               <e,t>
                      ╱     ╲             ╱     ╲
            λ₂ you forget about t_wh₂    ╱       ╲
                                OP₁ λ₁ after talking to t_OP₁(=PG)
```

t_{wh_2}, $<e,t>$, AdjunctP $<e,t>$, λ_2 you forget about t_{wh_2}, OP$_1$ λ_1 after talking to t_{OP_1}(=PG)
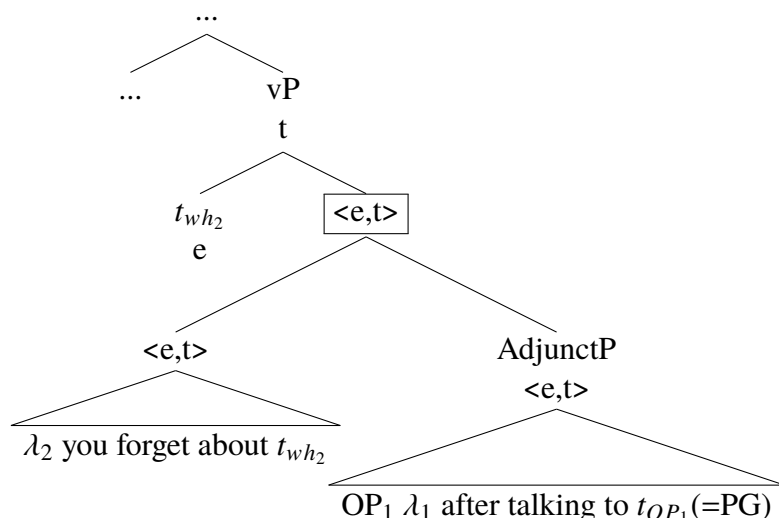
Nissenbaum provides arguments from phenomena including binding, NPI licensing, VP ellipsis, and VP fronting to establish that the PG-containing adjunct in such contexts is indeed adjoined at a clause-medial position, which he takes to be the vP edge, as (25) shows. In this structure, the (boxed) intermediate vP node created by merge of the adjunct to the site of successive-cyclic movement in vP denotes a function of type <e,t>, which is true of individuals that the addressee both talked to and forgot about. The intermediate type e trace of the A′-moved nominal phrase will saturate the individual argument of this function, "filling in" both the variable that corresponds to its trace in VP, and the trace of the null operator in the adjunct, resulting in a PG that is successfully interpreted.[10]

    This theory of PGs captures the fact that a PG can only be licensed if the smallest island that contains it is adjoined within the licenser's movement path. It is semantically necessary for the PG-forming operator to reach the edge of the island that will adjoin within the licenser's movement path, but if an embedded island intervenes,

---

[10]This exposition has assumed that successive cyclic A′-movement (and the corresponding Predicate Abstraction) occurs before merge of the PG-bearing adjunct in (25). In principle it is also plausible for merge of the adjunct to precede A′-movement, however. In this case, if Predicate Abstraction simply applies to the sister of the specifier of A′-movement in the fashion posited by Heim and Kratzer (1998), then the segment of vP to which the PG-bearing adjunct adjoins would not be type <e,t> as shown in (25), and Predicate Modification would fail. This incorrect hypothetical derivation is appropriately excluded by Nissenbaum's hypothesis that Predicate Abstraction does not apply to the sister of the specifier created by movement, but rather, to the lowest saturated projection of the phrase that movement targets. This concept is vital for the analysis of PGs in structures where multiple phrases successive cyclically move through vP: as we'll see shortly, in such contexts it is necessary for Predicate Abstraction to create open semantic argument positions that are not directly local to the corresponding moved phrases. See Nissenbaum (2000) for further discussion.

such movement fails. Furthermore, if the PG-containing constituent is semantically a predicate that must undergo Predicate Modification with a landing site of the licensing phrase, then it is straightforward why it cannot be interpreted in a position outside of the licenser's movement path. This account also accurately predicts that in a context with embedded islands, PG-forming operator movement within the larger island can feed the presence of an additional PG within the embedded island:

(26)   *Recursive PG licensing*

a.   John's a person who$_1$ I hang out with t_1 [because **[friends of PG$_1$]** admire PG$_1$].
(Adapted from Nissenbaum 2000, p. 26, ex. 13c)

b.   Guess [which food]$_1$ I can't help but eat t_1 [despite really fearing PG$_1$ **[because there are carcinogens in PG$_1$]**].

Next we will see how Nissenbaum relates this theory to contexts with multiple specifiers of vP, and thus makes the prediction that this chapter's results depend on.

2.1.1   Multiple specifiers and PG licensing

Nissenbaum's examination of PGs in multiple specifier structures focuses on multiple *wh*-questions. In English, such questions involve overt movement of whichever *wh*-phrase is initially highest (unless D-linking occurs), as (27a) below shows. This is in contrast to many Slavic languages like Bulgarian (27b) and Romanian, in which all *wh*-phrases in such contexts overtly move. In these languages lower *wh*-phrases land in a position below the initially highest one, an ordering that Richards (1997, 1999) attributes to tucking-in:

(27)   a.   *English multiple wh-question*
Who$_1$ did Ivan ask t_1 **what**?

b.   *Bulgarian multiple wh-question*

Kogo$_1$ **kakvo**$_2$ e   pital   Ivan t_1 t_2?
who.NOM what.ACC AUX asked Ivan

(Richards 1997 p. 277 ex. 75a)

Nissenbaum follows Richards in arguing that in both types of languages, all *wh*-phrases move in multiple *wh*-questions, the difference being that in English the movement responsible for forming the tucked-in specifier(s) is covert:

(28)   *Covert tucked-in wh-movement in English*
Who$_1$ $\varnothing_2$ did Ivan ask t_1 **what**$_2$?

If seemingly in situ *wh*-phrases truly do move, we might expect them to be capable of PG licensing. As Engdahl (1983) observed, it appears that they are not:

(29)   *Wh-in-situ cannot license a PG*
         * I forget who filed [**which articles**]$_1$ [without reading **PG**$_1$].
         (Engdahl 1983, ex. 34)

However, Nissenbaum observed that when overt *wh*-movement in English licenses the first of two PGs in one island, an in situ *wh*-phrase can license the second PG:

(30)   *Wh-in-situ licenses a second PG in the same island*
         [Which senator]$_1$ did you persuade t_1 to borrow [**which car**]$_2$ [after getting an opponent of PG$_1$ to put a bomb in **PG**$_2$]
         (Nissenbaum 2000, p. 12 ex. 8)

Thus Nissenbaum argues that English in situ *wh*-phrases do move. Only the licensing of the second of two PGs allows such movement to become apparent in English, Nissenbaum argues, because of the semantic complexities that arise when there are multiple specifiers of A′-movement. Next I summarize Nissenbaum's theory of the interpretation of multiple specifiers, which will lead us directly to the diagnostic prediction that this chapter will make central use of.

Importantly, Nissenbaum argues that when multiple phrases A′-move to specifiers of one head, the lowest saturated segment of that head's projection undergoes one instance of Predicate Abstraction for each moved phrase (31). Thus when vP is targeted by two intermediate A′-movements in a multiple *wh*-question, the vP segment that was the root before movement (labeled β in (31) below) is interpreted as a two-place predicate <e,<e,t>>. In this context, the trace in the lower specifier of A′-movement saturates the first argument of this two-place predicate, yielding a one place predicate (labeled α below) which is in turn saturated by the higher specifier's trace of A′-movement, resulting in a type t for the root node of the vP:[11]

---

[11]Since the trace of movement through the inner spec-vP (t_{wh2}) in (31) is the sister of β, the two semantic argument positions of the function denoted by β must have an order that is the opposite of the two specifiers of vP that saturate it (t_{wh1} t_{wh2} ... λ_2 λ_1). This ensures that when t_{wh2} combines with β, it correctly binds its corresponding trace in VP, rather than that of the outer specifier. After this occurs the result is α, a one-place predicate whose saturation by the trace of the outer spec-vP (t_{wh1}) binds the corresponding trace in VP as well, resulting in the correct interpretation.

The same relative ordering of moved phrases and corresponding semantic argument positions must also hold for the configuration discussed in (36) below, which involves multiple PG-forming operators. If the two null operators in a multi-PG structure like (36) below must always tuck-in (as Nissenbaum proposes) and apply Predicate Abstraction in reverse order in the way just described, we accurately predict Nissenbaum's observation that in a multiple PG adjunct as in a sentence like (30), the order of gaps and licensers must match, as shown in (i) below:

(31)   *vP after two successive-cyclic movements*

$$
\begin{array}{c}
vP \\
t \\
\diagup\quad\diagdown \\
t_{wh_1}\qquad \alpha \\
e\qquad\quad <e,t> \\
\diagup\quad\diagdown \\
t_{wh_2}\quad \beta \\
e\quad\; \boxed{<e,<e,t>>} \\
\diagup\qquad\qquad\diagdown \\
\lambda_2\,\lambda_1\;(S)\;\text{v-V}\;t_{wh_1}\;t_{wh_2}
\end{array}
$$

If an adjunct with one PG is type <e,t> as shown in (23) above, then such an adjunct can only combine via Predicate Modification with the node α in the multiple specifier structure in (31). Doing so results in the configuration in (32):

(32)   *Licensing of single PG by (trace of) A′-movement through higher spec-vP*

$$
\begin{array}{c}
vP \\
t \\
\diagup\quad\diagdown \\
t_{wh_1}\quad \boxed{<e,t>} \\
e\qquad\quad \diagup\quad\diagdown \\
\alpha\qquad\qquad \text{AdjunctP} \\
\boxed{<e,t>}\qquad\qquad \boxed{<e,t>} \\
\diagup\quad\diagdown\qquad\qquad \diagup\quad\diagdown \\
t_{wh_2}\quad \beta\qquad OP_3\;\lambda_3\;...\;t_{OP_3} \\
e\quad\; <e,<e,t>> \\
\diagup\qquad\qquad\diagdown \\
\lambda_2\,\lambda_1\;(S)\;\text{v-V}\;t_{wh_1}\;t_{wh_2}
\end{array}
$$
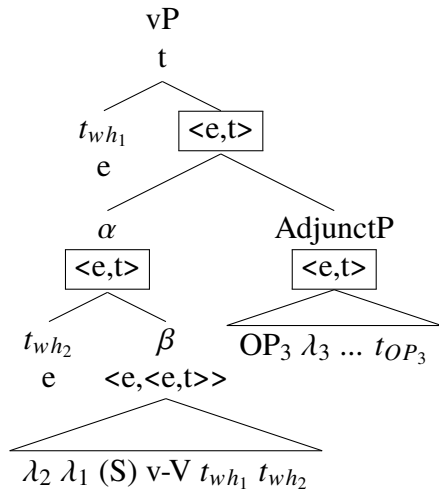
The node in (32) produced by merger of the PG-containing adjunct to α is a predicate that is saturated by the intermediate trace in the higher specifier of A′-movement (t_{wh_1}). Thus here the A′-moving phrase that formed the outer spec-vP licenses the

---

(i)   a.   ? [Which senator]$_1$ did you persuade t_1 to borrow [which car]$_2$ [after getting an opponent of PG$_1$ to put a bomb in PG$_2$]?
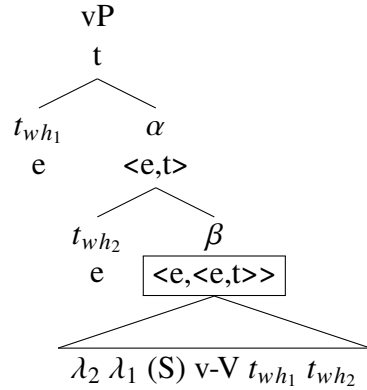   (Nissenbaum 2000, p. 107 ex. 21a)
    b.   * [Which car]$_1$ did you lend t_1 to [which senator]$_2$ [after getting an opponent of PG$_2$ to put a bomb in PG$_1$]?
                (Nissenbaum 2000, p. 110 ex. 27a)

Since this topic is not vital to this chapter, I suspend further discussion of this until chapter 7.

PG. This derivation appropriately models a sentence where the initial movement in an English multiple *wh*-question, which is overt, successfully licenses a PG:

(33)   *Overt wh-movement in multiple wh-question can license PG*
        Who$_1$ $\varnothing_2$ did Reginald ask t_1 about **what**$_2$ [in order to have a good reason to talk to **PG**$_1$]?

In contrast, as we saw in (29) above, an in situ *wh*-phrase in a multiple *wh*-question cannot license a lone PG. If the in situ phrase covertly moves to an inner specifier of the vP (and later on the CP), what has been said above accurately predicts this restriction. In order for the inner specifier of A′-movement in (32) above to license the PG, the PG-containing adjunct would have to merge with the node β, which is the sister of t_{wh_2}. If this were possible, t_{wh_2} would saturate the function output by Predicate Modification of the adjunct with β, and thus license the PG. However, in reality such a structure is not interpretable, as we see in (34) below. Since β is a segment of vP that is a two-place predicate due to the occurrence of two A′-movements, the one-place predicate instantiated by an adjunct with one PG cannot combine with β either by Predicate Modification, or by Functional Application. Thus the inner specifier of A′-movement cannot license the PG here:

(34)   *Type mismatch prevents licensing of single PG by lower spec-vP*
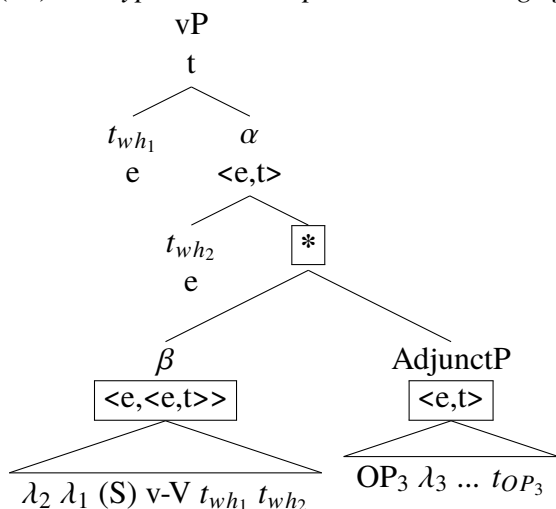
These concepts lead us to the correct prediction that, when overt *wh*-movement licenses the first of two PGs in one island, a covertly moving *wh*-phrase can license the second PG, as (30) above showed. Assuming that Predicate Modification can apply to any two phrases of the same semantic type (in the spirit of Partee and Rooth 1983, see also Nissenbaum and Schwarz 2011), the (trace of the) inner specifier of A′-movement in (34) will be able to license a PG in a situation where the PG-

containing adjunct is a two-place predicate, and thus can be interpreted by merging with β. This is achieved if the adjunct contains two PG-forming operators, which each move to the edge of the adjunct and each trigger an instance of Predicate Abstraction, with the result that the adjunct is type <e,<e,t>> as needed, as in (35):

(35)   *Movement of two operators inside adjunct facilitates two PGs*

$$\ldots$$

$$\ldots \quad \text{AdjunctP}$$
$$\text{<e,<e,t>>}$$

$$\text{OP}_3 \text{ OP}_4 \, \lambda_4 \, \lambda_3 \ldots t_{OP_3} \, t_{OP_4}$$
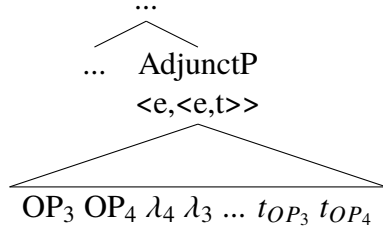
Merge of this two PG adjunct to β in (31) allows both specifiers of A′-movement to license one PG each, as (36) diagrams, and as the facts have shown us:

(36)   *vP after two successive-cyclic movements with a two PG adjunct*

$$\text{vP}$$
$$\text{t}$$

$$t_{wh_1} \qquad \alpha$$
$$\text{e} \qquad \text{<e,t>}$$

$$t_{wh_2} \quad \boxed{\text{<e,<e,t>>}}$$
$$\text{e}$$

$$\beta \qquad\qquad \text{AdjunctP}$$
$$\boxed{\text{<e,<e,t>>}} \qquad \text{<e,<e,t>>}$$

$$\lambda_2 \, \lambda_1 \, \text{(S) v-V} \, t_{wh_1} \, t_{wh_2} \qquad \text{OP}_3 \text{ OP}_4 \, \lambda_4 \, \lambda_3 \ldots t_{OP_3} \, t_{OP_4}$$
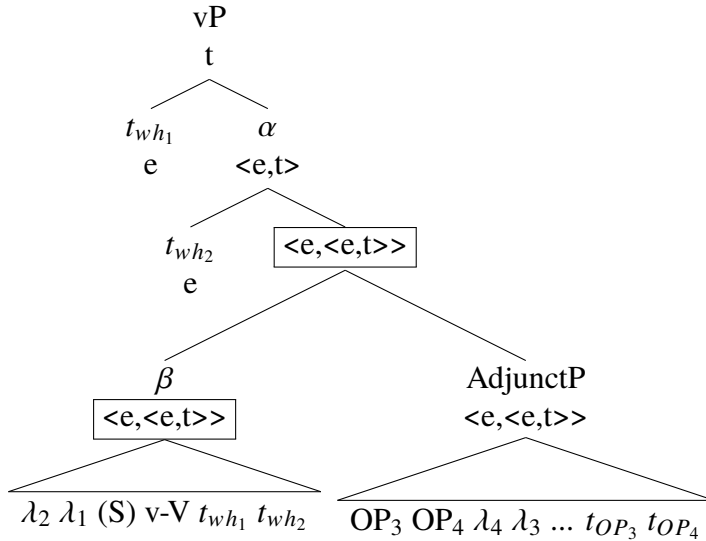
In summary, if covert *wh*-movement in English is a product of tucking-in to a lower specifier, the account described above predicts the fact that such movement cannot license a lone PG, but can license the second of two PGs contained in one island. Abstracting away from the fact that inner specifier formation happens to be covert in English multiple *wh*-questions, this understanding predicts that, in any multiple specifier configuration, the lower of two specifiers of A′-movement will be unable to license a lone PG. Consequently in any structure like (37) below in which α and β are both specifiers of successive-cyclic A′-movement from vP, we expect a single PG to be able to be licensed only by the higher specifier α, regardless of the

phonological status of the movements involved:

(37)   *Only the outermost A′-specifier of vP can license a lone PG*
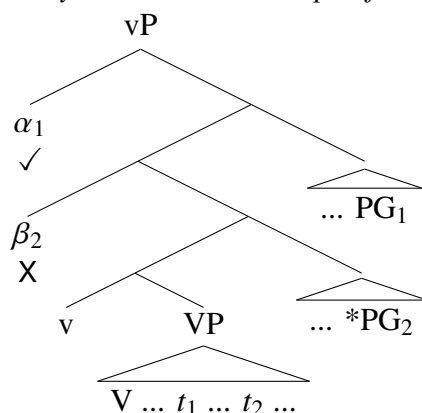
This is precisely as described by the prediction previewed at the beginning of this chapter, repeated below:

(38)   *Multiple Specifier Single Parasitic Gap Prediction*
     In multiple specifier constructions in which XP_1 is the highest specifier of vP and XP_2 is a low specifier of vP, only XP_1 can license a (single) parasitic gap in an adjunct to vP.

With an explicit prediction about the relationship between PGs and multiple specifiers now justified, in the next section, I show how these concerns reveal that overlapping moved phrases have the same relative order in both their intermediate and final positions, as described by (5) above. This finding both verifies the predictions of CL and raises questions about when tucking-in applies, which will extend to a number of topics about the properties of overlapping movement paths.[12]

---

[12]This dissertation has made use of the concept of successive cyclic movement from vP in a wide variety of circumstances. In many situations, it is difficult to be certain that successive cyclicity is required here as phase theories predict, rather than merely allowed. The concepts discussed in this section make it possible to provide argument that such movement is obligatory, based on further observations from Nissenbaum (2000). Observe (i), in which an adjunct contains two PGs—one licensed by heavy NP shift past the adjunct, and one licensed by *wh*-movement:

(i)   ? [Which book]$_1$ did Smith find t_1 on top of t_2, [after putting a copy of PG_1 next to PG_2], [the table in the corner]$_2$?
     (Nissenbaum 2000, p. 113, ex. 33)

Nissenbaum argues that heavy NP shift licenses PGs just like more typical leftward movement (an argument recently corroborated by Overfelt 2016). That is to say that heavy NP shift triggers Predicate Abstraction within vP just as leftward movement does, thus creating a position to which a

3   PGs reveal that final and intermediate orders match

To begin taking stock of the relevant facts, first consider once more an English sentence where topic/focus fronting and *wh*-movement are combined, such as (39a) below. In such examples the PCC holds (39b) and topic/focus movement must land above *wh*-movement as mentioned earlier, even if the PCC is obeyed (39c),[13] but set these constraints aside for the moment in order to focus on how PG licensing behaves in such configurations.

(39)   *Topic/focus movement plus wh-movement in English*
        a.   [This book]$_2$, who$_1$ should we talk to t_1 about t_2?
        b.   * [This student]$_1$, what$_2$ should we talk to t_1 about t_2?

PG-containing adjunct can be adjoined. In (i), in addition to heavy NP shift, *wh*-movement from vP also triggers an instance of Predicate Abstraction in this structure. These two instances of abstraction combine to create a two-place predicate within vP, and as expected, this allows the adjunct here to contain two PGs. Importantly, if either *wh*-movement or heavy NP shift here had the option of exiting vP non-successive-cyclically, then it should be possible for only one of these movements to trigger Predicate Abstraction, which would allow the merger of an adjunct containing just one PG. As Nissenbaum points out, however, in this context both of the PGs in the adjunct are obligatory:

(ii)   a.   * [Which book]$_1$ did Smith find t_1 on top of t_2, [after reading a review of PG$_1$], [the table in the corner]$_2$?
        b.   * [Which book]$_1$ did Smith find t_1 on top of t_2, [after wiping PG$_2$ with a sponge], [the table in the corner]$_2$?
          (Nissenbaum 2000, p. 114, ex. 37)

The obligatory presence of both PGs here signals the obligatory triggering of abstraction by both movements from vP, and thus, indicates that these movements were required to form intermediate landing sites in the vP edge. While Nissenbaum does not explicitly discuss this pattern in these terms, the account of it entails obligatory successive cyclicity, which phase theories predict.

    Another issue is raised by sentence (iib). If here it were possible for the *wh*-phrase to form the inner spec-vP and for heavy NP shift to form the outer one, with the adjunct sitting in between these two phrases, then we would expect heavy NP shift to be able to license the PG here. This configuration would be just like that in (32) above, if the outer specifier t_{wh_1} in (32) were a rightward linearized specifier of heavy NP shift. The fact that PG licensing is impossible in (iib) suggests that heavy NP shift must form an inner specifier of the vP. Notice that CL does not provide any reason why this should be so: the linear order of the moving *wh*-phrase and the phrase undergoing heavy NP shift are the same regardless of their structural relationship in the vP, since they surface on opposite sides of it. If heavy NP shift does not move beyond the vP itself (following the Right Roof Constraint, Akmajian (1975) a.o.), perhaps what matters here is that the relative c-command relationship between the two phrases be kept the same both in vP, and after further movement of the *wh*-phrase to spec-CP—a result that sits better with theories of shape preservation (Williams 1999, 2003; Müller 2002) than linearization preservation.

    [13]This rigidity of final landing sites is what we expect if such movements target dedicated phrases in the clausal periphery (Rizzi 1997, and many others), though see further footnote 15 below.

c. * What$_1$ [this person]$_2$ should we talk to t_2 about t_1?

We have seen in numerous preceding examples that *wh*-movement can license a PG. This is also possible for topic/focus movement, since PG licensing is available for all A′-movements, regardless of sub-type (Culicover and Postal 2001):

(40) *Topic/focus movement can license a PG*

a. [This book]$_2$, we should talk to someone else about t_2 [before writing our review of PG$_2$].

b. [These papers]$_2$, I asked the TA to grade t_2 [after making a lot of copies of PG$_2$].

Importantly, as previewed in the introduction, combining these two types of A′-movement in one clause restricts the possibilities for PG licensing. This fact is illustrated in the examples in (41-42) below. We see here that only the A′-moved phrase with the higher final landing site can license the single included PG, even when we ensure that the sentence would be semantically and pragmatically well-formed if the inner one were the licenser:

(41) *Only outer A′-moved phrase can license a PG*
[This book]$_2$, who$_1$ should we talk to t_1 about t_2...

a. ✓ ...before commenting on PG$_2$?
b. * ...before arranging a meeting with PG$_1$?

(42) *Only outer A′-moved phrase can license a PG*
[These papers]$_2$, [which TA]$_1$ did you ask t_1 to grade t_2...

a. ✓ ...before making copies of PG$_2$?
b. * ...after running into PG$_1$ in the elevator?

Given the prediction in (38) above, the fact that the outer moved phrase can license the PG in these examples, but the inner one cannot, indicates that the latter formed an inner specifier of the vP when these two phrases underwent successive-cyclic movement. This is precisely what CL predicts: for CL, the relative order of these two moving phrases must be the same both in their final landing sites in the left periphery, and in their intermediate positions in the vP edge. The PG licensing asymmetry in these sentences suggests that this is so: the inner of the two moved phrases in the left periphery behaves like the inner specifier of the vP.[14]

---

[14]Above, we saw evidence from Nissenbaum (2000) that, contra Engdahl (1983), covert *wh*-movement can license PGs in some circumstances: it simply fails to do so when there is only one PG present, because such movement forms an inner specifier. Conversely, in (41-42), we see instances of overt *wh*-movement which fail to license a lone PG when combined with topic/focus movement,

While combining topic/focus movement and *wh*-movement within one clause is not the only way to yield overlapping overt A′-chains in English, it is the simplest one.[15] The alternative is to use a multi-clausal derivation, which contains multiple clause edges that can be targeted for overt A′-movement. As Pesetsky (1982) notes, while such sentences have been taken to be ungrammatical by some works, many speakers do accept them provided that the PCC (which I address later on) is obeyed. Some such sentences are shown in (43-46) below, each of which includes a sentential adjunct with one PG. We see here that, as above, only the outermost of the two overtly A′-moved phrases is able to license the PG:

(43)   a.   *Successful PG licensing by higher moved phrase*
               Tell me [what article]$_2$ she said [who$_1$ I should talk to t_1 about t_2 [before reviewing PG$_2$]]?

        b.   *Failed PG licensing by lower moved phrase*
               * Tell me [what article]$_2$ she said [who$_1$ I should talk to t_1 about t_2 [after introducing myself to PG$_1$]]?

---

despite overt *wh*-movement being able to license PGs in isolation. This result indicates that the overt/covert distinction in A′-movement does not have anything to do with PG licensing: what really matters is the structure involved, and the mechanisms available for interpreting it.

     [15]The examples in (41-42) combining topic/focus movement and *wh*-movement involve the former movement landing in a position above the latter. Watanabe (1986) and Pesetsky (1989) point out that in embedded clauses, this pattern is reversed, with the latter movement obligatorily landing above the former (even if the PCC is respected):

(i)   a.   ? I know who$_2$ [this book]$_1$ we should give notes on t_1 to students of t_2.
        b.   * I know [this professor]$_2$ [which book]$_1$ we should give notes on t_1 to students of t_2.

In my judgment, the PG licensing asymmetry mentioned above holds here as well, as revealed by the fact that only the outer of the two moved phrases can license a lone PG in this context:

(ii)   a.   ? I know who$_2$ [this book]$_1$ we should give notes on t_1 to students of t_2 [if we want to please PG$_2$].
        b.   * I know who$_2$ [this book]$_1$ we should give notes on t_1 to students of t_2 [if we want to get people to cite PG$_1$].

This fact is significant because it indicates that in matrix contexts of this variety like (41-42), in which the topic/focus phrase is outermost, the inability of *wh*-movement to license a PG has nothing to do with the mere fact that topic/focus movement also occurs: when we switch to an embedded context like (i-ii) above where *wh*-movement is instead the outermost of the two, the PG-licensing ability of the two moved phrase reverses. This fact supports a theory under which all forms of A′-movement are equivalent in being able to license PGs in principle, unless the syntactic configuration in which they occur prevents them from doing so, as argued in this chapter.

(44)   a.   *Successful PG licensing by higher moved phrase*
           This is the guy who$_2$ I'm wondering [which painting]$_1$ to send t_1 to t_2 [in order to cheer up PG$_2$].

       b.   *Failed PG licensing by lower moved phrase*
           * This is the guy who$_2$ I'm wondering [which painting]$_1$ to send t_1 to t_2 [in order to get feedback on PG$_1$].

(45)   a.   *Successful PG licensing by higher moved phrase*
           [Which car]$_2$ do you know [who$_1$ to persuade t_1 to buy t_2, [in order to get to borrow PG$_2$ for free]]?

       b.   *Failed PG licensing by lower moved phrase*
           * [Which car]$_2$ do you know [who$_1$ to persuade t_1 to buy t_2 [after having a few drinks with PG$_1$]]?

(46)   a.   *Successful PG licensing by higher moved phrase*
           This is the book [which$_2$ I asked [who$_1$ you gave t_1 comments on t_2 [after making yourself a copy of PG$_2$]].

       b.   *Failed PG licensing by lower moved phrase*
           * This is the book [which$_2$ I asked [who$_1$ you gave t_1 comments on t_2 [after having a discussion with PG$_1$]].

This result indicates that the higher moved phrase on the surface in such sentences was also the higher spec-vP, exactly as CL predicts:[16]

---

[16]The configurations I focus on in this section all involve A′-movement of non-subjects. This is because subjects do not readily license PGs (Culicover and Postal 2001, a.o.). For instance, Engdahl (1983) notes that the following example is not accepted by the majority of speakers:

(i)   * This is the student [∅$_1$ everyone thinks [t_1 is clever] [because John said PG$_1$ was clever]].
       (Adapted from Engdahl 1983, ex. 59)

This example appropriately adjoins the PG-containing adjunct within the path of the A′-moving subject, so it is unclear why this example fails, unless something independent prevents PG licensing by subject movement in most cases. However, Engdahl also observes the following example with PG-licensing by movement of the subject of a copular predication, which is significantly improved:

(ii)   [Which caesar]$_1$ did Brutus imply t_1 was no good [while ostensibly praising PG$_1$]?
       (Adapted from Engdahl 1983, ex. 60)

While I cannot provide concrete evidence for the phasal status of copular clauses of this sort, in my judgment, PG-licensing under multiple movement in such a context reveals the expected contrast:

(iii)   This is the professor who$_2$ I know [which book]$_1$ we should claim t_1 will be helpful for students of t_2 [in order to please PG$_2$] / *[in order to improve sales of PG$_1$].

(47)   *Higher surface position → higher position in vP → can license PG*
        [This book]$_2$, who$_1$ should we [$_{vP}$ $\boxed{t_2}$ t_1 talk to t_1 about t_2...

        ...before commenting on $\boxed{\text{PG}_2}$]]?

The above examples have shown configurations where the two phrases competing to license the PG originate within the same clause. The same asymmetry in PG licensing holds even if the two moved phrases originate in different clauses, however, as exemplified by (48-50) below:

(48)   a.   *Successful PG licensing by higher moved phrase*
              These cookies$_2$, who$_1$ did you tell t_1 [that you adore t_2] [while buying a package of PG$_2$]?

        b.   *Failed PG licensing by lower moved phrase*
              * These cookies$_2$, who$_1$ did you tell t_1 [that you adore t_2] [while having a conversation with PG$_1$]?

(49)   a.   *Successful PG licensing by higher moved phrase*
              These are the beans [which$_2$ I know [who$_1$ you told t_1 [that you revile t_2] [after eating an expired can of PG$_2$]]].

        b.   *Failed PG licensing by lower moved phrase*
              * These are the beans [which$_2$ I know [who$_1$ you told t_1 [that you revile t_2] [after running into PG$_1$ in the hallway]]].

(50)   a.   *Successful PG licensing by higher moved phrase*
              [That book]$_2$, who$_1$ did you tell t_1 [that you hate t_2] [while burning a copy of PG$_2$ in the office yesterday]?

        b.   *Failed PG licensing by lower moved phrase*
              * [That book]$_2$, who$_1$ did you tell t_1 [that you hate t_2] [while discussing chemistry with PG$_1$]?

Given CL, in such examples we predict that the two A′-moving phrases will arrange themselves into an order corresponding to what their final relative order will be, as soon as the initially lower phrase moves into the vP in which the initially higher phrase originates. This is because once the two moving phrases occupy the same

---

Example (ii) is not fully parallel with (i), since it involves a PG in object rather than subject position. This begs the question of whether the configuration in (i) would be improved with an object PG, as in (iii). I suspect that this example is somewhat improved, but perhaps not fully.

(iii)   *? This is the student [∅$_1$ everyone thinks [t_1 is clever] [because John is always talking about PG$_1$]].

phase, they will be linearized with respect to each other, thus at that time and no later it is necessary for them to adopt whatever their final relative order will be. A sample derivation illustrating this prediction is shown in (51) below. The fact that only the higher moved phrase on the surface in such configurations can license a lone PG, as (48-50) showed, indicates that this analysis is correct.

(51)   *Phrases from separate clauses match final order once in same phase*[17]
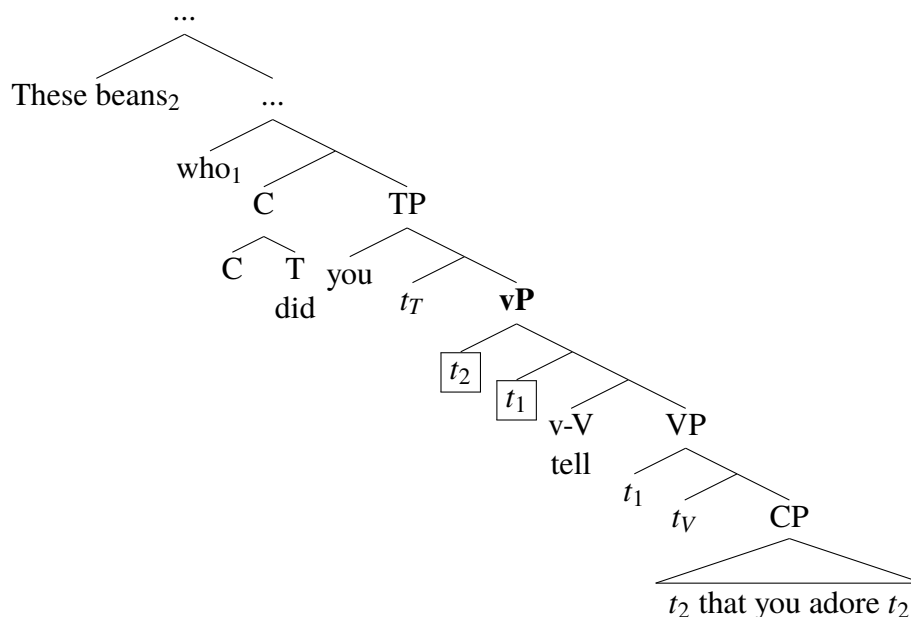
3.1   Confirming the status of the inner moved phrase

In the previous section, we saw that while covert *wh*-movement in English cannot license a PG which is alone (52a), such movement can license the second of two PGs which are contained in one island (52b):

(52)   a.   *Covertly moved wh-phrase cannot license a lone PG*
            * I forget who filed [which articles]$_1$ [without reading PG$_1$]?
            (Engdahl 1983, ex. 34)

        b.   *Covertly moved wh-phrase licenses a second PG*
            [Which senator]$_1$ did you persuade t_1 to borrow [which car]$_2$ [after getting an opponent of PG$_1$ to put a bomb in PG$_2$]?
            (Nissenbaum 2000, p. 12 ex. 8)

---

[17]I represent the indirect object in this sentence as originating in a specifier of the VP, since this accurately captures the word order of this construction, and because more fine-grained approaches to the structure of ditransitives would not contribute anything relevant to the analysis at hand.

As discussed above, Nissenbaum argues that this fact emerges from the semantic effect of multiple specifier formation: when two phrases successive cyclically A′-move from vP, the sister of the (trace of the) inner specifier is interpreted as a two-place predicate. This is the position that a PG-containing adjunct must be merged to in order for the inner specifier to license its PG, but successful semantic composition is only possible here if the adjunct contains two PGs, and is thus also a two-place predicate (35). For this reason, the inner of two A′-moved phrases will only ever reveal its expected PG-licensing ability when the relevant island contains two PGs. In (41-50) above, I provided an array of configurations with overlapping overt A′-movement, in which the moved phrase with the lower surface position is unable to license a lone PG. If as argued above the final relative order of these moved phrases reflects their relative order within vP, then the inner moved phrase should be an inner specifier of vP, which will reveal its PG licensing ability when an adjunct with two PGs is merged. This is indeed the case:

(53)   *Lower of two overtly A′-moved phrases can license a second PG*

   a.   [This book]$_2$, who$_1$ should we talk to t_1 about t_2, before giving comments on PG$_2$ to a student of PG$_1$?

   b.   This is the book which$_2$ I wondered who$_1$ we should send t_1 comments on t_2, after giving a copy of PG$_2$ to a friend of PG$_1$.

   c.   Guess [which cookies]$_2$ I discovered who$_1$ you told t_1 that you are addicted to t_2, [while buying a package of PG$_2$ for a friend of PG$_1$]?

This fact confirms that these constructions really involve a multiple specifier configuration at the vP level, as opposed separate movements through distinct projections within the verbal domain. If the later had occurred, there would not be a two-place predicate in vP that would facilitate a two-PG adjunct.[18]

3.2   Supporting evidence from binding

Under CL, we expect the final relative order for overlapping moved phrases to have already been formed within the first phase that those phrases both occupied, even if many phases intervene between their origination points and final landing sites. This prediction is corroborated by (54) below, in which a PG-containing adjunct includes

---

[18]Following footnote 11 above, we expect the order of the moved phrases and PGs in sentences like (53) to have to match. If their order does not match, the result is indeed degraded:

(i)   ?? This is the book which$_2$ I wondered who$_1$ we should send t_1 comments on t_2, after giving a student of PG$_1$ a copy of PG$_2$.

an anaphor bound by the subject of the embedded clause, ensuring that this adjunct is interpreted in the lower vP—the first phase of this multi-clausal derivation. Notice that the PG here can only be licensed by what is ultimately the phrase with the higher surface position, not the lower one. This shows that the relative c-command relation between these moved phrases that we see in their derived positions was indeed already established within the first phase of the derivation:

(54)   *Ordering established within the first vP as diagnosed by principle A*
        [**What book**]$_2$ did she wonder **who**$_1$ <u>you</u> should [$_{vP}$ t_2 t_1 ask t_1 about t_2 ...

   a.   [✓before making <u>yourself</u> a copy of PG$_2$]]?
   b.   [*after introducing <u>yourself</u> to PG$_1$]]?

Example (55) below uses binding in the opposite way to establish the same result. Here merge of the PG-containing adjunct in the matrix clause is excluded by principle B, forcing it to attach in the embedded clause, and yielding the same asymmetry in PG licensing that we saw in (54).[19]

(55)   *Ordering established within the first vP as diagnosed by principle B*
        [**What book**]$_2$ did <u>you</u> wonder **who**$_1$ we should [$_{vP}$ t_2 t_1 ask t_1 about t_2 ...

   a.   ✓ [before making <u>you</u> a copy of PG$_2$]]?
   b.   * [after introducing <u>you</u> to PG$_1$]]?

3.3   On PGs in subjects

So far this chapter has focused on PGs in sentential adjuncts, which are quite productive, and straightforward for semantic composition. However, the important PG licensing asymmetry shown in this environment also holds for PGs in subjects. This fact was observed by Pesetsky (1982) by way of the following contrast:

(56)   This volvo is one car which$_2$ I know who$_1$ to persuade [[owners of PG$_2$] / [friends of *PG$_1$] to talk to t_1 about t_2].
        (Adapted from Pesetsky 1982, ch. 4, ex. 81-82)

This fact is precisely what we expect, if Nissenbaum (1998, 2000) is correct that the configuration that permits PG licensing within a subject is the same as that

---

[19]Examples (54-55) rely on binding of a non-subject argument in the adjunct by the subject of the containing clause (presumably mediated by a co-indexed PRO in the adjunct), for which it is independently verifiable that principles A and B apply, in clause-bounded fashion as usual:

(i)   a.   You$_1$ went to the park [before taking pictures of yourself$_1$/*you$_1$].
       b.   You$_1$ said that we should go to the park [before taking pictures of you$_1$/*yourself$_1$].

involved with a PG in a sentential adjunct: both involve operator-hosting phrases which merge to vP and take advantage of Predicate Abstraction triggered within it when there is successive-cyclic movement from it.[20]

Notice that in (56), the PG-containing subject linearly precedes the gap left behind by the movement chain that licenses it. This is the opposite of what we see with PGs in sentential adjuncts, which linearly follow the licensing gap, since such adjuncts are linearized to the right of the clause they attach to. Importantly, it is apparently the case that the PG licensing asymmetry that is vital for this chapter applies to both PGs in sentential adjuncts and in subjects, regardless of this difference in how they are ordered with respect to the licensing gap. This is expected if the PG licensing asymmetry stems from structural facts, rather than from something like a linear constraint on the order of gaps and their fillers resulting from limitations of sentence processing. A similar point is made by Pesetsky (1982).[21]

3.4 Interim summary and remaining puzzles

So far, this chapter has used Nissenbaum's predictions about PGs to argue for the following generalization about derivations where multiple A′-chains overlap, which as explained above, is independently predicted by the CL approach:

(57)   *Overlapping Chain Generalization*
         If two A′-moving phrases XP_1 and XP_2 form specifiers of an intermediate phase YP such that XP_1 c-commands XP_2, then XP_1 c-commands XP_2 in all subsequent positions these phrases occupy.

While CL provides a rationale for this generalization, certain details about the derivations that yield it remain unclear. To see why, consider again English sentences combining topic/focus movement and *wh*-movement:

(58)   *Topic/focus movement plus wh-movement in English*
         [This book]$_2$, who$_1$ should we talk to t_1 about t_2?

As discussed in section 2, the theory in Richards (1997, 1999) predicts that the moving phrases here should form crossing paths in vP prior to forming their final landing sites (59a). I have argued that in reality, these moving phrases form nesting

---

[20]Semantic composition is more complex for PGs in subjects, however, but the syntactic configuration is the same as for PGs in sentential adjuncts. See Nissenbaum (1998, 2000) and Nissenbaum and Schwarz (2011) for discussion of the semantic mechanisms required for DP-internal PGs.

[21]Note that I do not argue that linear order constrains PG licensing. I do argue that facts about linear order correlate with PG licensing possibilities, but only because linearization constrains the set of syntactic derivations that PF can accept, and thus indirectly limits what structures will be available for a PG to be interpreted in.

paths within vP, such that they immediately come to match the order that they will have after reaching their final positions (59b). As we've seen, the same conclusion is applicable to numerous other sentences of the same general form.

(59)   a. *Richards: Movement to left periphery after tucking-in at vP*

[*this book*$_2$ *who*$_1$ S T [$_{vP}$ *who*$_1$ *this book*$_2$ t_S v-V t_1 t_2]]

   b. *This chapter: Multiple A′-movement from vP without tucking-in*

[*this book*$_2$ *who*$_1$ S T [$_{vP}$ *this book*$_2$ *who*$_1$ t_S v-V t_1 t_2]]

Earlier in this chapter, we saw evidence from Richards (1997, 1999) that crossed paths / tucking-in must sometimes hold for multiple movements to the same position (see (2) versus (12), for instance). However, this fact begs the question of why crossing paths are not required in vP for (58) and the multitude of sentences like it. If this were the case, such sentences would require the derivation in (59a), which in the context of CL would involve an inevitable crash at PF due to the moving phrases having different relative orders within vP and in their final positions.[22]

   At this point, what is necessary is an understanding of when overlapping paths are permitted to either cross or nest. I consider this issue next, using the PCC as a starting point for building a more general theory of crossing versus nesting.

4 On the PCC

In this section, I will provide a CL-based account of the PCC, as a prelude to a more formalized understanding of the distribution of crossing and nesting paths. It will be useful to first discuss a rationale for the PCC, since examining the derivations in which it arises will necessarily bring into focus several important issues about when crossing and nesting occur. The primary goal of this section is to provide a linearization reason for why nesting paths will sometimes be required in English, while the following section will consider in detail the general syntactic mechanisms that allow crossing or nesting to occur in a given context.

   The PCC was proposed by Pesetsky (1982) on the basis of the fact that in a variety of circumstances involving overlapping A′-movement, it is necessary for the

---

[22]Also note that given the ban on phrase-bound specifier-to-specifier movement argued for in previous chapters, it is not possible for the derivation of (58) to have involved tucking-in as in (59a) followed by re-arrangement in the vP edge in order to form the CL-obeying order in (59b).

paths to nest, resulting in the moved phrases ending up in a final order that is the reverse of the order they had in their origination positions:

(60)   *Forced order reversal as described by PCC*
     a.   [This book]$_2$, who$_1$ should we talk to t_1 about t_2?   [2 1 ... 1 2]
     b.   * [This student]$_1$, what$_2$ should we talk to t_1 about t_2?   [1 2 ... 1 2]

As we've seen, in configurations like (60), it is necessary for topic/focus movement to form a final landing site above *wh*-movement. Importantly, the PCC also holds when there is no rigidity of respective landing sites. For instance, the PCC also applies when one embedded *wh*-question is contained within another (61), despite the fact that the two instances of *wh*-movement in such a context are presumably driven by the same type of syntactic feature, and thus would in principle be suitable candidates to land in the periphery of either embedded interrogative CP:[23]

(61)   *PCC in double embedded question*
     a.   Tell me [$_{CP}$ [what subject]$_1$ you know [$_{CP}$ who$_2$ to talk to t_2 about t_1]].
     b.   * Tell me [$_{CP}$ who$_2$ you know [$_{CP}$ [what subject]$_1$ to talk to t_2 about t_1]].

Indeed, Pesetsky shows that the PCC applies very generally to contexts with overlapping A′-movement in English, regardless of the sub-types of A′-movement involved.

However, if Richards (1997) and Nissenbaum (2000) are correct about the derivation of multiple *wh*-questions in English, then such configurations violate the PCC since they involve crossing rather than nested dependencies, due to tucking-in:

(62)   *No PCC in a multiple wh-question*
     Who$_1$ ∅$_2$ did Ivan ask t_1 **what**$_2$?

What distinguishes English multiple *wh*-questions from PCC-obeying contexts is that in the former, only one phrase moves overtly, while in the latter, all moving

---

[23]Interestingly, while D-linking can create superiority-violating *wh*-questions (ia), D-linking is not capable of creating PCC violations (ib versus ic):

(i)   a.   Tell me [which subject]$_1$ you talked to [which teacher]$_2$ about t_1.
     b.   * Tell me [$_{CP}$ [which teacher]$_2$ you know [$_{CP}$ [which subject]$_1$ to talk to t_2 about t_1]].
     c.   ✓ Tell me [$_{CP}$ [which subject]$_1$ you know [$_{CP}$ [which teacher]$_2$ to talk to t_2 about t_1]].

While D-linking is beyond the scope of this chapter, this fact suggests that superiority violations permitted by D-linking are not derived by re-arranging the relevant *wh*-phrases before their movement to the clause periphery begins. If this were the case, then it should be possible to form (ib) by first reversing the order of *wh*-phrases before moving those phrases onward in PCC-obeying fashion as usual (though not in a surface-evident way).

phrases at least can be overt (modulo effects like the Doubly Filled Comp Filter). Richards (1997) takes advantage of this fact, hypothesizing that the absence of the PCC in multiple *wh*-questions, and the arising of the PCC in situations with multiple overt A′-movements, is connected to the impossibility of CP hosting multiple overt specifiers in English and languages like it.[24]
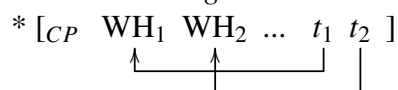
For Richards, nothing prevents a second movement to the same head in English if covert (63a), but since a second overt movement to a specifier of the same CP is by hypothesis unavailable (63b), the only possibility for a situation with overlapping overt *wh*-movements like (61) above is the following: the first C merged will trigger overt movement of the closer *wh*-phrase, with the C merged later on then attracting the lower *wh*-phrase (63c). Richards follows Kitahara (1991, 1997) in hypothesizing that this derivation is responsible for the PCC, and posits that the somewhat degraded status of the relevant sentences stems from the fact that they necessarily involve locality-violating movement of the lower *wh*-phrase:

(63)   *Distribution of overt versus covert movement in English (Richards)*

    a.   *Covert tucking-in*
$$[_{CP} \text{ WH}_1 \ \varnothing_2 \ ... \ t_1 \ \text{WH}_2 \ ]$$
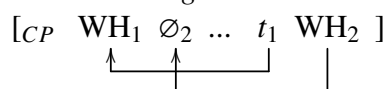
    b.   *No overt tucking-in*
$$* [_{CP} \text{ WH}_1 \ \text{WH}_2 \ ... \ t_1 \ t_2 \ ]$$

    c.   *Long distance nesting movement required for multiple overt extraction*
$$[_{CP2} \text{ WH}_2 \ \text{C2} \ ... \ [_{CP1} \text{ WH}_1 \ \text{C1} \ ... \ t_1 \ t_2 \ ]]$$
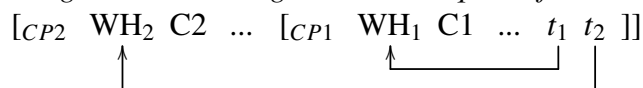
In contrast, Richards argues that the possibility of overt tucking-in within CP in Bulgarian, as revealed by its multiple *wh*-questions (2a), correctly predicts the absence of the PCC in this language (64). In brief, Richards' analysis is that since the first C merged can overtly attract both *wh*-phrases in Bulgarian, the lower *wh*-phrase

---

[24]An explanation for the PCC that depends on the phonological status of the movements involved, which I will develop later in this section, might lead us to expect that the PCC will not apply when one of the moved phrases involved is covert. This is not the case, however, as we see below:

(i)   a.   This is the saxophone $\varnothing_2$ I know [precisely which songs]$_1$ to never play t_1 on t_2.
    b.   * This is the saxophone $\varnothing_2$ I know [precisely which songs]$_1$ to never use t_2 to play t_1.

The occasional linearization-sensitivity of silent elements will be discussed comprehensively in chapter 8. For the meantime, following a suggestion I made in footnote 15 of chapter 3 with respect to a similar issue, it would suffice to say that the relevant element here was originally overt, and thus visible for linearization, before the application of a process like Free Deletion in Comp.

need not remain in situ until the merge of the second C. Rather, the second C will simply target the closest of the two phrases moved to the intermediate clause edge, and a PCC-style derivation never occurs:

(64)   *Anti-PCC in Bulgarian: Moved wh-phrases do not reverse their order*
   a.   **Koj**$_1$ se   opitvat da razberat [$_{CP}$ t_1 **kogo**$_2$ t_1 e   ubil t_2]]?
         who self try   to find.out   whom   AUX killed
   b.   * **Kogo**$_2$ se   opitvat da razberat [$_{CP}$ **koj**$_1$ t_1 e   ubil t_2]]?
          whom self try   to find.out   who   AUX killed
         (Richards 1997, p. 41, ex. 41)

   This account for the difference between English and Bulgarian is not compatible with contemporary developments in phase theory, nor the results of this dissertation, both of which mandate against a theory in which the initially lower A′-moving phrase is non-successive-cyclically attracted from its origination position in PCC derivations (63c). Next, I argue for an alternative account of the PCC that is compatible with the CL theory of phases, as well as the results argued for earlier in this chapter about the derivational history of such sentences. I will first provide a general rationale for the PCC, building on Richards' proposal that the covert/overt distinction for multiple specifiers is the central factor, and then propose a more formalized account of the distribution of crossing and nesting paths.

4.1   A linearization understanding of the PCC

I argue that if English generally regards tucked-in specifiers of A′-movement as covert, as discussed extensively above in reference to multiple *wh*-questions, then we come to a natural understanding of why the PCC must apply in contexts with overlapping overt A′-movement in this language. This analysis will require maintaining a distinction between inner specifiers derived in different ways, however.

   I hypothesize that when two phrases move to one head with crossing paths, spell-out of the containing phase linearizes the outer moved phrase at only the head of its chain, while the phrase that forms the inner specifier is linearized in only the tail of its chain, as in (65a). Further, I hypothesize that when two phrases move to one head with nesting paths, spell-out of the containing phase linearizes both phrases based on the head of their chains, as in (65b), in which α moves overtly

despite occupying an inner specifier after movement of β above it:[25] [26]

(65)   *Multiple specifier linearization in English*

   a.   *Crossing paths $\rightarrow$ covert inner specifier*

   XP
   α_1
   $\varnothing_2$
   X   ...
   ... $\varnothing_1$ β_2
   Linearization: $\alpha < X < \beta$
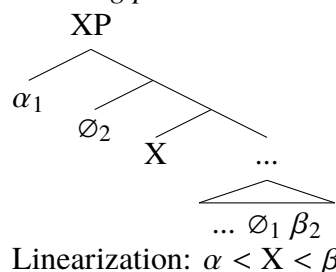
   b.   *Nesting paths $\rightarrow$ only overt movements*

   XP
   β_2
   α_1
   X   ...
   ... $\varnothing_1$ $\varnothing_2$
   Linearization: $\beta < \alpha < X$
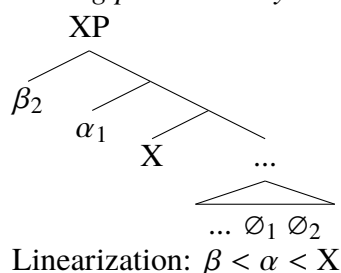
This is in contrast to a theory in which all inner specifiers, regardless of how they are formed, are treated as derived by covert movement. For CL, a theory along such lines would incorrectly predict the impossibility of ever overtly moving more than one phrase from a phase in languages like English, since a phase would only ever allow one phrase to be linearized at its edge. While I have not yet explained what allows tucking-in to be avoided in order to derive a double overt movement configuration like (65b), next I will show that positing the possibility of such a

---

[25]The schematic trees in (65) can be understood as illustrations of the structural descriptions under which each rule of chain linearization respectively applies. The illustration in (65a) is analogous to what is proposed in Pesetsky (2000), who along with Bobaljik (1995); Brody (1995); Groat and O'Neil (1996) and others, I follow in assuming that the overtness or covertness of a given movement step is determined at spell-out. This is in contrast to the (inverted) Y-model (Chomsky and Lasnik 1977) under which covert movement occurs post-spell-out.

[26]This conception of the rules that determine whether an instance of movement is overt or covert does not apply to circumstances like (i), in which two *wh*-phrases originate in separate clauses, and yet the lower *wh*-phrase undergoes no overt movement:

(i)   **Who** do you think told the kids to buy **what**?

Examples of this sort are especially challenging for a bottom-up theory of cyclic spell-out like the one pursued here. In the appendix to this chapter I further discuss the reasons why such sentences are puzzling, but ultimately I must leave them aside for future work.

movement pattern yields the correct linearization for PCC sentences in English. I will go on to consider the syntactic forces that determine the distribution of crossing and nesting in greater detail in the next section. For the purposes of exposition, for the meantime assume that crossing and nesting multiple specifier formation are both always available as far as the syntax is concerned.

To see how the hypothesis in (65) facilitates an understanding of the PCC in English and languages like it, first consider the schema for a vP in (66) below, where the phrases α and β both undergo A′-movement to the vP edge. Here β tucks-in below α. If tucking-in movement is regarded by English PF as covert (65a), then β will be linearized at the tail of its movement chain, at the right edge of the vP:

(66)   *Covert tucking-in within vP*
   $[_{vP} \quad \alpha \ \varnothing_{\beta} \ \text{S} \ \text{v-V} \ [_{VP} \ t_{\alpha} \ \beta \ ]]$

   Linearization: $\alpha < \text{S} < \text{V} < \beta$

If later in the derivation α moves to an overt landing site external to the vP, as in (67) below, there will be no issue: since α occupied the left linear edge of the vP in (66), CL allows α to overtly move on leftward in this way. However, the same does not apply to β. Since β did not overtly move to the left edge of vP in (66), it cannot overtly exit the vP in this way without causing a linearization contradiction, as (67) shows. This is because β cannot both follow the content of vP as (66) established, and also precede it as the later movement in (67) requires:

(67)   *Later overt movement from vP*
   $[\alpha \ ... \ \beta \ ... \ [_{TP} \ \text{S} \ [_{vP} \ ... \ \text{v-V} \ ... \ ]]]$
   Linearization: $\alpha < \beta < \text{S} < \text{V}$
   (Deviant if following (66) due to contradicting previous order $\text{S} < \text{V} < \beta$)

This contradiction would be avoided if in (66), β had moved above α within vP rather than tucking-in beneath it, as in (68) below. Following the linearization rule in (65b) above, neither movement is treated as covert in this situation:

(68)   *Nesting paths in vP yields two overt movements to spec-vP*
   $[_{vP} \ \beta \ \alpha \ \text{S} \ \text{v-V} \ [_{VP} \ t_{\alpha} \ \beta \ ]]$

   Linearization: $\beta < \alpha < \text{S} < \text{V}$

After (68), further overt movement of β and α is licit, since they both occupy the linear edge of the vP. Thus only a nesting derivation like (68), rather than one with

crossing paths like (66), facilitates multiple overt A′-movements from the vP.[27]

Importantly, notice that by forming nesting paths in (68), these moving phrases end up reversing their order within vP. Provided that their subsequent landing sites preserve their reversed ordering established within vP, the result is a CL-satisfying derivation fitting what the PCC describes as legal, in which the origination positions and final positions of the moved phrases are reversed with respect to each other:[28]

(69)   *PCC obeyed: Final positions reverse origination positions*
         $\checkmark$ [XP_2 ... YP_1 ... [$_{vP}$ t_2 t_1 ... t_1 t_2]]

This is precisely the derivation that the PG facts examined earlier in this chapter have shown us occurs in reality: overlapping moved phrases take on an order that matches their final relative order within the first vP they occupy, which in the case of a PCC-obeying sentence, involves immediate order-reversal followed by order-preserving movement onward.

4.2   Correct predictions and further puzzles

The above account of the PCC in English has, for the purposes of exposition, assumed that crossing and nesting multiple specifier formation are both possible in principle (modulo concerns of linearization). Note that this understanding accurately predicts the possibility of a crossing-paths derivation in English multiple *wh*-questions: since the non-initial *wh*-movement in such contexts has a covert landing site, no linearization problem arises from a lower *wh*-phrase tucking-in within vP before doing the same later on in CP. The above account of the PCC also accurately predicts that Bulgarian will not be subject to it, as we saw in (64). Since this language allows multiple overt specifiers in one phrase, it can simply allow overlapping overt A′-

---

[27]Assuming that the linearization rules described by (65) apply at spell-out, it is irrelevant whether the configuration in (68) was formed by attraction of α before β as described here (though superiority would lead us to expect this), or whether these phrases are attracted in the reverse order. Since for this view linearization is determined by the final structure of the paths those movements form, the order of operations does not matter. See footnote 32 below for a potentially desirable consequence of allowing attraction in either order in certain contexts.

[28]If the problem with (67) as a continuation of (66) is due to β moving covertly in vP and then onward overtly, then we might expect the relevant structures to simply involve covert movement of the initially lower moved phrase, rather than overt movement to a higher position in the way that we see in reality in PCC sentences. Note that in all PCC sentences in English, the two moved phrases end up in distinct final landing sites (assuming as footnote 13 states that topic/focus movement and *wh*-movement target distinct projections, following Rizzi). Given this fact, we will rule out the possibility of covert movement of the initially lower phrase under the assumption that a given landing site must have at least one overt goal, at least in the English contexts under consideration here. See Richards (2010, 2016) for work in this vein.

movements from vP to overtly tuck-in within this phase before then moving on. The problem posed by the covertness of tucked-in specifiers in English thus does not arise in Bulgarian, which therefore need not opt for a nesting path derivation.

However, if crossing and nesting multiple specifier formation are in free variation, then we have no explanation for why in English and Bulgarian multiple *wh*-questions, crossing paths are not merely an option, but a requirement.[29] The same holds for numerous obligatory tucking-in scenarios shown by Richards (1997) to never permit order reversal, including but not limited to local A-scrambling in Japanese, negative fronting in Bulgarian, multiple object shift in Germanic languages, and multiple clitic clustering in Tagalog and Serbo-Croatian.[30] In order to solve this puzzle, it is necessary to develop a more precise account of when crossing and nesting paths are permitted.[31] I do this in the next section, where I hypothesize that crossing paths are required in certain syntactic contexts, but in free variation with nesting in others (to the extent that concerns of linearization do not interfere).

5   Formalizing the distribution of crossing and nesting

Several principles are required to accurately formalize the distribution of crossing and nesting paths. First, let's consider the conditions under which crossing applies. Recall that Richards (1997, 1999) argues that crossing paths occur when one head attracts multiple phrases. Earlier in this chapter, we saw evidence from PGs that in English sentences that adhere to the PCC, the order that the moved phrases take on in their final landing sites is mimicked at their intermediate landing sites, in a way that entails a nested path structure within vP and thus an absence of tucking-in at that level of the derivation. This result leads us to the conclusion that tucking-in is not a universal property of situations where multiple phrases move to one head.

In response to finding, I propose that tucking-in is not a requirement of multiple movements to one head, but rather applies to movements that are triggered by one feature, as diagrammed in (70) below. Here the unvalued feature [uF1] on the head

---

[29]We have seen this fact for Bulgarian in (2a/12a) above. The following example shows the impossibility of a string that corresponds to nesting paths in an English multiple *wh*-question:

(i)   * What$_2$ $\varnothing_1$ did Ivan ask who$_1$ about t_2?

[30]Bruening (2001) argues based on facts about scope in ditransitives that multiple instances of Quantifier Raising also must involve crossing paths.

[31]Similarly, what I have just said accurately rules in the anti-PCC configuration in the Bulgarian (64a) above, but also rules in a PCC-style derivation for this Bulgarian configuration, which as we see in (64b), is unacceptable. I will show later in this chapter that the expected freedom of ordering arises in other contexts in Bulgarian (as well as Romanian, another multiple *wh*-fronting language).

X attracts, with crossing paths, the goals α and β which both bear the feature [F1]:

(70)   *Proposal #1: One feature attracts multiple goals → tucking-in*
         $[_{XP}\ \ \alpha_{[F1]}\ \ \beta_{[F1]}\ \ X_{[uF1]}\ \ ...\ \ t_\alpha\ \ t_\beta\ ]$
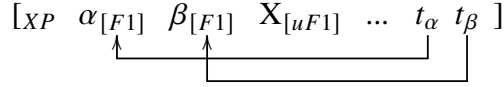
Additionally, I propose that when a head bears multiple movement-triggering features, the goals respectively attracted by those features are free to cross or nest. I hypothesize that if the requirement to nest is limited to multiple movements triggered by one feature, then when multiple goals each move due to attraction by separate features, the grammar does not constrain the form of their paths.[32] This concept is illustrated in (71), where the head X bearing the features [uF1] and [uF2] can attract the goals $\alpha_{[F1]}$ and $\beta_{[F2]}$ in a way yielding either crossing (71a) or nesting (71b) paths to its edge. The freedom of ordering in such contexts will be vital to capturing the full range of multiple movement configurations.

(71)   *Proposal #2: Attraction by separate features on one head → variability*

         a. $[_{XP}\ \ \alpha_{[F1]}\ \beta_{[F2]}\ \ X_{[uF1,uF2]}\ \ ...\ \ t_\alpha\ \ t_\beta\ ]$

         b. $[_{XP}\ \ \beta_{[F2]}\ \ \alpha_{[F1]}\ \ X_{[uF1,uF2]}\ \ ...\ \ t_\alpha\ \ t_\beta\ ]$

This feature-centric proposal about the distribution of crossing and nesting can only be meaningful when combined with an explicit proposal about how movement-triggering features are distributed. Following a great deal of work in contemporary syntax, I assume that whatever features are responsible for ultimately attracting a given phrase to its final position are also instantiated on the heads of intermediate phases crossed by that phrase's movement. For instance, in a *wh*-question, both C and v must bear a [uWH] feature capable of attracting the *wh*-phrase, while in a sentence combining topic/focus movement and *wh*-movement, v must bear two features—[uWH] and [uTop], corresponding to the features th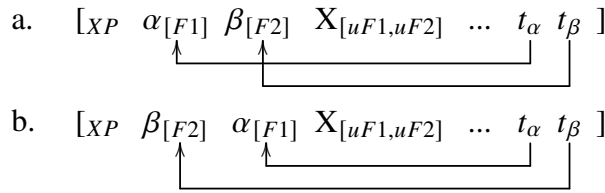at presumably pull the moving phrases in such a context to their respective positions in the left periphery. This is a straightforward consequence of Probe-Goal theories of syntactic dependencies under which movement is feature driven. However, it is worth asking what mechanisms are responsible for enforcing the necessary distribution of features.

---

[32]Alternatively, we might hypothesize that the freedom of specifier order in this situation stems from the possibility of the attracting features targeting their respective goals in either order. Since the attracting features in this case target goals with distinct features, it is plausible to assume that α in (71) will not serve as an intervener for the purposes of the attraction of β by [uF2], if that feature happened to trigger movement before [uF1]. A similar possibility is raised by Doggett (2004).

In this chapter, I will argue that the relevant facts about crossing and nesting emerge from the following two principles:[33]

(72)   *Principles of feature distribution*

      a.   *Attractor minimization principle*
          A given head enters the derivation with at most one movement-triggering feature, unless additional features are motivated by the principle in (72b).

      b.   *Intermediate attractor principle*
          If a given head X bears a movement-triggering feature [uF], and a phase head Y structurally intervenes between X and the corresponding goal phrase bearing [F], then Y may also bear [uF].

The principle in (72a), which is essentially an economy condition, prevents the addition of superfluous features in a way that we are about to see is necessary for the account of contexts with obligatory tucking-in. The additional principle in (72b) allows a head to bear additional movement-triggering features to the extent that they are necessary to facilitate successive cyclic movement.[34] Next, I will argue that the interaction of these two principles about feature distribution with the proposals about multiple specifier formation provided in (70-71) above accurately predicts the attested distribution of crossing and nesting paths.

---

[33]The principles in (72) are consistent with an understanding in which feature distribution is in principle free, but constrained by an economy condition motivating using only the minimal amount of features necessary for a convergent derivation in a given context. This concept is similar in character to the following constraint from previous literature:

(i)   *Economy of Movement Metric*
      If a derivation D_1 of a spell-out domain α converges without some movement operation, then D_1 blocks a derivation D_2 of α that includes that movement operation.
      (Chomsky 1992; Reinhart 2006; Overfelt 2015)

In the context of a theory in which movement is feature-driven, a principle of this variety is reducible to an economy constraint not on movement operations themselves, but on the introduction of features responsible for triggering movement. Such a concept will appropriately rule in only derivations with the fewest number of movement-triggering features needed for a convergent result.

[34]While a principle like (72b) is implicit in much syntactic work, ideally, it will stem from something independent. A possible explanation may be found in van Urk and Richards (2015), who argue that a prerequisite for extraction from a phase is that the phase is first targeted for agreement (see further the discussion in chapter 8). I hypothesize that this is so because each phase head must agree with the one below it in order to establish that the lower one carries matching movement-triggering features. For instance, in a *wh*-question, v must be generated bearing a [uWH] feature which, when interrogative C is merged, it matches against its own [uWH] under agreement with v. This process must recursively apply if the derivation includes more intermediate phases.

5.1 Deriving the distribution of crossing and nesting

Consider a multiple *wh*-question, which as we've seen, is characterized by multiple *wh*-phrases moving to the same CP with obligatory crossing paths:

(73)   a.   *English multiple wh-question*
           Who$_1$ $\varnothing_2$ did Ivan ask t_1 **what**$_2$?

        b.   *Bulgarian multiple wh-question*

           Kogo$_1$ **kakvo**$_2$ e   pital   Ivan t_1 t_2?
           who.NOM what.ACC AUX asked Ivan

           ([Richards 1997](#) p. 277 ex. 75a)

Given the principle in (72a), C in such a configuration must have a single [uWH] feature capable of attracting both *wh*-phrases. Given the principle in (72b), the v which intervenes between C and these *wh*-phrases also has [uWH]. First, [uWH] on v attracts these phrases, which being moved by the same feature form a crossing path configuration in vP, as (70) dictates. Precisely the same thing happens when these phrases are then targeted by the [uWH] feature on C, as in (74):

(74)   *Pure crossing paths in a multiple wh-question*

$$[_{CP} \ \text{WH1} \ \text{WH2} \ C_{[uWH]} \ \ ... \ \ [_{vP} \ \text{WH1} \ \text{WH2} \ v_{[uWH]} \ \ ... \ \ t \ t \ ]$$

If in this context it were possible for C to have two distinct [uWH] features (presumably instantiated on v as well) which separately attract the two *wh*-phrases, we would expect these phrases to be able to form a nesting path configuration within the vP, following the principle in (71). This is illustrated in (75), where these phrases nest within vP and then move on to CP in order-preserving fashion, both of which movements (71) permits in this situation:

(75)   *Hypothetical nesting multiple wh-question*

$$[_{CP} \ \text{WH2} \ \text{WH1} \ C_{[uWH1,uWH2]} \ \ ... \ \ [_{vP} \ \text{WH2} \ \text{WH1} \ v_{[uWH1,uWH2]} \ \ ... \ \ t \ t \ ]$$

In reality, such a result is not possible for (non-D-linking) multiple *wh*-questions, as shown again for concreteness in (76) below. The hypothetical derivation in (75) is appropriately excluded if the principle in (72a) blocks the introduction of a C which superfluously bears two distinct tokens of [uWH], rather than just one, since one feature is all that is needed to attract both of the *wh*-phrases in this context.

(76)   *Unacceptable nesting multiple wh-questions*

         a.   *Bulgarian*

                * Kakvo$_2$ kogo$_1$ e   pital   Ivan t_1 t_2?
                what.ACC who.NOM AUX asked Ivan

                (Richards 1997, p. 277 ex. 75b)

         b.   *English*
                * What$_2$ $\varnothing_1$ did Ivan ask who$_1$ t_2?

I hypothesize that all configurations with multiple movements to the same position that obligatorily form crossing paths involve those movements being driven by a single attracting feature, as in (74). The feature involved in triggering those movements will, of course, vary in accordance with the type of movement occurring.

Next, lets examine a context in which, in contrast to multiple *wh*-questions, the PCC applies, forcing nesting. First I'll consider a sentence with topic/focus movement and *wh*-movement, as discussed extensively above. Since the landing sites for these movements are rigidly ordered in the left periphery, I will assume that these movements target dedicated projections following Rizzi (1997) and many others, labeled XP and YP for convenience in (77) below. The respective movement triggering features on the heads of these projections are labeled [uTop] and [uWH]:

(77)   *PCC-obeying topic/focus plus wh-movement*
         [$_{XP}$ [These papers]$_2$ X$_{[uTop]}$ [$_{YP}$ who$_1$ Y$_{[uWH]}$ did you ask t_1 to grade t_2]]]?

In order to trigger successive-cyclic movement of these phrases, v must bear attracting features corresponding to both of these landing sites in the left periphery, given the principle in (72b) above. This is shown in (78):

(78)   *Pre-movement structure of the vP in (77)*
         [$_{vP}$ v$_{[uWH,uTop]}$ ... who [these papers]]

Since these phrases are attracted by distinct features, then given the proposal in (71) above, we predict that these phrases can in principle either form crossing or nesting paths within this vP. If these phrases move to the vP edge with crossing paths, the phrase that forms the inner spec-vP will move covertly, for the reasons laid out in section 4.1 above. As we saw in that section, since in such a context the inner moved phrase will move covertly within vP before overtly moving on from it, such a derivation will result in a linearization contradiction. In section 4.1 we also saw that, in contrast, no linearization issue arises if these phrases move in nesting fashion in vP. In this case both phrases move overtly, while also reversing their relative order in the way the PCC describes. After this initial reversal, these phrases can move

on to their final landing sites in order preserving fashion, as in (79). As we saw in section 4, this derivation accurately produces the PCC-obeying order characteristic of these and similar sentences.

(79)   *Nesting paths in vP permitted by distinct probes on v*

$[_{XP}$ [These papers]$_2$ X$_{[uTop]}$ $[_{YP}$ who$_1$ Y$_{[uWH]}$ S $[_{vP}$ t_2 t_1 v$_{[uTop,uWH]}$ V t_1 t_2]]]
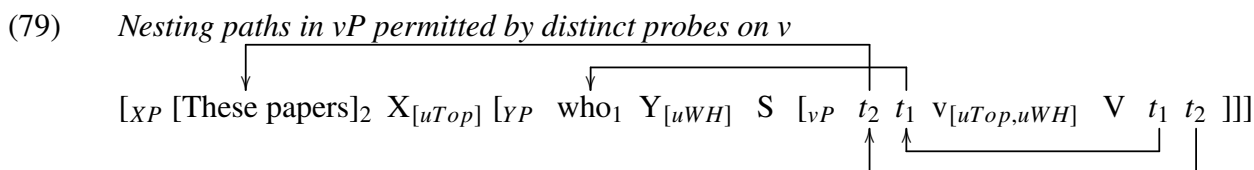
Further, as we saw in section 2, the PG diagnostic provides independent evidence that this derivation, with nesting paths in the initial vP, is the correct one.

This analysis easily extends to the majority of other contexts with two phrases undergoing separate instances of overt overlapping A′-movement in English, all of which obey the PCC (Pesetsky 1982). For instance, consider a sentence combining relativization with interrogative *wh*-movement (80):

(80)   *Relativization plus interrogative wh-movement: PCC obeyed*
      a.   This is the book [which$_2$ I asked [who$_1$ you gave t_1 comments on t_2].
      b.   * This is the book [which$_1$ I asked [who$_2$ you sent a review of t_1 to a student of t_2].

We saw in (44) above that this configuration, like others of the same variety, displays the PG-licensing asymmetry which indicates that the final order of moved phrases here was established in the initial vP where those phrases originated. As in the previous scenario in (79), the result that these moved phrases reverse order upon their movement within the initial vP is predicted, given (71), if these phrases' movements are driven by separate features. We could call these features [uREL] and [uWH] in (80), but the way these features are named is not important: however we label them, the features involved here are plausibly of distinct classes. This leads us to expect the possibility of an order-reversing, non-tucking-in derivation in (80)—a possibility that must be selected here, yielding the PCC, since a tucking-in derivation would pose a linearization problem as discussed in section 4.1.

The main circumstance where the applicability of this account is less obvious is with recursive embedded *wh*-questions. As shown in (61) above and repeated again once more below, the PCC holds for such sentences, despite the fact that the movements involved here are presumably featurally identical:

(81)   *PCC in double embedded question*
      a.   Tell me $[_{CP}$ [what subject]$_1$ you know $[_{CP}$ who$_2$ to talk to t_2 about t_1]].
      b.   * Tell me $[_{CP}$ who$_2$ you know $[_{CP}$ [what subject]$_1$ to talk to t_2 about t_1]].

If both of the *wh*-phrases in such a context are attracted by a single *wh*-feature, we would expect such a sentence to have a form like that of a multiple *wh*-question—that is, with overt movement of the initially highest *wh*-phrase, and covert tucking-in movement of the lower one. This result is impossible in this context, however:

(82)   *Double embedded question cannot behave like a multiple wh-question*
        * Tell me [$_{CP}$ who$_1$ $\varnothing_2$ you know [$_{CP}$ t_1 t_2 to talk to t_1 about [what subject]$_2$]].

The actual result in (81a) above will accurately be predicted by positing that the two CP edges that serve as landing sites for *wh*-movement in this context bear distinct *wh*-sensitive features. This could be formalized by counter-indexing those features, for instance as [uWH1] and [uWH2]. If in (81a) the outermost embedded C bears [uWH1] and the inner embedded C bears [uWH2], both of which are also instantiated on the intermediate v that they both c-command, then the two *wh*-phrases here can be attracted separately within the vP and thus form a nesting configuration before moving onward, creating the PCC-obeying result we see in reality in (81a). As discussed extensively above, nesting rather than crossing (and thus a derivation culminating in order reversal) is required in the initial vP here, given that in English crossing paths within the vP would result in an ordering contradiction due to the covertness of the inner specifier.

5.1.1   A further constraint on overtness/covertness

PCC-obeying overlapping movement raises a problem for the above analysis when we consider multi-clausal sentences more carefully. This problem arises in examples like (80-81) above, but the same point can be made more clearly by examining (83) below. The analysis of the PCC that I have proposed entails order-reversal of moving phrases within the first vP where they are both present, followed by order-preserving movement of those phrases to their final landing sites. The traces placed in (83) are faithful to this proposal, showing the initial order reversal within vP, followed by preservation of that reversed order in the intermediate CP and matrix vP that these phrases pass through on the way to their final positions in the matrix clause:

(83)   *Initial reversal followed by order preservation in all intermediate positions*
        [This book]$_2$, who$_1$ did you [$_{vP}$ t_2 t_1 say [$_{CP}$ t_2 t_1 that we should [$_{vP}$ t_2 $t_1$talk to t_1 about t_2]]]?

For these moving phrases to preserve the order reversal established in the initial vP, their subsequent movement steps must form crossing paths. This requires the lower moving phrase *who*$_1$ to tuck-in below *this book*$_2$ in the intermediate CP and matrix

vP edges, as the traces in (83) show. The problem is that if in English tucking-in is derived by covert movement as discussed above, it is unclear how movement of the lower phrase *who* in (83) manages to be ultimately overt in this configuration, as opposed to becoming covert after tucking-in in the embedded CP edge.[35]

In order to avoid positing that the moved phrase *who*$_1$ in (83) somehow covertly moves through lower specifiers of the intermediate CP and matrix vP before overtly moving to its final landing site (a derivation which would incur a linearization contradiction), I propose the following condition:

(84)   *Chain Consistency Condition*[36]
         The overtness or covertness of the initial step of a movement chain driven by a given feature [uF] must be matched by all later movement steps driven by [uF] until that sequence of movements terminates, regardless of how those subsequent movement steps would otherwise have been linearized.

This principle allows tucking-in of *who*$_1$ in the edges of the embedded CP and matrix vP in (83) to be overt, contrary to usual circumstances, since those movement steps were preceded by overt non-tucking-in movement of *who*$_1$ within the initial vP.[37] Thus the moved phrases in (83) can overtly move on from the initial vP in order-preserving fashion as required.

5.2   A return to Bulgarian and extension to Romanian

Earlier in this section, we saw an example from Richards (1997) showing that Bulgarian lacks the PCC, repeated below in (85). In this example, an embedded question within a matrix one results in two *wh*-phrases originating in the same clause landing in separate CP edges, without reversing their order. Order reversal is not only unnecessary, but impossible here:

---

[35]Note that it cannot be the case that order-reversal in fact applies at every phase edge in (83). If this were so, then these phrases would end up in the wrong final order, as (i) below shows:

(i)   * Who$_1$ [this book]$_2$, did you [$_{vP}$ t_2 t_1 say [$_{CP}$ t_1 t_2 that we should [$_{vP}$ t_2 $t_1$talk to t_1 about t_2]]]?

This result is unacceptable for multiple reasons, but the important point here is that a successive flip-flopping derivation cannot be attributed to (83), even if it weren't the case that CL rules it out.

[36]Fox and Nissenbaum (2018) precede me in positing a condition of this sort ("Early Determination") for covert movement. The formulation I offer here applies to overt movement as well.

[37]This principle also permits an element to undergo covert movement after landing in the terminal position of a given overt movement chain, though initiating overt movement after completing a covert movement chain will generally be ruled out by linearization for the reasons discussed in section 4.

(85)   *Anti-PCC in Bulgarian: Overlapping wh-phrases do not reverse order*

   a.   *No reversal*

Koj$_1$ se   opitvat da razberat [$_{CP}$ **kogo**$_2$ t_1 e   ubil t_2]]?
who self try   to find.out   whom   AUX killed

   b.   *Unacceptable reversal*

* **Kogo**$_2$ se   opitvat da razberat [$_{CP}$ **koj**$_1$ t_1 e   ubil t_2]]?
  whom self try   to find.out   who   AUX killed

(Richards 1997, p. 41, ex. 41)

I have proposed in reference to the English (81) above that in such situations where multiple interrogative *wh*-phrases end up in distinct landing sites, they are attracted by distinct features, such that either a crossing or nesting configuration should be available for them, given the proposal in (71). We saw in (81) above that a nesting derivation is in fact required in this English context (and many comparable ones). Following the concerns of section 4.1 above, this is because a crossing paths derivation for such a sentence would cause a linearization problem in English, due to the covertness of tucked-in specifiers in this language. However, since Bulgarian allows overt tucking-in, what I have said in section 4.1 wrongly predicts that this language should allow not only anti-PCC configurations like (85a), but also PCC-resembling ones like (85b).

In (70) above, I proposed that crossing paths are forced when a single feature on a given head attracts multiple goals. Given this proposal, the impossibility of order reversal in (85) would be expected if the two *wh*-phrases here are in fact attracted by the same feature, unlike in a comparable English recursive interrogative configuration. This would be the case if in Bulgarian, unlike English, any feature sensitive to interrogative *wh*-phrases necessarily interacts with, and thus attracts, all such phrases in its c-command domain.[38] This would preclude the possibility of a nesting path derivation in (85).[39]

If this hypothesis is correct for the multiple interrogative context in (85), then the predicted freedom of crossing and nesting should arise in Bulgarian when two *wh*-phrases undergoing overlapping movement are more clearly featurally distinct—in particular, when at most one of them is interrogative. Rudin (1988) reports Bulgarian examples of interrogative *wh*-movement overlapping with relativization that fit this

---

[38]This would instantiate what following Deal (2015a,b) we could classify as an *insatiable* feature.

[39]Another possibility is that an independent factor determines the order of *wh*-phrases in (85). Something along these lines is proposed by Billings and Rudin (1996), who argue that the order of *wh*-phrases in Bulgarian is governed by several (violable) constraints relating to animacy, case and other factors. Under their account, a preference for nominative *wh*-phrases to be initial will capture the contrast in (85), though it is unclear how general such a preference truly is.

prediction, as in (86). In (86a) below, we see an example where these two A′-movements form crossing paths, and in (86b), we see a situation where the same two types of movement reverse order instead, due to a corresponding reversal of their final landing sites:[40]

(86)   *Freedom of ordering for distinct A′-movements in Bulgarian*

   a.   *Order-preserving relativization with interrogative wh-movement*

         čoveka [**kojto**$_1$ ne znaeš   [**kakvo**$_2$ kazvat [če e   kupil t_1
         man   who.REL NEG know.2SG what   say.3PL that has bought
         t_2]]]

         'The man who you don't know what they say that he bought'
         (Adapted from Rudin 1988 ex. 8c)

   b.   *Order-reversing relativization with interrogative wh-movement*

         edna kniga [**kojato**$_1$ se čudja   [koj$_2$ t_2 znae   [**koj**$_3$ t_3 prodava
         a   book which wonder-1SG who   knows who   sells
         t_1]]]

         'A book which I wonder who knows who sells (it)'
         (Adapted from Rudin 1988 ex. 19)[41]

In sum, if overlapping moved phrases attracted by separate features are free to form crossing or nesting paths provided that one of those options isn't independently excluded by linearization (which is why crossing paths are banned outside of multiple *wh*-questions in English), this freedom of ordering in Bulgarian is precisely what we expect to see, at least some of the time: Since this language allows overt inner specifiers formed by tucking-in, we predict that it will allow sentences that, in English, would be unacceptable PCC violations.

---

[40]David Pesetsky (p.c.) points out that (86a) could involve a null resumptive pronoun in subject position rather than subject extraction from the embedded clause, given that Bulgarian allows *pro*-drop. While this possibility should indeed be controlled for, (86a) is in fact less important than (86b). Example (86a) involves crossing paths, which we have already seen to be possible in Bulgarian on the basis of multiple *wh*-questions, but example (86b), which is immune to Pesetsky's note of caution due to lacking subject extraction (since the embedded subjects remain in their CPs of origination) shows nesting paths. Nesting paths have been impossible in other Bulgarian examples we have seen so far. Thus we have evidence that both overt crossing and overt nesting paths can occur in Bulgarian in principle, which is the important point here.

[41]Example (86b) shows relativization crossing not one but two interrogative *wh*-phrases. While the crossing of just one such phrase would be enough to make the point here, Rudin chose to include more than one to emphasize the unboundedness of relativization in this particular configuration.

5.2.1 Correct predictions for Romanian

Like Bulgarian, Romanian is a language with multiple *wh*-fronting (87). According to Rudin (1988), in Romanian all interrogative *wh*-phrases move to specifiers of CP in precisely the same manner as in Bulgarian.

(87) *Romanian multiple wh-fronting*
      Cine$_1$ cui$_2$ ce$_3$ ziceai [că i-a   promis t_1 t_2 t_3]?
      who   who.DAT what said.2SG that him.DAT-has promised

      'Who did you say promised what to whom?'
      (Adapted from Rudin 1988, ex. 9a)

If this is so, then Romanian is also a language that in principle allows overt tucking-in, and thus we expect Romanian to behave exactly like Bulgarian as far as the PCC is concerned. Therefore we expect Romanian to show the same freedom of ordering that we saw in the Bulgarian examples in (86) above. This prediction is correct. In (88) below, we see that overlapping relativization and interrogative *wh*-movement in Romanian may either cross or nest:[42]

(88) *Freedom of ordering for distinct A′-movements in Romanian*

      a. *Nesting paths*

          Ion, [[**pe care**]$_1$ am   uitat [**cine**$_2$ mi-ai   spus că
          John, whom   have.1SG forgot who   me.DAT-have.2SG told that
          t_2 ţi   l$_2$-a   prezentat t_1]]
            you.DAT CLITIC-has introduced

          ''John, whom I forgot who you told me introduced to you''
          (Adapted from Comorovski 1986, ex. 4b)

---

[42]The *wh*-moved direct objects in these Romanian examples are clitic-doubled. Comorovski (1986) argues that this clitic is not something like a resumptive pronoun that occurs in the absence of movement, but rather that *wh*-movement is applying here as usual, with the doubled clitic being simply irrelevant. Example (88b) is also susceptible to the point raised in footnote 40 above: since Romanian is a *pro*-drop language, (88b) could involve base generation of the seemingly extracted subject with a resumptive *pro* in the embedded clause. This possibility is addressed by Comorovski, who argues that null subject resumption is not generally capable of ameliorating locality of movement violations in Romanian, meaning that if examples of this variety are considered *wh*-island violations, subject resumption is not responsible for their acceptability.

b.   *Crossing paths*

? Ion, [**care**$_1$ am   uitat   [[**pe cine**]$_2$ mi-ai   spus
John, who   have.1SG forgotten whom   me.DAT-have.2SG told

că t_1 ţi   l$_2$-a   prezentat t_2]]
that   you.DAT CLITIC-has introduced

'John, who I have forgotten whom you told me introduced to you'
(Donca Steriade, p.c.)

Example (88a) shows nesting paths, while example (88b) shows crossing paths, which we have also seen in the multiple *wh*-question in (87) above. Thus we have evidence that like Bulgarian, Romanian in principle allows both overt crossing and overt nesting, importantly in contrast to English.[43] [44]

---

[43]Comorovski (1986) reports numerous examples of the form in (88a), showing that overt nesting paths are indeed productive in Romanian, in addition to the crossing paths shown in (87) and (88b). This is importantly like Bulgarian but unlike English, the latter of which only allows overlapping overt movement to form nesting paths, as we've seen.

[44]The account of the PCC offered in this chapter predicts that this constraint should hold in a language like English regardless of the movement types involved. That is, whether two (or more) overlapping movement paths are both A′, both A, or one A and one A′, such overlapping paths will result in a multiple specifier structure, whose linearization will constrain the possibilities for the derivation in question. While I am not aware of a configuration that allows us to test whether overlapping A-movements obey the PCC, there is some evidence that the PCC does hold when A′-movement overlaps with A-movement. Holmberg et al. (2019) examine the properties of languages with symmetrical ditransitives, in which either the direct or indirect object is available for movement operations, most importantly for them passivization. As they discuss, some dialects of North-West British English are languages of this sort, and allow passivization of the theme of a ditransitive:

(i)   % [A book]$_1$ was given/sent/handed him t_1 (by Mary).
(North-West British English, Holmberg et al. 2019, ex. 8b)

Holmberg et al. (2019) show that in many languages of this variety, North-West British English included, passivization of the theme of a ditransitive removes the possibility of A′-moving the indirect object, as in (iia). A′-movement of the theme combined with passivization of the indirect object is acceptable, however, as in (iib).

(ii)   a.   * Who$_1$ was [the book]$_2$ given/sent/handed t_1 t_2 (by Mary)?
(North-West British English, Holmberg et al. 2019, ex. 9d)
     b.   [Which book]$_2$ was John$_1$ given/sent/handed t_1 t_2?
(North-West British English, Holmberg et al. 2019, ex. 9b)

Holmberg et al. (2019) argue that a particular approach to phase theory and the nature of multiple specifiers predicts this restriction. However, an examination of the structure of the paths in the above examples reveals a straightforward alternative explanation for the unacceptability of (iia): this example, unlike (iib), violates the PCC. The contrast in (ii) thus provides potential evidence that overlapping A- and A′-movements obey the PCC, as predicted.

5.3   On tucking-in and subjects

I conclude this section with an issue regarding interrogative subjects. For both Bulgarian and English (as well as Romanian as (87) above shows), in a multiple *wh*-question where an external argument is one of the *wh*-phrases, it appears to take the role of the outer specifier of CP, with any lower *wh*-phrases presumably tucking-in beneath it (though covertly in English), as (89-90) show. This tucking-in is apparently obligatory, since the *wh*-phrases may not appear in reversed order here:

(89)   *Multiple wh-question with subject wh in English*

      a.   Who sees whom?

      b.   * Whom who sees?

(90)   *Multiple wh-question with subject wh in Bulgarian*

      a.   Koj kogo vižda?
          who whom sees
          (Rudin 1988, ex. 5a)

      b.   * Kogo koj vižda?
           whom who sees
          (Billings and Rudin 1996, ex. 8b)

At the vP level, these sentences presumably involve tucking-in of the *wh*-object beneath the *wh*-subject in situ, as in (91), prior to these phrases moving on to CP:

(91)   *Tucking-in of wh-object below wh-subject in multiple wh-question*
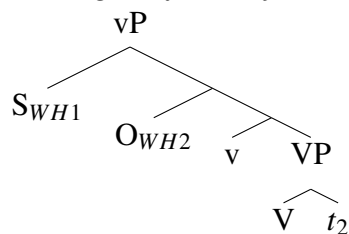
Tangentially, to my knowledge, a sentence like (iib) is also acceptable in mainstream English. In this sentence with overlapping A- and A′-movements, since the A′-moved phrase has the higher final landing site, we also expect it to be the higher of these phrases at the time when they passed through the vP edge. The fact that the *wh*-phrase in such a context can license a PG (in my judgment) verifies this expectation, just as the considerations of this chapter predict:

(iii)   [Which book]$_2$ was John$_1$ given t_1 t_2 for free [despite not being interested in PG$_2$] / [after buying PG$_2$ online]?

Though A-movement is known to independently be incapable of PG-licensing (Culicover and Postal 2001), importantly, if the A′-moved phrase had not formed the outer specifier of the vP here, then we would not expect it to license a PG either. However, evidently it does.

Earlier in this chapter, I proposed that tucking-in is only required in situations where movement of multiple phrases is triggered by one feature, as repeated below:

(92)   *One attracting feature for multiple goals → tucking-in*

$$[_{XP} \; \alpha_{[F1]} \; \beta_{[F1]} \; X_{[uF1]} \; ... \; \alpha \; \beta \;]$$

If external argument subjects like those in (89-90) originate in spec-vP, then the *wh*-object's movement to the lower specifier of vP in (91) was not preceded by movement of the subject, but rather, by external merge of the subject. In actuality, then, the *wh*-object's movement here is in fact the first movement targeting the vP. For this reason, it is not obvious that the rule described by (92) should apply in (91).

    This potential issue can be remedied by positing that tucking-in applies whenever any instances of merge, whether they be internal or external, interact with the same feature on the head of the projection being merged to. This concept can be formalized in the context of a theory where both movement and external merge are feature-driven. See Müller (2010) and references therein for discussion of such an approach, and Longenbaugh (2019, To appear) for recent work in this vein.

    To see how this perspective can be applied to the above puzzle, consider (93) below. In (93a) we have a v bearing an unvalued WH feature [uWH], along with a *wh*-object. If this structure is to become a multiple *wh*-question with a *wh*-subject, that subject must be merged into the specifier of this vP, as in (93b). Following the proposal that external merge is like internal merge in involving a featural relationship with the head targeted for merger, I assume that in (93b) the externally merged *wh*-subject values [uWH], to which [valWH1] is thus appended in (93b). We know that any valuation of [uWH] at this point cannot deactivate this feature: if [uWH] features were satisfied and deactivated after only one instance of merge, multiple *wh*-movement would not be possible. Therefore I assume that next, [uWH] targets the *wh*-object for movement. I hypothesize that because the *wh*-object is the second *wh*-phrase to interact with [uWH], it must tuck-in, as (93c) shows:

(93)   *Derivation of the vP for a multiple wh-question with a wh-subject*

       a.   *Merge of v*
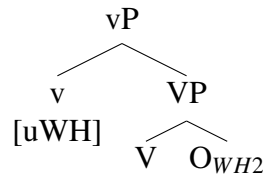
b. *External merge of external argument wh*

```
                        vP
              ┌──────────┴──────────┐
           S_{WH1}                  │
                        ┌───────────┴───────────┐
                        v                       VP
                      [uWH]               ┌─────┴─────┐
                    [valWH1]              V        O_{WH2}
```
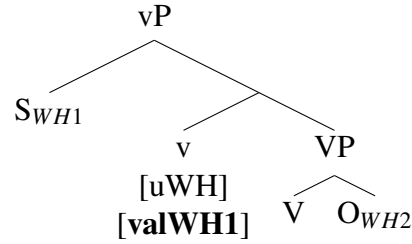
c. *Internal merge of wh-object*

```
                        vP
              ┌──────────┴──────────┐
           S_{WH1}                  │
                        ┌───────────┴───────────┐
                    O_{WH2k}                    │
                             ┌──────────────────┴──────────┐
                             v                             VP
                           [uWH]                     ┌──────┴──────┐
                          [valWH1]                   V           t_k
                          [valWH2]
```
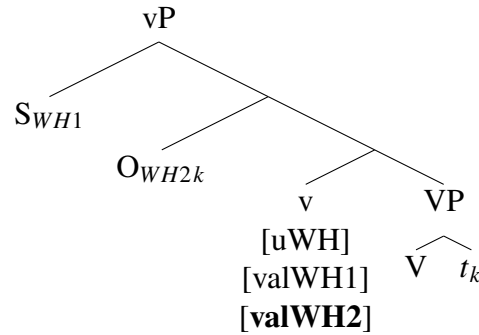
This understanding of external merge as dependent on a featural relationship with the selecting head, when combined with the proposal about tucking-in in (92) above, accurately models the facts about multiple *wh*-questions in (89-90) above.[45] [46] This account has relied on the assumption that the *wh*-subject is merged before movement

---

[45]Allowing a subject externally merged into the specifier of v to value a feature on v, as in (93b), contradicts the rationale for the ban on phrase-bounded edge-internal movement discussed in chapters 1 and 2. To review, Ko (2007, 2011, 2014) argues that this ban emerges from the widely held proposal that a probing feature can only search its c-command domain. If this is so, a movement-triggering feature can never target and attract a constituent that is already in the edge of the containing phrase, since the head bearing that feature does not c-command any material in the edge of that phrase. Similarly, since v does not c-command an external argument merged into its specifier, v should be unable to enter into a featural dependency with this constituent. One potential resolution to this tension is to posit that a feature on a given head can indeed make reference to elements in the specifier of the containing phrase, but that phrase-bounded edge internal movement is independently banned by anti-locality.

[46]A second route of analysis would be to posit that the clause-internal phase is not vP, but a slightly larger constituent, which we might identify as voiceP. If this is so, then in the multiple *wh*-questions under consideration, the *wh*-subject and *wh*-object will be attracted one after the other to the edge of the voiceP, in standard tucking-in fashion:

(i)   ... [$_{voiceP}$ S$_{WH1}$ O$_{WH2}$ voice [$_{vP}$ t_1 v [$_{VP}$ V t_2]]]

This analysis dissolves the issue by denying that subjects originate in the phase edge, but is not obviously compatible with certain results of the previous chapters, especially chapters 3 and 4, for which the concept that external arguments originate in a phase edge is vital.

of the *wh*-object, however. Since this issue is not a primary topic of this chapter (though nevertheless an important one) I will not justify this assumption here. See Longenbaugh (2019, To appear) for some recent relevant discussion of this topic.[47]

6 Stranding in edges with multiple moved phrases

In this section, I re-consider certain phenomena from chapters 2 and 3 in order to examine the predictions this chapter makes about the interaction of overlapping movement with stranding in phase edges. In chapters 2 and 3, I argued that potential stranding in a phase edge is banned when the stranded material would be crossed by an additional instance of movement from that phase. For instance, I argued that a subject's A-movement to spec-TP would cross over any material stranded in the edge of vP by A′-movement, in such a way that would result in a linearization contradiction, unless that stranded material is evacuated. That is, while such stranded material precedes the subject when vP is spelled-out, after the subject exits vP, the subject will be linearized as preceding vP and thus preceding that stranded material. For this reason, stranding in spec-vP in this situation is blocked, as (94) below shows with hypothetical stranding of an element α.

(94)   *Conflict between subject movement and stranding in the vP edge*

$$[_{CP} \text{ WH}(\checkmark\alpha) \text{ C S T } [_{vP} \text{ WH}(*\alpha) \text{ S v-V } \text{ WH}(\checkmark\alpha)]]$$

Furthermore, the ban on phrase-bounded edge-internal movement prevents the constituents of the vP edge from being re-arranged in such a way that would avoid this problem. As chapters 2 and 3 have shown, this understanding makes correct predictions about the distribution of the stranding of *wh*-adjoined adjuncts in mainstream English, as well as *all*-stranding in West Ulster English as reported by McCloskey (2000). Both of these phenomena permit stranding in the CP edge under A′-movement, but not the vP edge:

(95)   *Intermediate stranding of wh-adjoined adjunct*
        a.   *Stranding possible in CP edge*
            Tell me $[_{CP}$ (to the nearest pound) [**how much flour**]$_k$ (to the nearest pound) you said $[_{CP} t_k$ **(to the nearest pound)** (that) the bakery wants t_k (to the nearest pound)]].
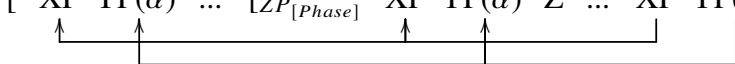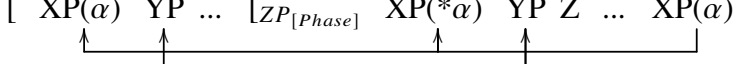
---

[47]Longenbaugh, following Müller, argues for a more complex ontology of featural triggers for syntactic operations than I have assumed here, and furthermore argues that a given instance of merge can satisfy multiple features at once. Such a theory is very likely applicable to the puzzle under discussion here, though I will refrain from pursuing such an account in detail in this context.

     b.   *No stranding in vP edge*
        How much flour (to the nearest pound) did you [$_{vP}$ (***to the nearest pound**) tell me [$_{CP}$ (to the nearest pound) that the bakery [$_{vP}$ (***to the nearest pound**) asked you for (to the nearest pound)]]]?

(96)   *West Ulster English all-stranding*[48]

     a.   *Stranding possible in CP edge*
        What$_k$ (all) did he say [$_{CP}$ t_k (**all**) that we should buy t_k (all)]?
        (McCloskey (2000), ex. 8)

     b.   *No stranding in vP edge*
        What$_k$ did he tell$_j$ [$_{vP}$ t_k (***all**) t_j his friends [$_{CP}$ t_k (all) that he wanted t_k]]?
        (McCloskey (2000), ex. 14e)

Abstracting away from the consideration of the vP in particular, this logic leads us to expect that stranding in any phase edge will be impossible if the stranded material is later crossed by another movement from within the same phase edge. This reasoning leads us to a prediction about intermediate stranding in contexts where multiple phrases undergo overlapping movement from a phase. That is, if two phrases XP and YP successive cyclically move from a given phase such that XP forms the outer specifier of that phase and YP forms the inner one, then stranding in the phase edge should only be possible for YP. This is because an element stranded by XP would be crossed over by the movement of YP, though material stranded by YP will not be crossed by anything. This prediction is illustrated in (97) below.

(97)   *Prediction for intermediate stranding with two overlapping moved phrases*

     a.   *Stranding by inner of two specifiers possible*
        [XP YP(α) ... [$_{ZP_{[Phase]}}$ XP YP(α) Z ... XP YP(α)]]

     b.   *Stranding by outer of two specifiers impossible*
        [XP(α) YP ... [$_{ZP_{[Phase]}}$ XP(*α) YP Z ... XP(α) YP]]

In (97a), intermediate stranding of the element α by YP in the edge of the phase ZP results in no problem, because throughout this derivation, the relative ordering XP < YP < α is maintained regardless of whether stranding occurs or not. In contrast, in (97b) stranding of α by the outermost moving phrase XP results in a linearization problem, because in this case α precedes YP at the spell-out of the phase ZP, but

---

[48]As chapter 2 discusses, the verb movement McCloskey posits for West Ulster English provides a second reason why the stranding in the vP edge attempted in (96b) is unacceptable.

would follow YP if left behind in the edge of ZP after YP moves on.

This prediction is verified by further facts about *wh*-adjoined adjunct stranding in English. These examples are complex since they require three clauses—two to provide landing sites for each overt A′-movement and one more for those movements to pass through and attempt stranding. There is nevertheless a contrast between an example where material is stranded in the lowest clause by the inner moved element (98a), as opposed to stranding by the outer one (98b), as predicted:

(98)   *Adjunct stranding with two moved phrases*

     a.   *Intermediate stranding by inner moved phrase*
         This is the bakery which$_2$ they reported [**how much flour**]$_1$ they thought [$_{CP}$ t_2 t_1 (**to the nearest pound**) that they sent t_1 to t_2].

     b.   *No intermediate stranding by outer moved phrase*
         Tell me [**how much flour**]$_2$ they said who$_1$ they thought [$_{CP}$ t_2 (***to the nearest pound**) t_1 that they should ask t_1 to buy t_2].

This prediction can also be tested with English possessor extraction, which in chapter 3, I argued requires stranding of a non-subject possessum exited by extraction in the edge of an embedded CP, as shown once more in (99):

(99)   *Non-subject stranding in clause edge by possessor extraction*
        Who$_2$ do they think [$_{CP}$ [__$_2$**'s book**]$_1$ Mary read t_1]?
        (Chapter 3, ex. 7b)

As we see in (100) below, when an additional A′-movement is added to such a context, possessor extraction from the inner moved phrase permits such stranding (100a), but extraction from the outer one does not, as expected (100b):

(100)   *Stranding by possessor extraction only from inner moved phrase*

     a.   *Intermediate stranding by inner moved phrase*
         This is the book which$_1$ you mentioned **who$_3$** you think [$_{CP}$ t_1 [t_3 **'s friend**]$_2$ you might talk to t_2 about t_1].

     b.   *No intermediate stranding by outer moved phrase*
         * This is the guy **who$_3$** you mentioned [which book]$_2$ you think [[$_{CP}$ [t_3**'s friend**]$_1$ t_2 we might send t_2 to t_1].

As the traces in (98) and (100) indicate, I have assumed that the movements in these configurations obey the PCC in the initial CP, after which one of the moving phrases and an extracted element from the other phrase undergo order-preserving movement into the higher clause. This initial order reversal followed by order preservation is precisely what the results of the earlier part of this chapter lead us

to expect. Importantly, all the sentences in (98) and (100) have this same path structure, but only those sentences in which the intermediately stranded material is not crossed over within the embedded clause's edge are acceptable. This is what we expect given the prediction in (97), which is an automatic consequence of the concepts defended in this chapter, and this dissertation more generally.
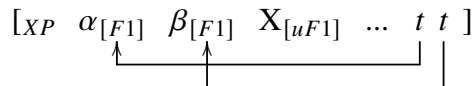
7 Conclusion

In this chapter, I examined configurations with overlapping A′-movement paths. Using certain predictions about PGs and multiple specifiers from Nissenbaum (2000), I argued that successive cyclic movement in such derivations is constrained in the way described by the following generalization, which CL predicts:

(101)   *Overlapping Chain Generalization*
        If two A′-moving phrases XP_1 and XP_2 form specifiers of an intermediate phase YP such that XP_1 c-commands XP_2, then XP_1 c-commands XP_2 in all subsequent positions these phrases occupy.

This result raised a question about about the distribution of crossing and nesting derivations, which I argued is resolved by understanding crossing paths as a requirement of multiple movements triggered by one feature, and crossing versus nesting as being in free variation otherwise:

(102)   *One attracting feature for multiple goals → tucking-in*
        $$[_{XP}\ \alpha_{[F1]}\ \beta_{[F1]}\ \mathrm{X}_{[uF1]}\ ...\ t\ t\ ]$$

(103)   *Attraction by separate features on one head → variability*
        a. $$[_{XP}\ \alpha_{[F1]}\ \beta_{[F2]}\ \mathrm{X}_{[uF1,uF2]}\ ...\ t\ t\ ]$$

        b. $$[_{XP}\ \beta_{[F2]}\ \alpha_{[F1]}\ \mathrm{X}_{[uF1,uF2]}\ ...\ t\ t\ ]$$
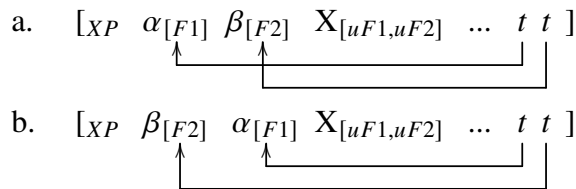
I argued that this understanding facilitates an account of the PCC in English when combined with independent concepts about the distribution of covert movement in this language, as well as the absence of the PCC in the multiple *wh*-fronting languages Bulgarian and Romanian. I also considered these results in the context of a few other topics about the nature of overlapping paths. I will further explore several of the concepts addressed in this chapter in chapter 7, which examines a particular under-analyzed configuration involving parasitic gaps in relative clauses.

Before that, however, in the next chapter I will shift focus to an investigation of the linearization of rightward extraposition.

The following appendix discusses a remaining puzzle of great interest to the proposals of this chapter, as well as bottom-up theories of cyclic spell-out in general.

8   Appendix: A puzzle about multiple *wh*-questions

Following Richards (1997, 1999) and Nissenbaum (2000), in this chapter I took the covertness of non-initial A′-movements in English to be a result of the fact that such movement tucks-in beneath the phrase that moves first:

(104)   *Covert non-initial wh-movement*
         Who$_1$ $\varnothing_2$ did Ivan ask t_1 **what**$_2$?

A problem for this view arises in multiple *wh*-questions in which the *wh*-phrases originate in separate clauses. In such sentences the initially higher *wh*-phrase overtly moves to the specifier of the interrogative C, while the initially lower phrase remains in situ, at least on the surface, as in (105). In other words, such multi-clausal multiple *wh*-questions behave exactly like their mono-clausal counterparts:

(105)   *Lower WH separated from higher by clause boundary remains in situ*
         a.   **Who**$_1$ speculated (that) Mary bought **what**$_2$?
         b.   [**Which student**]$_1$ thinks (that) their parents donated [**which school supplies**]$_2$?
         c.   **Who**$_1$ did you notify t_1 (that) you finished [**which books**]$_2$?

This fact is a puzzle for a theory like the one pursued in this dissertation, for which syntactic structures are constructed and linearized bottom-up, phase-by-phase. Under such a view of the grammar, it is unclear, for instance, how the initially lower *wh*-phrase in such a context can have the foresight to not first undergo overt movement, and then switch to covert movement at the time that it becomes local to the higher *wh*-phrase, as in (106):

(106)   *No partial overt movement of lower WH*
         a.   * **Who**$_1$ speculated **what**$_2$ (that) Mary bought t_2?
         b.   * [**Which student**]$_1$ thinks [**which school supplies**]$_2$ (that) their parents donated t_2?
         c.   * **Who**$_1$ did you notify t_1 [**which books**]$_2$ (that) you finished t_2?

This fact thus presents a potential look-ahead problem for a theory like that adopted in this chapter, as well as any theory involving bottom-up cyclic spell-out.

This issue would be avoided in a top-down approach to syntactic derivations, but it is unclear how the insights of phase theories can be recast in terms of a top-down grammar. This problem could also be resolved by a theory in which the linearization of movement chains is determined representationally rather than derivationally. In the context of such a system, it would suffice to state that at the end of the derivation, the structurally higher *wh*-phrase must be pronounced at the head of its chain and the structurally lower one must be pronounced at the tail of its chain. See Bošković (2001, 2002) for work in this direction.

While a representational theory of how movement is linearized is not obviously compatible with the results of this dissertation, a reconciliation might come about by positing that linearization in the sense focused on here is in fact evaluated prior to any insertion of phonological information. If the considerations of linearization actually play out at the level of abstract syntactic terminals, without any reference to whether or not those terminals have been assigned phonological form yet or not, then the positions in which those terminals are actually pronounced will be irrelevant to the evaluation of ordering contradictions. As discussed in chapter 8, some works indeed argue that linearization is evaluated prior to the insertion of morpho-phonological form in this way. This perspective of the grammar is likely independently necessary given that, as chapter 8 discusses, it appears that the considerations of CL at least some of the time make reference to the position of silent material. Since it is difficult to identify additional facts about sentences like (105) based on which something more concrete might be said about them, I will end this discussion here.

Chapter 6

Ordering extraposition

1 Introduction

The previous chapters have all examined leftward movement, which I've argued is subject to numerous word order constraints that emerge naturally from CL, and its interaction with independent limitations on when syntactic movement can apply. In this chapter, I argue that this approach to cyclic spell-out also deepens our understanding of rightward displacement, which patterns unlike its more familiar leftward counterpart in several interesting ways. I will focus primarily on rightward displacement of non-arguments, generally adjuncts, which I will use the term *extraposition* to refer to. A stereotypical instance of extraposition is rightward displacement of relative clauses, exemplified in (1) below. Here we see relative clauses which appear at the right edge of the sentence, separated from the DP they are interpreted as modifying, which I'll refer to as the *source* of extraposition:

(1)   *Relative clause extraposition*

     a.   *English*
           I bought [**a book** t_1] yesterday [**that someone recommended to me**]$_1$.

     b.   *Wolof (Niger-Congo)*
           Gis-naa [**ab fas** t_2] démb   [**wu nga sopp**]$_2$.
           see-1SG a   horse   yesterday that you like
           'I saw a horse yesterday that you like'

The concept that extraposition is constrained by concerns of linear order has precedent in a few previous works (Jenks 2011, 2013a,b; Nissenbaum 2000; Fox and Pesetsky 2009). As far as I know, no recent literature has gathered all of the relevant facts about the relationship between extraposition and word order, and considered a linearization account of them in detail. The goal of the present chapter

is to fill this gap. Here I take stock of the observations of previous research on this topic as well as some convergent new facts (primarily from Wolof and Javanese[1]) as a basis for considering how rightward displacement and linearization interact.

A considerable amount of research in syntactic theory has argued that distinct derivations are responsible for rightward displacement of adjuncts and of arguments (see Baltin (2006) for an overview). In brief, while rightward displacement of arguments has many properties typical of phrasal movement,[2] non-argument extraposition has characteristics suggesting the absence of a movement dependency of the usual variety. In response to such findings, a number of works have argued that non-argument extraposition uniquely involves an exceptional case of base generation (Culicover and Rochemont 1990; Fox and Nissenbaum 1999; Nissenbaum 2000; Fox 2002; Sheehan 2010; Overfelt 2015, a.o.). In this chapter, I will argue that an analysis under which extraposition is derived by late external merge (Lebeaux 1988, a.o.) after (covert) movement of the source phrase (Fox and Nissenbaum 1999; Nissenbaum 2000; Fox 2002; Overfelt 2015 a.o.) is best suited to capture a particular cross-linguistic word order generalization about this phenomenon. This analysis of extraposition was initiated by Fox and Nissenbaum (1999), and while this chapter will maintain their essential insights, it will also implement extraposition in a considerably different way, as we'll see.

The primary goal of this chapter is to justify and account for the generalization about extraposition in (2) below, which has been hypothesized, in different ways, by at least Jenks (2011, 2013a,b), Nissenbaum (2000), and Fox and Pesetsky (2009):

---

[1]All Wolof data in this chapter, unless otherwise cited, comes from my fieldwork with speakers Lamine Diallo, Lamine Toure, and Aisha Toure at MIT in 2018, while all Javanese data reported here comes from my fieldwork with native speaker Ismartilah Drummond at MIT in 2019. This chapter also owes a great deal to Danny Fox, as described further in footnote 3.

[2]Though similar to leftward phrasal movement in many ways, rightward displacement of arguments is nevertheless more constrained. Particularly prominent among its limitations is the *right roof constraint*, which states that rightward movement of a given element can proceed no higher than the edge of the local cyclic domain (Akmajian 1975, a.o.). Sabbagh (2007) and Overfelt (2015) argue (focusing respectively on right node raising and heavy NP shift) that this constraint is only apparent: rightward movement is potentially unbounded, but becomes more restricted in certain circumstances where its derivation cannot be appropriately successive-cyclic. Both works argue, in different ways, that CL is central in constraining possible rightward movement derivations, a result which is directly aligned with the framework that this dissertation advances. These works do not deal much with non-argument extraposition, however, for which the right roof constraint is generally thought to hold as well. I will leave the right roof constraint aside in this chapter, in order to focus on the conditions under which extraposition is possible in the first place.

(2)   *Source ordering generalization (SOG)*
       Extraposition of a non-argument constituent α from a source phrase XP must result in α having the same relative linear order with respect to the rest of the content of XP that it would have had in the absence of extraposition.

The SOG describes the fact that only elements that independently would surface at the right edge of the source phrase in a given context are capable of extraposition—a fact likely first observed by Nissenbaum (2000) on the basis of contrasts between English and Hebrew, as we'll see shortly. This concept is illustrated in (3) below:

(3)   *Predictions of the SOG*
       a.   *Extraposition from right linear edge of source*
             ✓ [$_{XP}$ X ... α] ... α
       b.   *No extraposition from linear interior of source*
             * [$_{XP}$ α X ...] ... α

A telling instance of this generalization comes from Wolof (Niger-Congo). In this language, indefinite determiners are pre-nominal, as in (1b) above, while definite ones and demonstratives are post-nominal, as in (4a) below, following any relative clauses. My fieldwork research has found that relative clause extraposition in Wolof is possible only with pre-nominal determiners, as (1b) showed, not post-nominal ones and demonstratives (4b), with an interesting exception that we'll see later on:

(4)   *Post-nominal determiner blocks extraposition in Wolof*
       a.   Gis-naa [fas   [**wu nga sopp**] wi/wee] démb.
             see-1sg horse that you like   the/that yesterday
             'I saw the/that horse yesterday that you like'
       b.   * Gis-naa [fas t_1 wi/wee] démb   [**wu nga sopp**]$_1$.
             see-1sg horse   the/that yesterday that you like
             'I saw the/that horse yesterday that you like'

The concept that linear crossing interrupts extraposition is reminiscent of other constraints on displacement we've seen throughout this dissertation, particularly with regard to intermediate stranding as analyzed in chapter 2. While the nature of extraposition will necessitate an account rather different than for crossing effects with leftward movement, I will argue that in this domain as well, a linearization-based explanation facilitates a number of desirable consequences.

As mentioned above, here I will pursue the late merge theory of extraposition. Previous works arguing for this analysis, such as Fox and Nissenbaum (1999)

and Nissenbaum (2000), assume that covert material is visible for linearization: a concept that is unexpected, if linear ordering is an aspect of PF, from which covert material is by definition absent. The linearization of covert material is particularly important to Nissenbaum (2000), who uses this concept to account for asymmetries in extraposition that mirror the SOG. In this chapter, I will pursue an account of this generalization that does not rely on the assumption that covert material is ordered. I will argue in the latter part of this chapter that such an understanding provides insight into several other facts about extraposition as well.

1.1   Chapter contents

In section 2, I provide cross-linguistic evidence for the SOG. In section 3, I provide preliminary discussion of the nature of the SOG and its relationship to attested theories of extraposition, ultimately leading to my adoption of the late merge theory. In section 4, I argue that combining the late merge theory of extraposition with CL yields an explanation for the SOG which eschews the concept of ordered covert operations. Here I also address related issues about where extraposed phrases can surface, adopting a modified version of a constraint on late merge and linear order from Nissenbaum (2000). In section 5, I show how this account of the SOG facilitates an understanding of extraposition from overtly moved phrases. In section 6, I consider an interaction between extraposition and ellipsis observed by Takahashi and Ohtaka (2017), which further supports the proposal that covert movement is not ordered. In section 7, I discuss a constraint on the depth of late merge that is both required for this analysis and independently proposed (in various forms) in previous literature, before concluding in section 8.[3]

2   Evidence for the Source Ordering Generalization

In this section I provide the evidence for the SOG, incorporating observations from previous research with a few new patterns. As far as I know, the first proposal of a word order constraint on extraposition comes from Nissenbaum (2000), who argues based on the contrast between English adjectives and relative clauses in (5) that only

---

[3]This chapter owes a considerable debt to lecture notes from a syntax seminar taught by Danny Fox and David Pesetsky in 2003, and to numerous discussions with Danny Fox in particular, who deserves credit for the majority of the concepts that facilitate this chapter's findings. The general ideas in sections 4, 5, and 6 are all inspired by those lecture notes and by Danny's thoughts, though my decisions about how exactly these ideas must be implemented and justified differs in various details. The incorporation of the linear condition on late merge from Nissenbaum (2000) into the theory of extraposition in sections 4 and 5 is also my own innovation.

rightmost constituents of DP can be extraposed. As we see in (5), while adjectives and relative clauses are both presumably adjuncts, the former originates preceding N and cannot extrapose, while the latter originates after N and can extrapose. PP adjuncts of NP, which also originate post-nominally, can extrapose as well (5e).

(5)   *Left/right asymmetry in extraposition: Relative clause versus adjective*

      a.   I saw [a **man** t_1] yesterday [**who was very tall**]$_1$.

      b.   * I saw [a t_1 **man**] yesterday [**very tall**]$_1$.
          (Adapted from Nissenbaum 2000, pg. 208, ex. 33)

      c.   I saw [a **dog** t_1] yesterday [**that was wearing a blue collar**]$_1$.

      d.   * I saw [a t_1 **dog**] yesterday [**extremely large and intimidating**]$_1$.

      e.   I saw [a **dog** t_1] yesterday [**with a silly hat on**]$_1$.
          (My examples)

Since the adjective undergoing attempted extraposition in (5b) is fairly phonologically small, we might posit that this example is unacceptable due to a prosodic problem. This is a reasonable hypothesis, given that another rightward displacement phenomenon, heavy NP shift, is well-known to be unacceptable if the shifted nominal does not have enough phonological weight. In order to dispense with this hypothesis, I have included example (5d), in which we see that an increase in phonological weight does not make adjective extraposition acceptable. This is what we expect, if what we are really dealing with here is a word order restriction.

Nissenbaum (2000) also shows (citing personal communication from Danny Fox) that in Hebrew, which has post-nominal adjectives, examples like (5b/d) are possible. This fact is also noted by Fox and Pesetsky (2009), who report that while extraposition of an originally post-nominal adjective or modifier succeeds in Hebrew (6a-b), extraposition of an originally pre-nominal modifier does not (6c):

(6)   *Left/right extraposition asymmetry in Hebrew*

      a.   ? Yosef raʔa [**iša** t_1] etmol   [**gvohaa beyoter**]$_1$.
           Yosef saw   woman   yesterday tall   in-more
           'Yosef saw a very tall woman yesterday'

      b.   ? Yosef raʔa [**anašim** t_1] etmol   [**rabim   meʔod**]$_1$.
           Yosef saw   people   yesterday many-pl a-lot
           'Yosef saw very many people yesterday'

      c.   * Yosef raʔa [t_1 **anašim**] etmol   [**harbe meʔod**]$_1$.
           Yosef saw   people   yesterday many   a-lot
           'Yosef saw very many people yesterday'
           (Fox & Pesetsky 2009, ex. 6-7)

Fox and Pesetsky (2009) go on to show, reporting an observation they credit to Alex Grosu, that in English adjective extraposition from the compound quantifier phrase *someone* is possible. This is expected, since adjectives originate on the right side of this variety of DP, as we see in (7a). I observe that the same facts hold for adjectives with analogous quantifier phrases, as (7) illustrates:[4]

(7)   *Post-nominal adjectives of certain quantifiers can extrapose*
    a.   I saw [**someone / somebody** [**quite tall**]] yesterday.
    b.   ? I saw [**someone / somebody** t_1] yesterday [**quite tall**]$_1$.
    c.   Mary met [**everyone / everybody** [**somewhat interesting**]] just now.
    d.   ? Mary met [**everyone / everybody** t_1] just now [**somewhat interesting**]$_1$.
    e.   I talked to [**nobody / no one** [**particularly unusual**]] tonight.
    f.   ? I talked to [**nobody / no one** t_1] tonight [**particularly unusual**]$_1$.

We can tell that the adjectives in (7) are unlikely to be reduced relative clauses, since if this were the case, they should be able to post-nominally merge in typical DPs/QPs as well, contrary to fact:

(8)   *No post-nominal adjectives in typical English DPs/QPs*
    a.   * I saw a person **quite tall** yesterday.
    b.   * Mary met every person **somewhat interesting** just now.
    c.   * I talked to no people **particularly unusual** tonight.

Further potentially relevant facts are evident for extraposition from possessors. In my judgment, a relative clause construed as modifying a pre-nominal possessor cannot be extraposed in English. This is natural if the containing possessum DP

---

[4]The extraposed adjectives in example (7) all include either an intensifier, or some other element relating to a degree. In my judgment, the absence of a degree-encoding item makes this variety of extraposition significantly degraded:

(i)   a.   * I saw **someone** today **tall**.
    b.   * Mary met **everyone** just now **interesting**.
    c.   * I talked to **nobody** tonight **unusual**.

Even if such extraposition is for reasons yet to be examined contingent on the presence of a degree, note that the degree-encoding element in examples (5b/d) and (6c) above does not facilitate extraposition in those contexts. This shows that word order is indeed the operative factor.

    In the context of the late merge theory of extraposition that this chapter will adopt, the fact that only degree-encoding adjectives can extrapose may indicate that only these adjectives are capable of late merge. This would suggest, in the context of the late merge theory of Lebeaux (1988), that basic adjectives are not adjuncts, or at least are not elements that combine with their host via Predicate Modification. I cannot comment on the viability of this hypothesis here.

qualifies as the source of extraposition in this context, since the relative clause would need to linearly cross the content of this DP in order to extrapose:

(9)   *No relative clause extraposition from a possessor*

    a.   I'm gonna read [[**someone** [**who studies ancient pottery**]]'s new article] today.

    b.   * I'm gonna read [[**someone** t_1]'s new article] today [**who studies ancient pottery**]$_1$.

    c.   I wrote down [**my grandma** [**who loves meat**]'s favorite brisket recipe] last week.

    d.   * I wrote down [**my grandma** t_1's favorite brisket recipe] last week [**who loves meat**]$_1$.

Additionally, in my judgment, extraposition from a DP embedded within another notably improves when the structure differs so that the relative clause of the embedded DP is at the linear edge of the containing one:

(10)   *Extraposition from embedded DP permitted from edge of containing DP*

    a.   I published [an article about [**an unusual person** [**who lives in a beehive**]]] this month.

    b.   I published [an article about [**an unusual person** t_1]]] this month [**who lives in a beehive**]$_1$.

    c.   Let's paint a portrait of [**a student** [**who is wearing a big hat**]] today.

    d.   Let's paint a portrait of [**a student** t_1] today [**who is wearing a big hat**]$_1$.

It should be noted that these facts are not guaranteed to fall under the purview of the SOG, since they involve nested DP structures, for which it is not completely clear how the relevant notion of "source" should be defined. This pattern is nevertheless analogous enough to the SOG to be worth mentioning here.[5]

---

[5]Yet another analogous fact is that in a nested DP structure, adding a right-leaning adjunct to the outer DP blocks extraposition from the inner DP:

(i)   a.   I published [an article about [**an unusual person** [**who lives in a beehive**] [that's 100 pages long]] this month.

    b.   * I published [an article about [**an unusual person** t_1] [that's 100 pages long]]] this month [**who lives in a beehive**]$_1$.

    c.   Let's create [a sculpture of [**a monkey** [**who is wearing a straw hat**]][made out of granite]] today.

    d.   * Let's create [a sculpture of [**a monkey** t_1][made out of granite]] today [**who is wearing a straw hat**]$_1$.

Further evidence for the SOG comes from Javanese (Austronesian). Javanese permits extraposition of both relative clauses (11) and adjunct PPs (12):

(11)   *Javanese relative clause extraposition*[6]

    a.   [Srabi   ne [**sing di senengi Budi**]] sing endi sing di tuku
        pancake PRT which PST likes   Budi   which one which PST buy
        dheweke?
        he
        'Which pancake that Budi liked did he buy?'

    b.   Srabi   ne t_1 sing endi sing di tuku dheweke, [**sing di**
        pancake PRT   which one which PST buy he   which PST
        senengi Budi]$_1$?
        like   Budi
        'Which pancake did he buy that Budi liked?'

(12)   *Adjunct PP extraposition in Javanese*

    a.   Kowe wis   maca [**buku** [**ngenani kewan**]] nang sekolah.
        you   already read book about   animals at   school
        'You read a book about animals at school.'

    b.   Kowe wis   maca [**buku** t_1] nang sekolah [**ngenani kewan**]$_1$.
        you   already read book   at   school about   animals
        'You read a book about animals at school.'

In Javanese adjectives are post-nominal, and importantly, they can extrapose as well.

A similar pattern occurs in coordinate structures. In my judgment, if only one of two coordinated DPs serves as a source for extraposition, it must be the rightmost one:

(ii)   a.   I painted a tree and a **horse** today *[with a bird's nest in it] / [**with an apple in its mouth**].
    b.   I painted a horse and a **tree** today [**with a bird's nest in it**] / *[with an apple in its mouth].

The facts in (i), at least, are reminiscent of the right roof constraint (see footnote 2). However, assuming as I will in this chapter that DPs are phases, the acceptable sentences in (i) and (9-10) violate the right roof constraint, since for these extraposition passes through multiple DPs. It is worth asking to what extent the SOG and the right roof constraint might emerge from the same principles. In this chapter I will argue that the SOG arises from the nature of linearization, and as mentioned in footnote 2, Sabbagh (2007) and Overfelt (2015) reach the same conclusion about the right roof constraint. A direct comparison between these works and this chapter is not straightforward, but this point of similarity suggests that something more general might be said about this body of issues.

    [6]Javanese examples like (11) are likely clefts. For instance, (11a) might be more faithfully translated as "*It's which pancake which is the one which Budi likes?*". Since I lack in-depth data about this construction, I will not analyze it further here.

This is unsurprising, given the facts from English and Hebrew shown above:

(13)   *Adjective extraposition in Javanese*

    a.   Aku uwis   mangan [gado-gado [**enak   banget**]] nang sekolah.
            I   already eat   gado-gado delicious very   at   school
            'I ate delicious gado-gado at school.'

    b.   Aku uwis   mangan [gado-gado t_1] nang sekolah [**enak**
            I   already eat   gado-gado   at   school delicious
            banget$]_1$.
            very
            'I ate delicious gado-gado at school.'

Importantly, the post-nominal quantifier *kabeh* ("all/every"), which follows nominal-internal adjuncts, prevents them from extraposing (14). This is what we expect, if extraposition can only succeed for an element that would otherwise inhabit the edge of the source DP in the context in question:

(14)   *No adjunct extraposition from the non-edge of DP in Javanese*

    a.   Kowe wis   maca [buku [**ngenani kewan**] **kabeh**] nang sekolah.
            you   already read book about   animals every   at   school
            'You read every book about animals at school.'

    b.   * Kowe wis   maca [buku t_1 **kabeh**] nang sekolah [**ngenani**
             you   already read book   every   at   school about
            kewan$]_1$.
            animals
            'You read every book about animals at school.'

Additional facts of a similar character are examined by Jenks (2011, 2013a,b), who investigates quantifier float in East Asian languages. Jenks focuses on Thai facts like (15), in which we see that a quantifier can be displaced to the right of the noun phrase with which it is construed:

(15)   *Rightward quantifier displacement in Thai*

    a.   Nák.rian **thúk-khon**   ʔaan nàŋsɯ̌ɯ mɯ̂ɯawaanníi.
            student   every-CLF$_{Person}$ read book   yesterday
            'Every student read a book yesterday.'

    b.   Nák.rian ʔaan nàŋsɯ̌ɯ mɯ̂ɯawaanníi **thúk-khon**.
            student   read book   yesterday   every-CLF$_{Person}$
            'Every student read a book yesterday.'
            (Adapted from Jenks 2013b, ex. 1)

265

Jenks argues that in Thai, numeral quantifier float of this variety is essentially ex-
traposition, in contrast to the proposal for other languages like Korean and Japanese
that Q-float is derived (at least in some cases) by stranding (Saito 1985; Ko 2014;
Miyagawa 2017, a.o.) Jenks states that this form of rightward quantifier float is
attested in languages that independently allow N < Q order such as Khmer, Tibeto-
Burman, and Southwestern Tai, but not those that only allow Q < N order such as
Vietnamese, Chinese, Hmong-Mien, as well as North and Central Tai.

The facts Jenks reports for Malay (citing personal communication from Dan
Kaufman) demonstrate this concept neatly, since the positions for numeral quantifiers
are variable in this language. In particular, both N < Q and Q < N order are attested
for objects (16), while only Q < N order is permitted for subjects (17):

(16)   *N < Q and Q < N order permitted for Malay object*

    a.   Saya beli [[**tiga ekor**] sapi].
        I   buy three CLF   cow
        'I bought three cows'

    b.   Saya beli [sapi [**tiga ekor**]].
        I   buy cow three CLF
        'I bought three cows'
        (Jenks 2013a, ex. 29)

(17)   *Only Q < N order permitted for Malay subject*

    a.   [[**Tiga ekor**] sapi] makan semua rumput.
        three CLF   cow eat   all   grass
        'Three cows ate all the grass'

    b.   ?? [Sapi [**tiga ekor**]] makan semua rumput.
        cow three CLF   eat   all   grass
        'Three cows ate all the grass'
        (Jenks 2013a, ex. 30)

As expected, quantifier float is only possible from objects in Malay (18):

(18)   *Q-float permitted only from object in Malay*

    a.   Saya beli **sapi** kemarin   [**tiga ekor**].
        I   buy cow yesterday three CLF
        'I bought three cows yesterday'

    b.   * **Sapi** makan rumput [**tiga ekor**].
        Cow eat   grass   three CLF
        'Three cows eat grass'
        (Jenks 2013a, ex. 31b)

Ultimately, Jenks proposes the generalization in (19), which he hypothesizes may essentially be a sub-case of Nissenbaum's original observation that extraposition is only possible for elements that originate at the right edge of the source DP.

(19)   *Quantifier Float Generalization*
Rightward quantifier float is only attested in languages which allow a DP-internal order where N precedes the (numeral) quantifier/classifier.
(Adapted from Jenks 2013, ex. 39)

While Jenks does not consider data from all languages mentioned in great detail, and thus potentially conflates instances of Q-float that are not syntactically homogeneous, Jenks' observation is clearly analogous to what the SOG predicts.

Wolof is another language in which extraposition is productive. A few examples of relative clause extraposition in this language are shown below:[7]

(20)   *Wolof relative clause extraposition*

    a.   Lekk-naa [**gato** [**bu nex lool(u)**]] démb.
            ate-1sg   cake that tasty very   yesterday
            'I ate a very tasty cake yesterday'

    b.   Lekk-naa [**gato** t_1] démb   [**bu nex lool(u)**]$_1$.
            ate-1sg   cake   yesterday that tasty very
            'I ate a very tasty cake yesterday'

    c.   Sacc-naa [**xar** [**bu ndaw**]] tey.
            stole-1sg sheep that small   today
            'I stole a small sheep today'

    d.   Sacc-naa [**xar** t_1] tey   [**bu ndaw**]$_1$.
            stole-1sg sheep   today that small
            'I stole a small sheep today'

    e.   Gis-naa [**kër** [**gu yaatu**]] démb.
            see-1sg house that big   yesterday
            'I saw a big house yesterday'

    f.   Gis-naa [**kër** t_1] démb   [**gu yaatu**]$_1$.
            see-1sg house   yesterday that big
            'I saw a big house yesterday'

The above examples show extraposition from bare nominals. Extraposition is also

---

[7]See Torrence (2013) for background on the syntax of Wolof. Though I will not discuss the characteristics of Wolof in detail here, I note that as far as I know, no previous work on Wolof has studied the possibility of extraposition in this language.

possible from DPs containing pre-nominal material, as we see for an indefinite determiner and numeral below. Since Wolof relative clauses follow N, just as in languages like English, the fact that extraposition is possible here is not surprising:

(21)   *Wolof extraposition compatible with pre-nominal material*

    a.   Gis-naa [**ab** fas   [**wu nga sopp**]] démb.
           see-1sG a   horse that you like   yesterday
           'I saw a horse that you like yesterday'

    b.   Gis-naa [**ab** fas t_1] démb   [**wu nga sopp**]$_1$.
           see-1sG a   horse   yesterday that you like
           'I saw a horse yesterday that you like'

    c.   Sacc-naa [**ñaar-i xar**   [**yu ndaw**]].
           stole-1sG two   sheep that small
           'I stole two small sheep'

    d.   Sacc-naa [**ñaar-i xar** t_1] tey/démb   [**yu ndaw**]$_1$.
           stole-1sG two   sheep   today/yesterday that small
           'I stole two small sheep today/yesterday'

In contrast, Wolof demonstratives and definite determiners follow N, and any co-occurring relative clauses:

(22)   *Wolof relative clauses cannot follow post-nominal D*

    a.   Sacc-naa xar   [**bu ndaw**] **bi**.
           stole-1sG sheep that small   the
           'I stole the small sheep'

    b.   * Sacc-naa xar   **bi** [**bu ndaw**].
             stole-1sG sheep the that small
           'I stole the small sheep'

    c.   Lekk-naa gato [**bu nga indi**]   **bi**.
           eat-1sG   cake that you brought the
           'I ate the cake that you brought'

    d.   * Lekk-naa gato **bi**   [**bu nga indi**]$_1$.
             eat-1sG   cake the that you brought
           'I ate the cake that you brought'

As expected given the SOG, these post-nominal elements block extraposition:

(23)   *No extraposition with post-nominal determiner/demonstrative in Wolof*

     a.   Gis-naa [fas   [**wu nga sopp**] **wi/wee**] démb.
           see-1SG horse that you like   the/that yesterday
           'I saw the/that horse that you like yesterday'

     b.   * Gis-naa [fas t_1 **wi/wee**] démb   [**wu nga sopp**]$_1$.
           see-1SG horse   the/that yesterday that you like
           'I saw the/that horse yesterday that you like'

     c.   Sacc-naa [juroom-i xar   [**yu ndaw**] **yii/yee**]   tey.
           stole-1SG 5   sheep that small   the/these today
           'I stole the/those five small sheep today'

     d.   * Sacc-naa [juroom-i xar t_1 **yii/yee**]   tey   [**yu ndaw**]$_1$.
           stole-1SG 5   sheep   the/these today that small
           'I stole the/those five small sheep today'

A post-nominal quantifier also blocks extraposition:

(24)   *No extraposition with post-nominal Q in Wolof*

     a.   Sacc-naa [xar   [**yu ndaw**] **yëpp**].
           stole-1SG sheep that small   all
           'I stole all the small sheep'

     b.   * Sacc-naa [xar t_1 **yëpp**] tey   [**yu ndaw**]$_1$.
           stole-1SG sheep   all   today that small
           'I stole all the small sheep today'

     c.   Lekk-naa [mango [**yu rëy**] **yëpp**] démb.
           eat-1SG   mango that big   all   yesterday
           'I ate all the big mangos yesterday'

     d.   * Lekk-naa [mango t_1 **yëpp**] démb   [**yu rëy**]$_1$.
           eat-1SG   mango   all   yesterday that big
           'I ate all the big mangos yesterday'

     While Wolof demonstratives are typically post-nominal as shown above, they can become pre-nominal under focus. In this context, relative clause extraposition becomes possible, as we see below:

(25)   *Fronted focused demonstrative in Wolof permits extraposition*

     a.   Gis-naa [(**yii**)   góór [**yu njool**] (**yii**)].
           see-1SG (THESE) men that tall   (these)
           'I saw these / THESE men who were tall'

b. Gis-naa [**yii** góór t_1] démb   [**yu njool**]$_1$.
    see-1SG THESE men   yesterday that tall
    'I saw THESE men yesterday who were tall'

c. Lekk-naa [**bii** gato [**bu nex lool**]] tej.
    ate-1SG   THIS cake that tasty very today
    'I ate THIS very delicious cake today'

d. Lekk-naa [**bii** gato t_1] tej   [**bu nex lool**]$_1$.
    ate-1SG   THIS cake   today that tasty very
    'I ate THIS very delicious cake today'

e. Gis-naa [**bii** muus t_1] demb   [**bu rey lool**]$_1$.
    saw-1SG THIS cat   yesterday that big very
    'I saw THIS cat yesterday that was very big'

f. Indi-nga   [**bii** mango t_1] demb   [**bu rey lool**]$_1$.
    brought-2SG THIS mango   yesterday that big very
    'You brought THIS mango yesterday that was very big'

This pattern indicates that the mere presence of a demonstrative does not somehow preclude extraposition in Wolof. The demonstrative only blocks extraposition if it would prevent it from obeying the SOG, as expected if concerns of linear order constrain the distribution of extraposition.

To conclude this section, we have seen evidence from a variety of languages that the following generalization holds:

(26)   *Source ordering generalization (SOG)*
        Extraposition of a non-argument constituent α from a source phrase XP must result in α having the same relative linear order with respect to the rest of the content of XP that it would have had in the absence of extraposition.

Next, I will discuss the nature of the SOG and its relationship to attested theories of extraposition, ultimately settling on the late merge theory. I also summarize previous claims from Nissenbaum (2000) about the relationship between extraposition and word order, as a prelude to the CL account that I argue for.

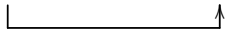# 3   Preliminary discussion and background

Before presenting this chapter's account of the SOG, it will be useful to consider what the shape of this generalization suggests about the nature of extraposition, which must be established in order to provide a context for an account of its restrictions. Consider the following illustration of the generalization in question once more:

(27)   *Predictions of the SOG*

    a.   *Extraposition from right linear edge of source*

        ✓ [$_{XP}$ X ... α] ... α

    b.   *No extraposition from linear interior of source*

        * [$_{XP}$ α X ...] ... α

In the context of CL, assuming the source phrase XP to be a phase, we indeed expect the derivation in (27a) to be successful. This is because in this case α crosses no content in XP by moving from its right edge. We also expect the illicitness of (27b), since such extraposition from the interior of XP, which crosses over some of its content on the way out, should create an ordering contradiction. This line of reasoning raises another possibility, however: if the problem with (27b) is that it involves the extraposed phrase crossing some material within XP upon exiting it, then we would expect a configuration like this to be licit if the extraposed phrase were to successive cyclically move to the right edge of XP before moving on (28):

(28)   *Hypothetical successive cyclic extraposition: Predicted to be licit*

    [$_{XP}$ α X ... α] ... α

Unlike extraposition, the origination position for leftward movement need not be at the linear edge of the domain(s) it exits, precisely because such movement can successive cyclically pass through the edge of each phase it crosses:

(29)   *Successive-cyclic leftward movement*

    [$_{XP}$ α X ... [$_{YP}$ t Y ... [$_{ZP}$ t Z ... t]

Since the possibility of the successive cyclic extraposition derivation in (28) does not accurately predict the arising of the SOG, the necessary conclusion is that extraposition cannot be derived by successive cyclic movement.[8]

    One way of understanding this result would be to posit that extraposition is indeed rightward phrasal extraction, but that a locality issue (for instance a form of anti-locality, discussed extensively in chapter 2) prevents it from being successive cyclic. Prima facie, it is not obvious that there is one relevant notion of locality that can be invoked for all of the facts reported in the previous section, however. Another possibility is that extraposition cannot be successive cyclic because it is in fact not

---

[8]In the context of a consideration of the right roof constraint, Baltin (2006) citing discussion in Akmajian (1975) also notes this consequence: if extraposition could be successive cyclic, why would the right roof constraint ever apply? As mentioned in footnote 2 above, I leave the right roof constraint aside here, but this is nevertheless a parallel worth noting.

extraction. While a locality account of the lack of successive cyclic extraposition may be tenable in principle,[9] I will not pursue such an account here, since there is independent evidence that extraposition involves a non-extraction derivation.

Several works have argued that extraposition in fact involves base generation of the seemingly displaced non-argument constituent outside of the source phrase, to a position within the clause (see for instance Culicover and Rochemont 1990, 1997; Sheehan 2010; Reeve and Hicks 2017 and references therein). Under such an analysis, the term "source" would be a misnomer, since there is no syntactic dependency between the extraposed constituent and the phrase it semantically modifies. This analysis is illustrated in the simplified tree diagram in (30) below, in which we see that there is no movement relationship between the extraposed relative clause, and the DP *a book* which it semantically modifies:

(30)   *High base generation account of extraposition*

---
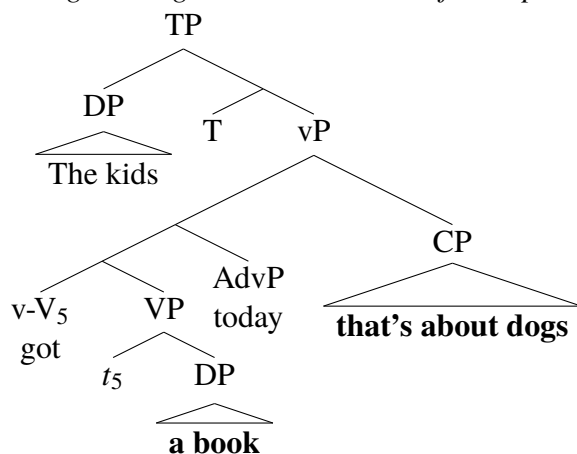
[9]An account like the following, for instance, is conceivable. The majority of the facts in section 2 involve extraposition of either an adjective or relative clause. Both of these elements have been argued to be type <e,t> constituents, which are interpreted by Predicate Modification under merger to NP, which has the same semantic type (Heim and Kratzer 1998). If there is no material separating an adjective or relative clause from the right edge of the DP, then under CL, nothing prevents rightward extraction of such items. In contrast, if other material does intervene between the adjective / relative clause and the edge of DP, it would be necessary for the adjective / relative clause to move around that material before exiting DP, assuming that DP is a phase. Such movement would plausibly require a step of movement within the NP itself, or alternatively, movement from the edge of NP to spec-DP. The first of these options is ruled out by the ban on edge-internal movement, discussed extensively in chapter 2 (Ko 2014, a.o.) while the second option is precluded by the version of anti-locality in Bošković (2005), among others, as discussed in chapter 4. Given these locality considerations, there is no way to successive cyclically move a constituent adjoined to NP around other material within the same DP, and thus extraposition would be necessarily non-successive-cyclic, as was implicit in (27) above. This hypothesis is analogous to the proposal in Bošković (2005) about the distribution of left branch extraction, which I discuss (and criticize) in the next chapter.

Another variety of base generation analysis posits that extraposition does involve movement, but not of the extraposed phrase: rather, the source phrase undergoes covert movement, and the extraposed phrase is externally merged to the head of that movement chain (Fox and Nissenbaum 1999; Nissenbaum 2000; Fox 2002; Johnson 2012; Overfelt 2015). This analysis depends on the concept of late external merge inside of moved constituents, a concept introduced by Lebeaux (1988). This analysis is illustrated in the simplified diagram in (31) below, where we see covert movement of *a book* (encoded by the crossing-out of the terminal nodes of its derived position), which facilitates overt external merge of a relative clause:

(31)   *Covert movement + late merge analysis of extraposition*
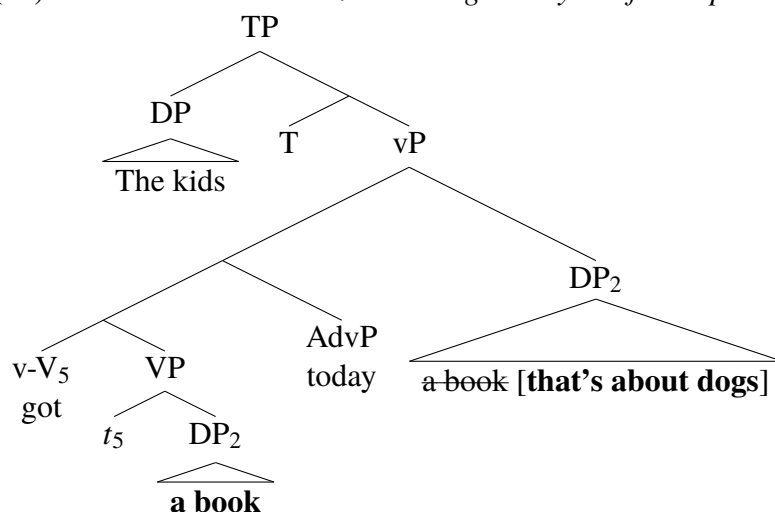
Under the original proposals of Fox and Nissenbaum (1999), extraposition is parasitic on covert movement that adjoins to the VP, though the relevant constituent would be the vP under contemporary assumptions, as the diagram in (31) reflects.

Let's pause to consider the difference between these two versions of the base generation analysis. For the pure base generation analysis in (30), there is no syntactic dependency between the extraposed constituent and its source. Under such an account of extraposition, it is unclear why the SOG should hold: why should the extraposed constituent be sensitive to its pre-displacement position within the source, if there is no syntactic relationship whatsoever between these two phrases? In contrast, under the covert movement plus late merge analysis in (31), there is, in fact, a direct syntactic relationship between the source and extraposed phrase. The latter analysis thus provides a concept to take advantage of in accounting for the word order relationship between an extraposed phrase and its source.[10]

---

[10]Also notice that the analysis in (31) captures the semantic relationship between the source and the extraposed phrase automatically, while the analysis in (30) requires something more to be

Next, I will summarize some evidence for the late merge analysis of extraposition, and overview previous proposals about extraposition and word order, before going on to provide a revised understanding of the SOG.

3.1 The relationship between late merge and extraposition

Lebeaux (1988) argued that adjuncts can be externally merged late, to phrases that have already been constructed and undergone movement, based on contrasts like that in (32) below. In (32a), a moved *wh*-phrase contains a relative clause with an R-expression co-indexed with the subject. This sentence is acceptable, in contrast to the one in (32b), which is the same as (32a) aside from the fact that the constituent containing the relevant R-expression is a complement rather than an adjunct:

(32)   a.   *R-expression in adjunct of moved phrase*
            ✓ [Which cakes [that **John**$_2$ ate]]$_1$ did **he**$_2$ find t_1 very tasty?

        b.   *R-expression in complement of moved phrase*
            */?? [Which rumor [that **John**$_2$ ate all the cakes]]$_1$ did **he**$_2$ deny t_1?

For both of these sentences, the subject c-commands the gap left by *wh*-movement. Thus there is a stage of the derivation at which the moved phrase, and thus the offending R-expression, was presumably c-commanded by the subject. If principle C must be satisfied at every stage of the derivation, both (32a) and (32b) should be unacceptable, though only the latter is in reality. For this reason, Lebeaux proposed that it is possible to externally merge an adjunct such as a relative clause to a DP that has already been constructed and undergone A′-movement. By merging the relative clause late in this way, the R-expression in the relative clause in (32a) was never c-commanded by the subject, and hence no principle C violation arises.

(33)   *How to avoid principle C*

        a.   *Step 1: Move*
            [Which cakes]$_1$ did he$_2$ find t_1 very tasty?

        b.   *Step 2: Merge adjunct to moved phrase*
            [Which cakes [that John$_2$ ate]]$_1$ did he$_2$ find t_1 very tasty?

In contrast, such a derivation is not possible for the complement in (32b), which must be merged during the initial construction of the containing DP. Thus a principle C violation in this example is unavoidable.

In Lebeaux's contexts, late merge targets an overtly moved phrase. As stated above, Fox and Nissenbaum (1999) proposed that late merge can also apply after

---

stipulated to ensure that the extraposed constituent semantically modifies the source DP.

covert movement (in particular Quantifier Raising), and that doing so derives adjunct extraposition in the way diagrammed in (31) above, as described next.

Fox & Nissenbaum gather several pieces of evidence for the late merge analysis of extraposition from binding, scope, and extraposition from coordinate structures. Nissenbaum (2000) and Fox (2002) provide further evidence for this analysis from the interaction of extraposition and antecedent contained deletion, and Overfelt (2015) goes on to show that this analysis makes correct predictions about the behavior of negative polarity items (NPIs) in extraposed relative clauses. I will not summarize all the evidence for this analysis here, but it will be convenient to demonstrate the evidence from principle C, since this will be referenced again later on.

Fox & Nissenbaum point out, citing Taraldsen (1981), that extraposition can circumvent violations of principle C in the way shown in (34) below. We also see here that failing to perform extraposition while holding everything else constant prevents principle C circumvention, as expected if this effect is indeed dependent on the application of extraposition here:

(34)   *Principle C avoidance by adjunct extraposition*
   a.   I gave him$_1$ a picture yesterday [**from John's$_1$ collection**].
   b.   ??/* I gave him$_1$ a picture [**from John's$_1$ collection**] yesterday.
   c.   I gave him$_1$ an argument yesterday [**that supports John's$_1$ theory**].
   d.   ??/* I gave him$_1$ an argument [**that supports John's$_1$ theory**] yesterday.
   e.   I told you that he$_1$ will accept the argument, when you and I last spoke, [**that I presented to John$_1$**] yesterday.
   f.   ??/* I told you when you and I last spoke that he$_1$ will accept the argument [**that I presented to John$_1$**] yesterday.
         (Fox & Nissenbaum 1999, ex. 11)

For Fox & Nissenbaum, in (34) movement of the source phrase, which happens to apply covertly, creates a position to which an adjunct can be late merged. This allows that adjunct to contain an R-expression which, due to never being c-commanded by the co-indexed phrase in question, never incurs a principle C violation. Under this analysis, example (32a) above and those in (34) are directly analogous, the only difference being the PF status of the movements that facilitate late merge of the adjunct. Extraposition derived by covert movement in the way just described is illustrated in the simplified tree diagram in (35) below, which represents (34c):[11]

---

[11]Similar evidence from principle C is available in Javanese and Wolof. In Javanese, while principle C applies without movement (ia), it is avoided under extraposition (ib):

(35)   *Principle C escape through extraposition via late merge*
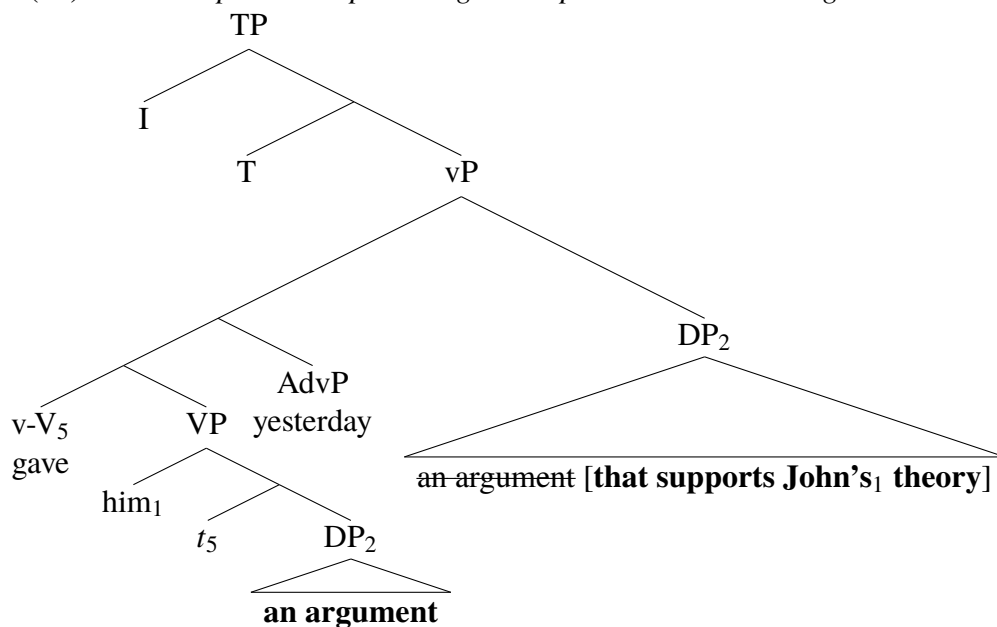
This will suffice as an introduction to the late merge analysis of extraposition. Next I will summarize some previous proposals about word order and extraposition, before detailing this chapter's analysis of the SOG.

(i)   a.   Dheweke weruh Budi.
           he   saw   Budi
           'He$_{1/*2}$ saw Budi$_2$'

      b.   Srabi   ne t_1 sing endi sing di tuku dheweke, [sing di senengi Budi]$_1$?
           pancake PRT   which one which PST buy   he   which PST like   Budi
           'Which pancake did he$_1$ buy that Budi$_1$ liked?'

Similarly, in Wolof, principle C normally holds (iia-b), but extraposition circumvents it (iic):

(ii)   a.   Gis-na muusu Ada.
           see-3SG cat   Ada
           'She$_{1/*2}$ saw Ada$_2$'s cat'

      b.   Jang-na [téére [bu xaritu Roxaya sopp lool]] démb.
           read-3SG book that friend Roxaya like very yesterday
           'She$_{1/*2}$ read a book that Roxaya$_2$'s friend really likes yesterday'

      c.   Jang-na [téére t_1] démb   [bu xaritu Roxaya sopp lool]$_1$.
           read-3SG book   yesterday that friend Roxaya like very
           'She$_{1/2}$ read a book yesterday that Roxaya$_2$'s friend really likes'

3.2 Previous proposals about extraposition and word order

Fox and Nissenbaum (1999) propose that extraposition goes to the right because covert movement, despite being phonologically vacuous, is linearly rightward. Under this view, the fact that covertly moved material goes to the right only becomes apparent when overt material is late merged within it, as in cases of extraposition. This hypothesis is already evident in the diagram in (35) above, for instance. This concept is carried over by Nissenbaum (2000), who additionally posits that covert material is in general visible for the computation of linear order—an assumption that I will attempt to dispense with in what follows. First, however, we must see what empirical results this assumption achieves.

As we saw in the introduction, Nissenbaum was likely the first to observe (based on English and Hebrew) the generalization that only rightmost elements within DP can extrapose. Nissenbaum argues that this generalization is predicted from the principle in (36) below, given rightward covert movement, in addition to the more general assumption just mentioned that covert material is visible for linearization.[12]

(36)   *Linear Edge Condition (LEC)*
        New material late merged inside of an already completed phase is possible
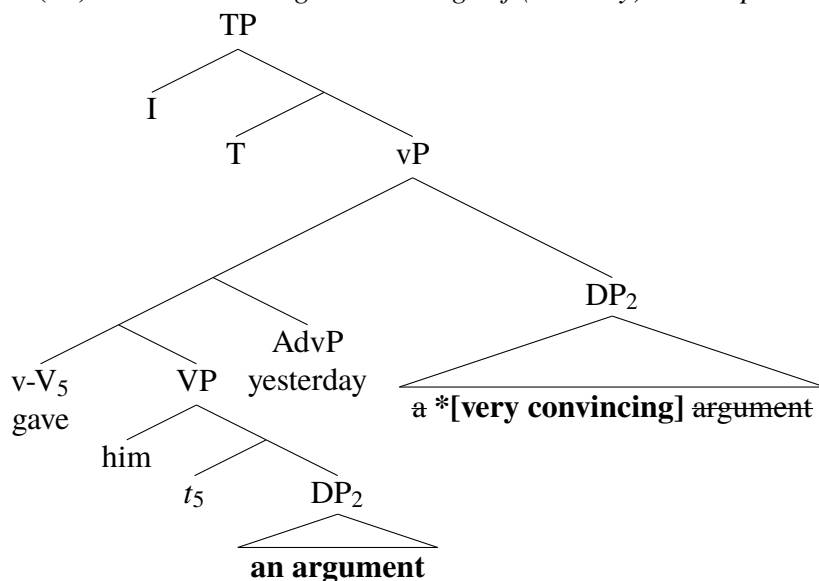   only at that phase's linear edge.
        (Paraphrased from Nissenbaum 2000, p. 201, ex. 22)

Notice that the late merged relative clause in (35) above satisfies the LEC, since it appears at the right edge of the covert instance of the DP *an argument*. Importantly in contrast, since English adjectives are typically merged to a position preceding N, adjective extraposition would involve late merge into the linear interior of this DP, as illustrated in (37) below. This violates the LEC, which therefore predicts the impossibility of extraposing a pre-nominal adjective. As we saw in section 2, extraposition of pre-nominal adjectives is indeed unacceptable, though post-nominal adjective extraposition is permitted, as expected.

---

[12]In (36) I have modified Nissenbaum's LEC for clarity of exposition, but (36) does the same work as his original formulation, which is stated in (i):

(i)   For any syntactic object SO accessed in an array, merge of new material is possible inside
        SO only at the linear edge.

(37)   *No late merge to non-edge of (covertly) moved phrase: Adjective*

```
                            TP
                     ┌──────┴────────┐
                     I               
                   ┌─┴──┐            
                   T    vP          
                      ┌──┴──────────────┐
                      │                  DP₂
                      │              ┌────┴──────────┐
                      │            AdvP              
                  ┌───┴──┐       yesterday          
                v-V₅    VP                          
               gave   ┌──┴────┐   a *[very convincing] argument
                     him      │
                        ┌─────┴──┐
                       t₅       DP₂
                              ┌──┴──┐
                            an argument
```

This understanding will successfully extend to predict the cross-linguistic SOG illustrated in the first section. However, it critically relies on the assumption that covert material is subject to linearization. This claim might be regarded as unusual, if linear ordering is only a property of the phonological representation, from which covert material is by definition absent. Therefore in what follows, I will offer a CL-based account for the SOG which eschews the linearization of covert material, and in doing so, I argue, extends to insights about several other facts about extraposition that we will examine later in this chapter.

The concept that extraposition is constrained by CL has precedent in Jenks (2011, 2013a,b), who notes building from Nissenbaum's original observations that rightward displacement appears to always be order preserving. Jenks does not provide much detail about what a CL account of extraposition would look like, however.[13] I will attempt to flesh out such an analysis here, while avoiding the proposal that covert material is linearized as just mentioned.[14]

---

[13]Jenks posits that rightward quantifier float in Thai, at least, is derived by movement of a full DP followed by distributed deletion (Faneslow and Ćavar 2002). Other works such as Sheehan (2010) have also suggested that some instances of extraposition are formed by distributed deletion. I will not pursue such an account here. However, note that extraposition by distributed deletion is still expected to obey the SOG, since the relative positions of overt terminals created by that operation will need to be order-preserving, as mentioned in the appendix to chapter 4.

[14]The following account owes a great deal to notes from a syntax seminar taught by Danny Fox and David Pesetsky in 2003. For a sketch of an account that assumes the ordering of covert movement in a different way than Nissenbaum (2000), see Fox and Pesetsky (2009), which relies on a number of assumptions that I attempt to leave aside here. The present goal is to see how much can be achieved

4   Extraposition and Cyclic Linearization

As mentioned in the discussion surrounding the diagram in (27) above, the SOG emerges in part from the concept that the source of extraposition is a phase: if it were not, order preservation would not be relevant. The instances of extraposition that this chapter focuses on have a DP as their source. Consequently, it is necessary here to state that DP is a phase (Heck and Zimmermann 2004; Bošković 2005, 2014, 2016; Newell 2008; Syed and Simpson 2017, a.o.).

To illustrate the derivation of extraposition I argue for, here I will consider extraposition from an object DP. The first thing to occur is the construction and spell-out of that DP. In (38) below, linearization within the object DP establishes the ordering information *a < dog*. The derivation of the vP then occurs. As stated above, I will assume that in order for extraposition to arise, the source DP covertly moves. In (38) such movement is shown as occurring to the left due to the constraints of the tree format, but I assume that such movement has no effect on linear order at the time when this vP spells-out and linearizes its contents. This is indicated by the crossing-out of the terminal nodes of the head of this covert movement chain.[15]

(38)   *Covert movement of DP within vP*
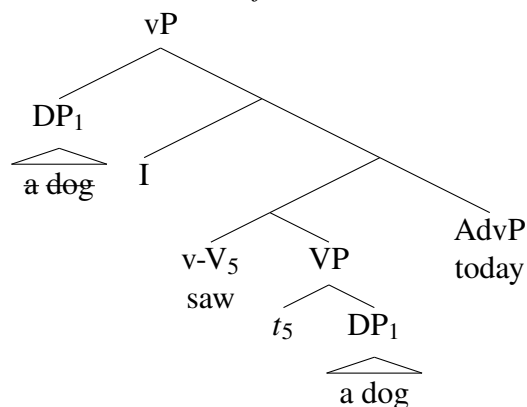
  1. Spell-out of DP: a < dog
  2. Spell-out of vP: I < saw < a < dog < today

For reasons that will become clear shortly, I hypothesize that extraposition is the result of late merging to a covertly moved DP which is in a segment of the clause (vP/CP) that has already undergone spell-out. I assume that when this occurs, in or-

---

without attributing ordering to covert material. This topic will come up again in chapter 8, however.

[15]I am assuming what has been termed the single-output model, in which the overtness or covertness of a given movement process is established at spell-out (Pesetsky 2000; Bobaljik 1995; Brody 1995; Groat and O'Neil 1996), as opposed to a model in which covert movements follow overt movements due to following the application of spell-out.

der for the late merged material to be incorporated into the linearization information established so far, the condition in (39) applies:

(39) *Local re-linearization condition*
      Late external merge into previously spelled-out material updates the linearization information of the minimal phase containing the late merge site.

As just described above, the diagram in (38) shows a vP that has already undergone spell-out, containing a covertly moved object DP. Since everything in (38) has already undergone linearization, late merging a relative clause to the head of the covert movement chain of this DP will require local re-linearization of that instance of DP. Since English relative clauses follow N, this re-linearization establishes that the late merged adjunct follows the N of the source phrase, as we see in (40):

(40) *Late merge and re-linearization of DP following (38)*
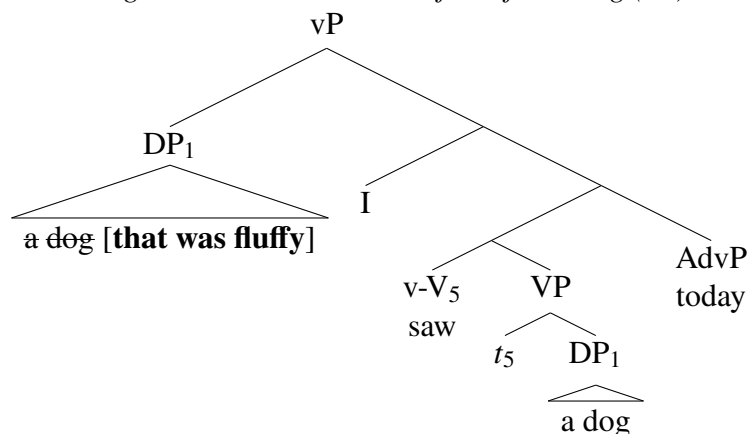
1. Spell-out of DP: a < dog
2. Spell-out of vP: I < saw < a < dog < today
3. Ordering of DP after late merge: a < dog < **that < was < fluffy**

Local re-linearization of DP triggered by late merge to its derived position only adds information about how the relative clause is ordered with respect to that DP. This process does not change the fact that this DP was, at the time when vP spelled-out, linearized as following V. The late merged relative clause, in contrast, was not present at the time when vP spelled-out and established where this DP is to be ordered. Consequently, in this context DP has been linearized with respect to the content of vP, and the late merged relative clause has been ordered with respect to DP, but importantly, this relative clause has not been linearized with respect to vP.

I argue that the linearization information in (40) is compatible with extraposition. Since the late merged relative clause was not ordered with respect to the content of

vP, nothing rules out its surfacing at the right edge of the vP. In such a situation the relative clause would also still surface rightward of the pronounced location of its source phrase, as the linearization information established in (40) requires (a < dog < that < was < fluffy). While I will shortly further constrain the orderings that are possible in such a situation, it will be useful to first be more specific about how the word order characteristic of extraposition is established in this context.

In (40), linearization has produced only a partial ordering: while every element is linearized with respect to something else here, some elements remain un-ordered with respect to each other. In particular, the late merged relative clause has not been ordered with respect to the verb, subject, or adverb in this vP. I hypothesize that the ambiguity inherent to a partial ordering is illegible at PF, and that therefore a partial ordering resulting from late merge is always resolved into an unambiguous total ordering.[16] Given order preservation, that total ordering will have to be consistent with the linearization instructions previously established in the derivation in question. As mentioned above, the linearization information established in (40) is consistent with the late merged relative clause being ordered at the right edge of the vP. I assume that when a partial ordering as in (40) produced by late merge is translated into a total ordering that treats the late merged material in this way, the result is extraposition, as (41) shows:

(41)   *Total order with extraposition derived from partial order after late merge*
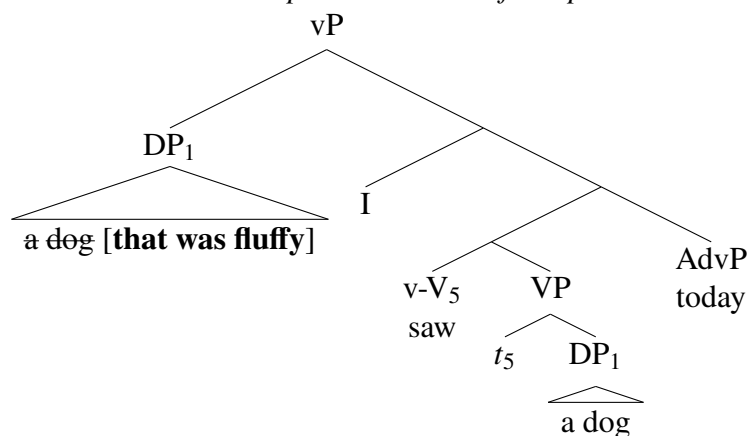
1. Spell-out of DP: a < dog
2. Spell-out of vP: I < saw < a < dog < today
3. Updated ordering of DP: a < dog < that < was < fluffy
4. **Total ordering for vP:** I < saw < a < <u>dog</u> < today < <u>that < was < fluffy</u>

With extraposition thus achieved, the derivation can continue on as usual.

---

[16]This is in contrast to a conceivable theory in which partial orders are licit, and simply result in optionality of word order.

A gap in the above exposition is that while selecting a total order yielding extraposition was indeed licit in the scenario just examined, nothing said so far prevents a different total order from potentiality being chosen. For instance, the linearization information in (40) would also be consistent with a total ordering in which the relative clause surfaces adjacent to the source DP, preceding the adverb. While I will address this issue shortly, what is important at the moment is that the late merge derivation in (40) allows extraposition to occur, as (41) shows. Furthermore, these concerns lead us directly to an explanation for the SOG, as I show next.

The extraposition derivation just provided will fail with late merge of a pre-nominal element, such as an adjective in a typical English DP. This is shown in (42) below, which minimally modifies (41) by including a late merged adjective *very fluffy* rather than a relative clause. After covert movement of DP and spell-out of the containing vP, late merge of the adjective and the resulting local re-linearization will establish that the adjective precedes the N of its source. If the total order established for (42) were to place this adjective at the right edge of the vP, there would be no ordering contradiction as far as the linearization of the vP itself is concerned, since the adjective was never ordered with respect to this vP. However, such a result would contradict the DP-internal ordering for this element established when it was late merged. Thus the only total order licit for (42) is one in which the late merged adjective surfaces internal to DP:[17]

---

[17]Notice that the linearization information in (42) dictates nothing about where the adjective is pronounced, beyond that it must precede N. While this ordering information is indeed inconsistent with an adjective being extraposed to the right, in a context where D is null in the source phrase, this configuration would in fact be consistent with the adjective being linearized at the left periphery of the sentence. This would be as we see in (i) below, which is unacceptable in English:

(i)   * [Very fluffy]$_1$ I saw [t_1 dogs].

This incorrect prediction will not emerge if null D is linearized in English. While we might justify this by saying that null D does indeed have a phonological exponent, which is $\varnothing$, it is difficult not to see this hypothesis as back-tracking towards allowing covert material to be visible to linearization. In chapter 8, I will discuss other evidence that this may be partially correct, however.

An alternative is that the rightwardness of extraposition is enforced by a default mechanism that causes late merged content to surface to the right of material that it was not directly linearized with respect to. In the case of late merge of a pre-nominal adjective this would necessarily yield an ordering contradiction, since the adjective would be ordered to the left of N before the application of this default mechanism. See Fox and Pesetsky (2009) and Johnson (2012) for further discussion about what might be responsible for making extraposition necessarily rightward (at least in the languages under consideration here).

(42)   *Extraposed order impossible for late merged adjective*
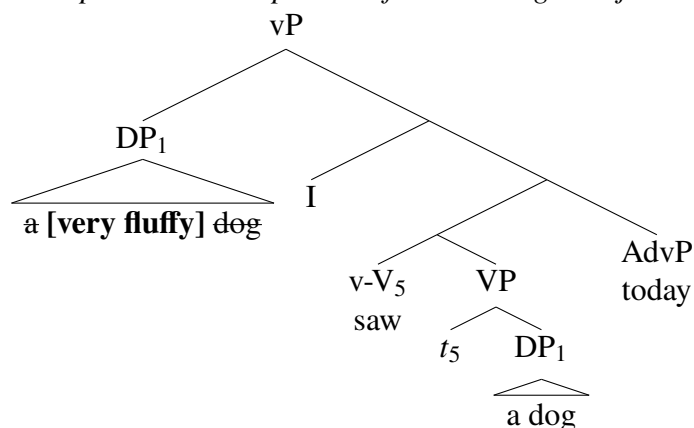
1. Spell-out of DP: a < dog
2. Spell-out of vP: I < saw < a < dog < today
3. Re-linearization of DP after late merge: a < very < fluffy < dog
4. Permitted total ordering: I < saw < a < very < fluffy < dog < today
Illicit total ordering: * I < saw < a < <u>dog</u> < today < <u>very</u> < <u>fluffy</u>

   This line of reasoning generalizes to predict the impossibility of extraposition for elements that do not merge at the right linear edge of the source phrase in question. This is as described by the SOG, which we've seen holds cross-linguistically based on the available data. Notice that this account of the SOG did not rely on the proposal that covert material and operations are subject to linearization. Later in this chapter, I will argue that this perspective has benefits for the understanding of several other facts about extraposition.[18] Before that, however, I will address a further puzzle, previewed above, about the positions that late merged material can surface in.

4.1   No string vacuous / non-peripheral extraposition

As described above, the linearization information established in (40) is compatible with the formation of a total order that places the late merged relative clause at the right edge of the vP, in the fashion of extraposition, as in (41). However, the fact that the relative clause in (40) was not directly linearized with respect to the vP at the time it was merged means that this derivation would also have been compatible with a total ordering that places the relative clause anywhere within the vP, provided that it follows the source DP. For instance, a final ordering which places the relative

---

[18]The account of extraposition provided here does not, in fact, predict that covert movement of the source phrase should be a necessary condition for extraposition. I argue that this issue is resolved by understanding late merge as being unable to apply too deeply, a concept with precedent in many works, as discussed in section 7.

clause to the left rather than the right of the co-occurring adverb should also be licit for (40), which would yield the seemingly un-extraposed word order in (43) below:

(43)   *Hypothetical alternative total ordering for (40)*
         I saw a dog [**that was fluffy**] today.

This is an incorrect result. To see why, examine once more the following facts from Fox and Nissenbaum (1999):

(44)   *Principle C avoidance by adjunct extraposition*
         a.   I gave him$_1$ a picture yesterday [**from John's$_1$ collection**].
         b. ??/* I gave him$_1$ a picture [**from John's$_1$ collection**] yesterday.
         c.   I gave him$_1$ an argument yesterday [**that supports John's$_1$ theory**].
         d. ??/* I gave him$_1$ an argument [**that supports John's$_1$ theory**] yesterday.
         e.   I told you that he$_1$ will accept the argument, when you and I last spoke,
                [**that I presented to John$_1$**] yesterday.
         f. ??/* I told you when you and I last spoke that he$_1$ will accept the argument
                [**that I presented to John$_1$**] yesterday.
                (Fox & Nissenbaum 1999, ex. 11)

If it is possible to late merge an adjunct to a covertly moved phrase but not pronounce it at the right edge of the containing vP, then there is no reason why principle C should necessarily apply to an un-extraposed adjunct, though the unacceptable variants of the examples in (44) show that it indeed does. What this teaches us is that late merged material not only can, but must, surface at the right edge of the structure it is merged into.

   I suggest that this fact can be understood as a manifestation of the linear condition on late merge proposed by Nissenbaum (2000), introduced in (36) above and repeated once again in (45) below:

(45)   *Linear Edge Condition (LEC)*
         New material late merged inside of an already completed phase is possible
         only at that phase's linear edge.
         (Paraphrased from Nissenbaum 2000, p. 201, ex. 22)

This formulation of the condition is appropriate in the context of Nissenbaum's proposals. However, this definition is not compatible with the derivation of extraposition that I have proposed above, for which late merge must be able to target non-peripheral positions in spelled-out phases (though the late merged material does ultimately surface at the periphery). For this reason, I revise the wording of Nissenbaum's condition as follows (46):

(46)   *Revised Linear Edge Condition (RLEC)*
        An element α late merged in a spelled-out phase XP cannot surface in the linear interior of XP, and thus must appear at the edge of or external to XP.

Given the RLEC, the representative extraposition derivation in (40-41) above cannot result in an un-extraposed order like that in (43), since this would involve the late merged relative clause surfacing in the linear interior of the containing vP.[19] [20]

The RLEC is in essence a descriptive generalization that should be derived from other principles. Nissenbaum (2000) provides evidence from a variety of patterns that a condition of this form is indeed necessary, though he does not explicitly work out the basis for it. For further discussion of the factors potentially responsible for the rightwardness of extraposition, see Fox and Pesetsky (2009) and Johnson (2012). Since the empirical results that this chapter is concerned with will emerge correctly under a variety of plausible justifications for the RLEC, I will not work out the motivation for this concept here.[21]

This concludes the core analysis of this chapter. The main point of this section has been to argue for an understanding of the SOG as an order preservation effect,

---

[19]This discussion about the proposed extraposition derivation in (40-41), and the undesirable hypothetical result for it in (43), has assumed that the adverb in question is present before late merge of the relative clause. However, another way of deriving (43) would be to late merge the adverb after late merge of the relative clause. It is not clear that late merge of an adjunct of vP is possible, however. If it were, late merge of an adjunct containing an R-expression co-indexed with the subject should be able to avoid principle C by targeting a fronted vP. This does not appear to be possible:

(i)   a.   * [Criticize Mary [**on Bill**$_2$**'s birthday**]]$_1$ he$_2$ never would t_1.
       b.   [Criticize Mary [**on his**$_2$ **birthday**]]$_1$ Bill$_2$ never would t_1.

This should be tested more exhaustively, however.

The puzzle posed by hypothetical string-vacuous extraposition is also noted by Fox (2002), who suggests that a parsing preference motivates against processing strings as extraposition that do not indicate its presence in a surface-evident way. While this hypothesis would if correct remove the need for the RLEC in this section, the RLEC will be useful in a few other places in this chapter.

[20]The RLEC independently rules out late merge of the adjective in the derivation represented by (42) above. This is because in this context the adjective, not being extraposable, must ultimately surface directly in the position to which it is merged, in between D and N. The adjective will of course incur no violation if it is merged early, however, though extraposition is precluded if this occurs.

[21]We might understand the RLEC as a condition that bans late merge from separating elements that linearization has already established are adjacent to each other. Note that such a condition would not emerge directly from CL, since CL deals with relative precedence, not direct adjacency. It is conceivable that linearization also establishes information about adjacency, but such adjacency would have to be violable (perhaps in Optimality Theoretic fashion), or else extraction of material from a previously spelled-out phase would never be possible. The notes from Danny Fox and David Pesetsky mentioned in footnote 3 provide a brief sketch of a system along these lines, though I will not pursue such an approach in this chapter.

and to account for it without relying on the concept that covert material and covert movement are linearized. In what follows, I will extend this understanding to several other facts about the distribution of extraposition.

5   Extraposition from overtly moved phrases

The previous section's account of the SOG was framed in terms of what could be called ascending extraposition, in which the extraposed phrase is in a structurally higher position than the source phrase (due to covert movement before late merger). While this variety of extraposition is likely the most frequently discussed in recent literature, it is not the only one. As we'll see in this section, it is also possible for the source of extraposition to be realized in a position at least as structurally high as the phrase extraposed from it. This occurs when the source of extraposition overtly moves. For instance, in (47) below, we see extraposition from phrases that have undergone overt A′-movement:

(47)   *Extraposition from moved wh-phrases*

    a.   **Which book** did you buy recently [**that you've been telling everyone about**]?

    b.   I know [[**which book**] you bought yesterday] [**that was very expensive**].

We can tell that the extraposed relative clauses in such examples are at a structural height parallel to that of their source. This can most easily be shown via principle C, as in (48) below. We see here that A′-movement over the subject allows an R-expression co-indexed with the subject to occur within a relative clause extraposed from the A′-moved phrase. This is consistent with the relative clause being late merged after movement of the source phrase to its terminal position, spec-CP, which is higher than the subject and thus allows principle C circumvention:

(48)   *A′-movement feeds principle C escaping extraposition*

    a.   * **He**$_1$ invited [several girls t_2] to the party [that **John**$_1$ dated in high school]$_2$.

    b.   [How many girls t_3]$_2$ did **he**$_1$ invite t_2 to the party [that **John**$_1$ dated in high school]$_3$?
       (Baltin 2006, from Culicover and Rochemont 1997)

Example (49) provides further examples of this effect:[22]

---

[22]This principle C circumvention effect likely rules out an analysis of extraposition from overtly

(49) *Extraposition from moved phrase escapes principle C*

a. [What books t_3]$_2$ did **she**$_1$ buy t_2 yesterday [that **Sarah**$_1$ has been thinking about for a very long time]$_3$?

b. I wonder [[how much of the cake t_3]$_2$ **he**$_1$'ll eat t_2 on Saturday, [that **John**$_1$ said he's been saving for the weekend]$_3$].

For Fox and Nissenbaum (1999) and Nissenbaum (2000), who focus on extraposition from covertly moved phrases, extraposition is rightward because the source phrase undergoes covert movement linearly to the right. Under their analysis, if the source phrase undergoes overt movement, then it should not be possible for extraposition to occur, contrary to fact.[23] In contrast, I argue that the analysis of extraposition provided in the previous section straightforwardly predicts this possibility.

Lebeaux's original evidence for late merge came from examples like (50) below, in which a relative clause adjoined to an overtly A′-moved phrase avoids a violation of principle C with respect to the subject:

(50) *No principle C violation for R-expression in adjunct of moved phrase*
[Which cakes [**that John**$_2$ **ate**]]$_1$ did he$_2$ find t_1 very tasty?

I will suggest that the difference between this example, and examples like (48b-49) above, is simply a matter of whether late merge to the moved DP precedes, or follows, the linearization of the containing CP.

Consider again a CP with an overtly A′-moved DP in its specifier. If an adjunct is late merged to the moved DP before the containing CP spells-out, then that adjunct will be ordered with respect to the content of CP, and will thus simply be linearized in the position where it was merged. This is what we see in examples like (50). In contrast, if CP spells-out before late merge applies, then the adjunct will only be linearized with respect to the DP that it merges to, but not with respect to the containing CP, following the local re-linearization condition stated in (39) above. This is illustrated in (51) below, which ignores the vP level of the derivation for convenience. Here we see that a relative clause late merged after the spell-out of the containing CP is not ordered with respect to the content of that CP. Therefore the relative clause can be assigned a surface order in which it appears at the right periphery of the CP, as (51) shows. This derivation is thus compatible with an acceptable total order instantiating extraposition from an A′-moved phrase:

moved phrases as involving stranding of the relative clause lower in the structure via extraction of the source. See Sheehan (2010) for an overview of arguments against stranding analyses of extraposition.

[23]This puzzle has been noted by both Fox (2002) and Baltin (2006). Baltin suggests that one might posit covert rightward movement that adjoins to CP and facilitates extraposition, before a second instance of overt A′-movement leftward. Baltin criticizes such an analysis as lacking independent evidence, and I agree with his assessment.

(51)   *Movement → spell-out of CP → late merge*
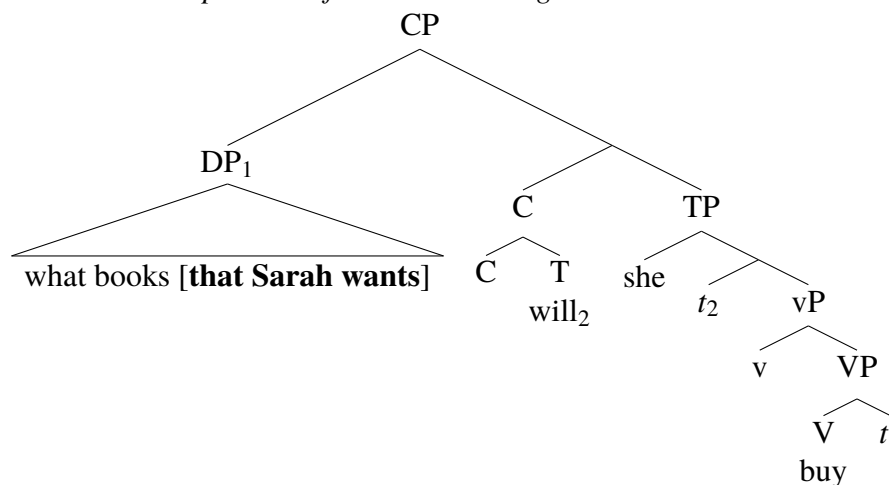
1. Spell-out of DP: what < books
2. Spell-out of CP: what < books < will < she < buy
3. Late merge to DP: what < books < that < Sarah < wants
4. Total order: ✓ <u>what < books</u> < will < she < buy < <u>that < Sarah < wants</u>

In languages like English, subjects A-move to spec-TP.[24] A derivation directly analogous to (51), differing only in involving late merge to an A-moved subject rather than an A′-moved phrase, will successfully achieve extraposition from a subject. A few sentences illustrating this possibility are shown in (52):

(52)   *Extraposition from subject*

    a.   **A book** appeared **which was written by Chomsky**.
       (Baltin 2006, ex. 2b)

    b.   **Someone** picked some books up which were lying on the table **who really didn't want to**.
       (Baltin 2006, ex. 17a)

    c.   **Those students** will pass this course **who complete all of their assignments on time**.
       (Baltin 2006, ex. 24)

Finally, extraposition from both A′-moved phrases and A-moved subjects cannot appear clause medially, as (53) and (54) below respectively show:

---

[24]I have generally not shown A-movement of subjects in the diagrams of this chapter, since this concept is largely irrelevant to the issues discussed here, but I do assume that such movement occurs.

(53)   *No clause-medial extraposition from A′-moved phrase*
  a. * **Which toy** did [**that they really like**] you buy them?
  b. * **Which toy** did you [**that they really like**] buy them?
  c. * **Which toy** did you buy [**that they really like**] them?
  d. ✓ **Which toy** did you buy them [**that they really like**]?

(54)   *No clause-medial extraposition from subject*
  a. * **A man** has just [**who I knew at school**] bought this shop.
    (Sheehan 2010, ex. 39)
  b. * **A man** has [**who I knew at school**] just bought this shop.
  c. * **A man** has just bought [**who I knew at school**] this shop.
  d. ✓ **A man** has just bought this shop [**who I knew at school**].

The impossibility of clause-medial extraposition has been observed in previous literature (Borsley 1997; Büring and Hartmann 1997), and is generally expected given the RLEC, so I will not comment further on this here.[25] [26]

---

[25]Given the RLEC, we predict that extraposed phrases should be able to be structurally clause-medial as long as they can surface in a peripheral position. This could be achieved by late merging a relative clause to a moved subject's base position in spec-vP. Evidence for such a derivation comes from the fact that it is possible for a relative clause extraposed from a subject to contain a negative polarity item (NPI) licensed by sentential negation (Linebarger 1980; Takahashi 2018):

(i) a. A doctor **won't** [make it in today [who knows **anything** about toenails]].
  b. * A doctor **won't** [[who knows **anything** about toenails] make it in today].
  c. A doctor **won't** [make it in today [who has **ever** done this procedure before]].
  d. * A doctor **won't** [[who has **ever** done this procedure before] make it in today].

Takahashi collects similar facts about the interaction of extraposition with ellipsis from Culicover and Rochemont (1990), and extraposition and scope from Reeve and Hicks (2017), that are consistent with the same conclusion. If this configuration is derived by late merge to a lower copy / trace of the moved subject, then we also expect to be able to late merge to an intermediate position created by A′-movement through spec-vP. To use NPIs as a diagnostic for such a configuration, it is safest to avoid contexts with potentially interrogative semantics, since this alone could serve to license the relevant NPI. In (ii) below I attempt extraposition of a relative clause containing the NPI *ever* from a topicalized phrase, which does not appear to be acceptable. I lack an explanation for this.

(ii) a. I **don't** want to eat cookies [that have **ever** been in contact with nuts].
  b. * [The cookies [that have **ever** been in contact with nuts]]$_1$, I **don't** want to eat t_1.
  c. * [The cookies t_2]$_1$, I **don't** want to eat t_1 [that have **ever** been in contact with nuts]$_2$.

[26]While extraposed constituents cannot occur medially within the same CP that their source inhabits, they can occur sentence-medially when they appear at the edge of an embedded clause which is itself nested within a higher clause (Sheehan 2010, ex. 40b):

(i) The fact [$_{CP}$ that [a man t_1] has bought this shop [**who I knew at school**]$_1$] is weird!

6   Extraposition, ellipsis, and covert movement

Here I discuss a finding of Takahashi and Ohtaka (2017) which I argue corroborates the proposal that covert movement is not linearized. Takahashi and Ohtaka examine the interaction of extraposition and ellipsis, and argue for the principle in (55):

(55)   *Anti-lacunarity*
    A host to which late merger applies must be pronounced when the late merged item is pronounced.
    (Takahashi and Ohtaka 2017, ex. 4)

Evidence for this conclusion comes from contrasts like that in (56) below. In (56a), we see that VP ellipsis can apply to the exclusion of a sentential adjunct, which serves as a landmark for detecting the presence of extraposition in the following example (56b). There is no ellipsis in (56b), but if VP ellipsis is applied in this context, as in (56c), the result is unacceptable:

(56)   *VP ellipsis vs. extraposition*
   a.   I found an article while browsing but you didn't ~~find an article~~ while searching.

   b.   Although I saw an article while browsing [which interested me], you didn't see **an article** while searching [**which interested you**].

Since in this situation extraposition is derived in the embedded clause before the construction of the matrix one, we do not expect the RLEC to be applicable. However, David Pesetsky (p.c) suggests that if the source of extraposition moves from an embedded clause, extraposition within that embedded clause is unacceptable, as illustrated in (ii). We also see here that in contrast, in this situation extraposition to the right edge of the higher clause containing the source is acceptable:

(ii)   [Which books t_2]$_1$ did Mary shout [that we should buy t_1 soon (*[**that Sarah wants**]$_2$)] as we were leaving ([**that Sarah wants**]$_2$)?

This fact is predicted by the theory of extraposition in section 4. Extraposition within the embedded clause will be derived by late merge of the relative clause to the source phrase, after spell-out of the local embedded CP. If the source moves on into a higher clause after this occurs, it will necessarily carry the late merged relative clause along with it. When the higher CP is linearized, the fact that the moved source DP contains the relative clause at that time means that the relative clause will be linearized within the higher clause, and thus determined to precede the embedded one. Since this same relative clause was previously linearized at the right edge of the embedded clause, the result is an ordering contradiction: the relative clause cannot both surface at the right edge of the embedded clause and within the higher one. Alternatively, if late merge does not occur until the moved phrase reaches its final landing site in the higher clause, then the result will simply be extraposition that surfaces at the right edge of the higher clause, which as we saw in (ii) is permitted.

    c.   * Although I saw an article while browsing [which interested me], you didn't ~~see~~ **an article** while searching [**which interested you**].
        (Adapted from Takahashi & Ohtaka ex. 8, citing Sheehan 2010)

The unacceptability of (56c) instantiates the principle in (55), since ellipsis here deletes the source of extraposition. Example (56b) is especially important here because it establishes that extraposition is capable of targeting a position above the region to which VP ellipsis applies. The fact that ellipsis of the source in (56c) is unacceptable despite the fact that extraposition can target a position beyond the ellipsis site suggests an explanation along the lines of (55). This is as opposed to an explanation in which extraposition cannot co-occur with VP ellipsis due to being adjoined within a domain that ellipsis applies to, as some have argued (Asakawa 1979; Guéron 1980; Baltin 2017).

    I argue that the theory of extraposition proposed in this chapter facilitates a deduction of the constraint in (55). To review, I have argued here that extraposition occurs when late merge to a given moved DP occurs after spell-out of the containing vP or CP. In particular, I argued that after such late merge occurs, the resulting partial order is used as a basis to derive a total order in which, due to the RLEC, the late merged material (if SOG-obeying) is ordered at the periphery of the structure built so far. This is shown once more in the partial derivation in (57) below:

(57)   *Total order with extraposition derived from partial order after late merge*
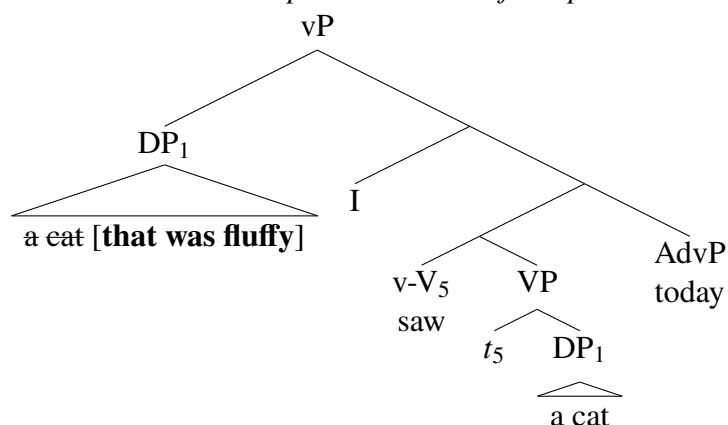
1. Spell-out of DP: a < cat
2. Spell-out of vP: I < saw < a < cat < today
3. Ordering of DP after late merge: a < cat < that < was < fluffy
4. **Total ordering for vP:** I < saw < a < <u>cat</u> < today < <u>that < was < fluffy</u>

In the first three linearization steps in this derivation, the source DP is ordered with respect to the rest of the vP, and the late merged relative clause is ordered with respect to that DP. I hypothesize that in this scenario, the fact that the source DP has been

linearized internal to the vP allows that DP to serve as an anchor for establishing a total order in which the late merged relative clause is also unambiguously ordered with respect to the vP's contents. Correspondingly, I propose that if the phrase that is the source of extraposition is silenced by ellipsis, the phrase that is late merged in the course of deriving extraposition lacks an anchor that would allow a total ordering that includes it to be established.

To see this concretely, consider example (56c) above, which illustrates the restriction under discussion. The matrix vP in which ellipsis takes place in (56c) is diagrammed in (58) below. Assuming that ellipsis is the product of a feature that instructs spell-out to ignore a given element (Merchant 2001, 2004; Aelbrecht 2010), here covert movement of the source phrase to facilitate extraposition is followed by V and the base position of the source phrase failing to be linearized when the vP is spelled-out, as illustrated by crossing out in (58). Since the source DP is thus absent from the linearization information of this derivation, there is no element that is ordered with respect to both the containing vP, as well as the late merged relative clause. I propose that for this reason, it is not possible to translate the ordering information in this derivation into a total order which assigns the late merged relative clause a position relative to the vP.

(58)   *Total order indeterminate after ellipsis of the source*
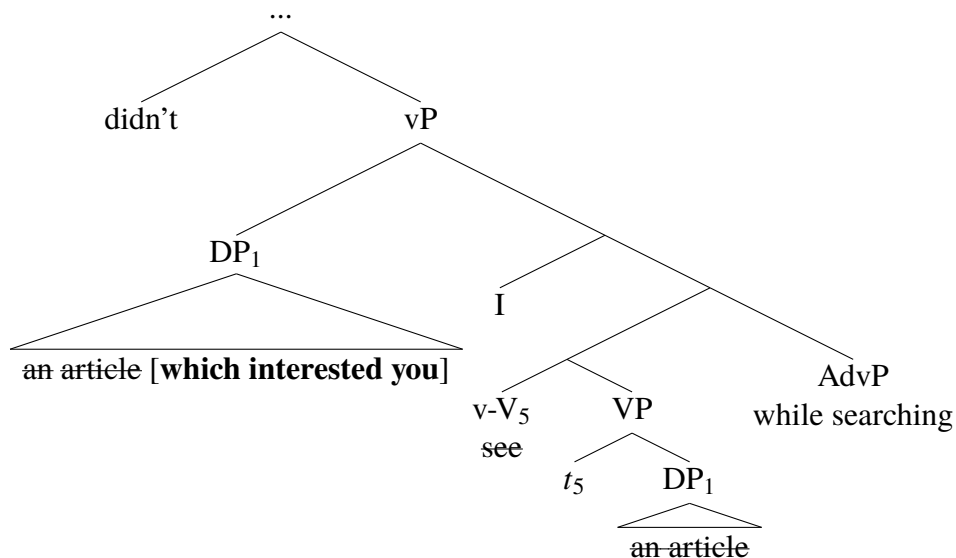
1. Spell-out of DP: an < article
2. Spell-out of vP: I < while < searching (verb and base position of source DP un-linearized due to ellipsis)
3. Ordering in DP after late merge: an < article < which < interested < you
4. **Total ordering for vP:** ???

I hypothesize that in the absence of a total ordering, the linearization of the derivation remains incomplete, causing a PF crash. This concept makes sense of Takahashi and Ohtaka's observation that extraposition is blocked by ellipsis of its source (55).

Importantly in contrast, if covert movement were inherently ordered to the right as Fox and Nissenbaum (1999) originally suggested, then an extraposed constituent should be granted a linear position regardless of whether or not its host is actually a part of the PF representation or not, since it inhabits a constituent that is itself ordered with respect to the containing clause.

These concepts predict that if ellipsis does not entirely silence the content of the source DP, extraposition should be licit. This prediction is verified by the fact that NP ellipsis in the source is compatible with extraposition, as (59) shows:

(59) *NP ellipsis within source does not block extraposition*
      We were each supposed to find three interesting articles. Although I found three articles while browsing [which interested me], you didn't find three ~~articles~~ while searching [**which interested you**].

Since NP ellipsis in English always leaves behind some overt remnant of the DP in which it applies, NP ellipsis within a source of extraposition does not entirely remove it from the PF representation, and thus, some material remains that can serve as an anchor for the ordering of the extraposed material. These considerations also predict that VP ellipsis should be fully compatible with extraposition, provided that the source phrase moves out of the ellipsis site. I argue that this is indeed so:

(60) *Extraposition permitted by movement of source from ellipsis site*
      I know what ingredients John will buy at the store that he needs to make cupcakes. However, I don't know **what ingredients** Mary will ~~buy at the store~~ **that she needs to make muffins**.

6.1 The relative order of extraposed material and its source

In the context of CL, we expect the linear order established between extraposed material and its source to have to be preserved by any subsequent movement processes. This concept leads to correct predictions about certain interactions between VP fronting and extraposition observed in previous literature.

Consider the Dutch (61) below. Example (61a) establishes the possibility of adjunct extraposition in Dutch. In example (61b), we see that a fronted VP that contains extraposed material, but not the source phrase, is unacceptable. In (61c), by contrast, the fronted VP carries along both the source phrase and the extraposed constituent, which are separated by the V of this moved VP. As Sheehan (2010) discusses, the same pattern has been observed for a number of Germanic languages.

(61)   *VP fronting cannot reverse source and extraposed phrase: Dutch*

        a.   Ik heb [**een man** t_1] gezien [**uit India**]$_1$
             I have a   man   seen   from India

        b.   * [$_{VP}$ Gezien [**uit India**$_1$]]$_2$ heb ik [**een man** t_1] t_2
                seen   from India   have I a   man

        c.   [$_{VP}$ **Een man** t_1 gezien [**uit India**]$_1$]$_2$ heb ik niet t_2
              a   man   seen from India   have I not
            (Sheehan 2010 ex. 44, from Koster 2000)

We see in (61a) that the extraposed PP ends up to the right of its source under typical circumstances. In (61b), the fact that VP fronting carries the extraposed phrase to a position preceding the overt position of its source is expected to give rise to an ordering contradiction, since this unacceptable derivation reverses the order of these constituents that extraposition established. In contrast, as (61c) shows, by performing VP fronting that carries along both the source and the extraposed phrase, the relative order of these two constituents is preserved, and no ordering contradiction arises.

     An analogous fact of another variety is shown in the English (62) below. Here we see that VP fronting that strands an extraposed constituent is acceptable (62a), while including the extraposed phrase with VP-fronting is unacceptable (62b). Since the latter of these configurations reverses the relative order of the source and of the extraposed material, we indeed expect an ordering contradiction to arise here:

(62)   *VP fronting cannot reverse source and extraposed phrase: English*

        a.   They said that a man would come in, and [come in]$_1$ **a man** did t_1 **who had lived in Boston**.

        b.   * They said that a man would come in who had lived in Boston, and [come in **who had lived in Boston**] **a man** did.
            (Culicover and Rochemont 1990, ex. 36-37)

     To conclude this section, here I have argued that the finding of Takahashi and Ohtaka (2017) that extraposition is blocked by ellipsis of its source corroborates the proposal that covert movement is un-ordered: if it were ordered, it should be possible to appropriately linearize extraposition regardless of whether its source is pronounced or not. The fact that an extraposed phrase is ordered (at least) with respect to its source also accounts for the unacceptability of certain configurations combining extraposition and VP fronting.

7 Limiting the depth of late merge

In this section, I discuss a final topic about the relationship between extraposition and late merge. As discussed above, Fox and Nissenbaum (1999), as well as several following works, argue that extraposition is parasitic on (covert) movement of the source phrase. In addition to the evidence shown above, another correct prediction of this account is that the source of extraposition will scope at least as high as the extraposed phrase (as independently shown by Williams (1974) for comparative and result phrase extraposition). This is exemplified by (63) below. Here extraposition from an indefinite DP forces it to have high scope, here resulting in a reading where there is a specific picture that is being looked for:

(63)   *Extraposition from DP forces high scope*
        (Fox & Nissenbaum 1999, ex. 18a/19a)

    a.   I looked for a picture t_1 very intensely [from John's factory]$_1$.
          ($\exists$> look for, *look for > $\exists$)
    b.   I looked for a picture t_1 very intensely [by this artist]$_1$.
   ($\exists$> look for, *look for > $\exists$)

See Fox and Nissenbaum (1999); Nissenbaum (2000); Fox (2002); Overfelt (2015) and references therein for discussion of further facts which show that extraposition involves (covert) movement of the source phrase. Importantly, the theory of extraposition I provided in section 4 does not, by itself, predict that movement of the source will be required for extraposition to occur. I discuss this issue here.

In this chapter I have assumed, as Nissenbaum (2000) does, that late merge can apply within previously constructed phases. This concept is natural under CL, for which unlike Chomsky's proposals about phases, spell-out does not make the contents of a phase inaccessible to later operations. However, this proposal opens up the possibility of deriving extraposition without moving the source phrase. To see why this is predicted, consider (64) below. In (64a), we see a vP in which an object DP has been merged, but not moved in any way. Both DP and vP have undergone spell-out here. Next, in (64b), we see late merger of a relative clause to this DP. The updated linearization of the object DP establishes that the relative clause follows its contents, but since such re-linearization is local as extensively discussed above, the relative clause is not ordered with respect to the vP. Thus the linearization information established in this derivation is consistent with the relative clause ultimately being assigned a surface position at the right periphery of the vP, despite no movement occurring here, as (64b) shows:

(64)   *Hypothetical extraposition derivation without movement*

    a.   *Construction of vP, no covert movement*
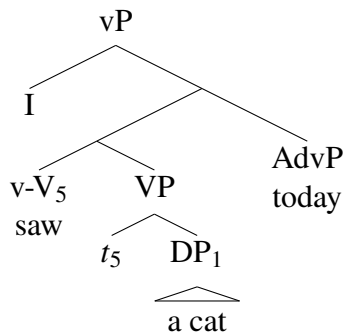
        1. Linearization of DP: a < cat
        2. Linearization of vP: I < saw < a < cat < today

    b.   *Late merge to un-moved object DP and potential extraposition*
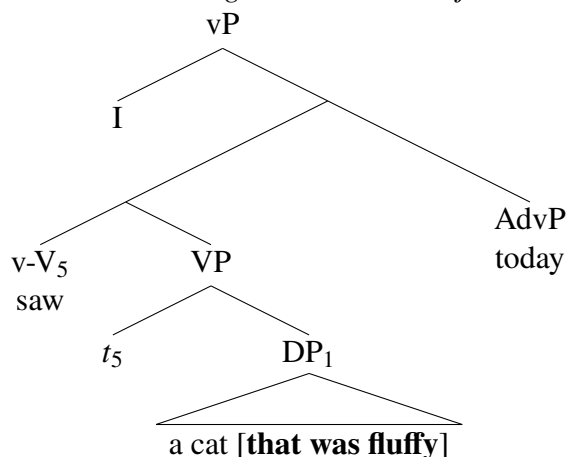
        1. Linearization of DP: a < cat
        2. Linearization of vP: I < saw < a < cat < today
        3. Late merge and re-linearization of DP: a < cat < that < was < fluffy
        4. Possible final order: I < saw < <u>a < cat</u> < today < <u>that < was < fluffy</u>

If Fox and Nissenbaum (1999) and following works are right that extraposition is indeed contingent upon movement of the source phrase, then a movement-less extraposition derivation like that in (64) must be ruled out.

    This problematic derivation will be appropriately banned if there is a principle that prevents late merge that is, in some relevant sense, too deep within a previously spelled-out constituent. For instance, if late merge within the complement but not the edge of a spelled-out vP is banned, then late merge to the object DP in a derivation like (64) will be impossible. In contrast, late merge to a DP that has moved (either overtly or covertly) to the vP edge will be licit, and thus extraposition will always

co-occur with movement of the source.

The idea that late merge cannot apply too deeply has been argued for in a variety of works, such as Tada (1993); Sauerland (1998); Stepanov (2001); Stanton (2016) and Safir (2018). A clear manifestation of this restriction is discussed in Sauerland (1998), in terms of examples like (65) below.

(65) *No late merge to embedded positions: Evidence from stacked modifiers*

a. [Which computer [compatible with his$_1$] [that Mary$_2$ knew how to use]]$_3$ did she$_2$ tell [every boy]$_1$ to buy t_3?

b. * [Which computer [compatible with Mary$_2$'s] [that he$_1$ knew how to use]]$_3$ did she$_2$ tell [every boy]$_1$ to buy t_3?

In these examples, the moved *wh*-phrase contains two modifiers: one bearing a pronoun *him* bound by the lower quantifier phrase *every boy*, and another containing an R-expression *Mary* co-referent with the subject *she*. In the acceptable example (65a), the pronoun is contained in the inner modifier, while the R-expression is contained in the outer one. In this configuration, it is possible to have first merged the inner modifier in order to satisfy the c-command condition on variable binding, and then late merge the outer modifier after *wh*-movement, thus avoiding a principle C violation. In contrast, in the unacceptable (65b), the modifier containing the R-expression is the inner one, while the modifier containing the pronoun is the outer one. This example would be acceptable if it were possible to merge the outer modifier before *wh*-movement, and the inner one after *wh*-movement. However, if late merge to an embedded position is banned, then it will not be possible for the inner modifier to be late merged in the way required in (65b). As expected given this hypothesis, (65b) is indeed unacceptable.

The works cited above vary in how they constrain the depth of late merge. However, all share the intuition that counter-cyclic merge is indeed limited in how counter-cyclic it can be, generally targeting a position as close as possible to the root node of a given structure. If deep late merge is indeed generally banned, then late merge to and thus extraposition from a phrase that has not moved to the edge of the structure built so far will largely be ruled out, as desired.[27]

---

[27]This concept also suggests an explanation for the fact that in contexts with extraposition from multiple phrases, extraposition from the structurally higher phrase must surface outside of extraposition from the structurally lower one (Bianchi 2000; de Vries 2002; Sheehan 2010):

(i) (Sheehan 2010, from Bianchi 2000:137)

a. ? **A man** entered the room last night [that I had just finished painting] [**who had blond hair**].

b. * **A man** entered the room last night [**who had blond hair**] [that I had just finished painting].

Several proposals about the constraints on late merge are attested in the literature. Therefore establishing a single empirically adequate definition of the relevant limitation will be an important task for future work.[28] Since this is an issue that goes beyond the topic of extraposition, I will not explore it further here.

8 Conclusion

The primary goal of this chapter has been to gather the cross-linguistic empirical evidence for, and argue for an explanation of, the following generalization:

(66)   *Source ordering generalization (SOG)*
         Extraposition of a given element α from a source phrase β must result in α having the same relative linear order with respect to the rest of the content of β that it would have had in the absence of extraposition.

In particular, I have argued for an account of the word order characteristics of extraposition that importantly differs from proposals in Fox and Nissenbaum (1999) and Nissenbaum (2000) in not attributing linear order to covert material and covert movement. I argued that this perspective on extraposition extends naturally to an account of a number of additional facts about this phenomenon, such as the possibility of extraposition from moved phrases, and the fact observed by Takahashi and Ohtaka (2017) that extraposition is blocked by ellipsis of its source. I went on to argue that the fact that extraposition depends on movement of the source, though not an automatic consequence of the theory developed here, does emerge from the independently attested concept that the depth of late merge is limited.

If deep late merge is not permitted, then late merge to the structurally lower phrase will presumably need to apply first, yielding the inner instance of extraposition, before the same can occur to the structurally higher phrase, then yielding the outer instance of extraposition. The RLEC will require a second extraposed phrase to surface to the right of an instance of extraposition that was formed earlier, and consequently multiple extraposition can derive a sentence like (ia), but not like (ib).

[28]See further Fox (2017), who discusses facts potentially involving deep late merge to an embedded position within a moved DP, which appear to contradict the constraint indicated by Sauerland's examples in (65) above.

Chapter 7

A digression about parasitic gaps in relative clauses

1   Introduction

Throughout this dissertation, we have seen a wide variety of situations in which linearization and locality constrain extraction and discontinuity. Successive cyclicity was prominent among these topics, which in chapter 5, I investigated using parasitic gaps (PGs) as a means of diagnosing the properties of derivations with overlapping movement paths. In this chapter, I set aside CL and consider several of the concepts from chapter 5 in terms of a phenomenon which has received little study in previous research—PG licensing in relative clauses by extraction from the same NP, as in (1):

(1)   *PG in relative clause licensed by extraction from NP*
    a.   Who$_1$ did Mary take [pictures of t_1 [that weren't that flattering to PG$_1$]]? ([Citko 2014](), ex. 105)
    b.   That's the manager which$_2$ I know [an employee of t_2 [who's got a very intense grudge against PG$_2$]].
    c.   Let me tell you who$_3$ I've noticed [an aspect of t_3 [that makes me want to avoid PG$_3$]].
    d.   [This person]$_4$, I painted [a portrait of t_4 [that unfortunately didn't satisfy PG$_4$]].

Very little previous research has examined this configuration. [Citko](2014), discussed shortly below, reports example (1a) as being relatively acceptable, though the other examples in (1) are mine.[1] I have found that many (though not all) speakers

---

[1]The only other work I know to have mentioned such configurations is [Matushansky](2005),

do indeed accept such examples, which I will dedicate this chapter to examining.

Recall that under the theory of PGs in Nissenbaum (2000) that I adopted in chapter 5, a PG is licensed when the constituent that contains it adjoins to a position targeted by (successive cyclic) A′-movement. In this configuration, Predicate Abstraction is triggered both at the landing site of the licensing phrase's A′-movement, and at the edge of the constituent containing the PG, allowing the two to be combined via Predicate Modification. The schema in (2) below shows this scenario playing out at the vP level. Here the intermediate trace of A′-movement through the vP saturates the function created by Predicate Modification, and in doing so, binds the variable corresponding to the 'true' gap as well as the PG:

(2)   *PG licensing by successive cyclic movement from vP*
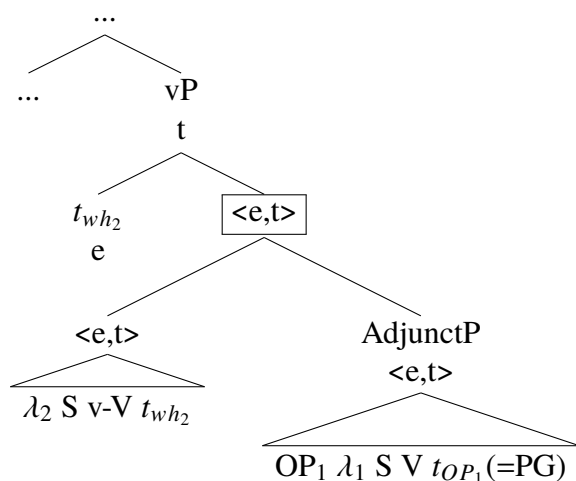
While for Nissenbaum and for chapter 5 of this work PG-licensing by successive

---

who reports the following judgment:

(i)   *Who$_1$ did Mary tell a story about t_1 [that really impressed PG$_1$]]?
       (Matushansky 2005, p. 168)

Citko (2014) cites this example and offers (1a) as an alternative, but does not explain what might distinguish the two. In general, I have found that these examples are more acceptable when the relative clause is phonologically larger. For instance, (ii) below, in my judgment, improves on (i):

(ii)   Who$_1$ did Mary tell a story about t_1 [that genuinely and absolutely impressed PG$_1$]]?

I suggest that what makes (ii) superior to (i) is that by including more phonological material between the gap of relativization and the PG, they become easier to distinguish, making it less taxing to determine their respective referents. Examples of this form are certainly somewhat marked, which I suspect simply stems from the fact that they are inherently multi-gap structures, and thus naturally have a relatively high processing burden.

cyclic movement from vP was most relevant, we do not expect this method of PG licensing to be exclusive to the vP. That is, the same mechanisms will allow any landing site of A′-movement to license a PG, provided that there is a possible PG-bearing constituent available to merge to that position. Legate (2003) uses this reasoning to diagnose successive cyclicity in various verbal contexts, and similarly, I will argue that these concepts lead us to conclude that the sentences in (1) reveal the possibility of successive cyclic movement through NP.

Relative clauses, at least restrictive ones like those in (1), are typically analyzed as adjuncts of NP (and not DP). Therefore given the theory of PGs adopted here, the sentences in (1) reveal that successive cyclic movement through the edge of NP is available. If it were not, there should be no way for the PGs in (1) to be licensed. The licensing configuration in question is shown in (3) below:

(3)   *Successive cyclic extraction from NP and a PG in a relative clause*
       $[_{CP}$ WH_4 ... V $[_{DP}$ D $[_{NP}$ t_4 $[_{N''}$ $[_{N'}$ N t_4] $[_{RC}$ ... PG_4]]]]]

As mentioned in chapters 3 and 4, there is precedent for the proposal that NP is a phase—a concept that entails the possibility of successive cyclic extraction from this constituent. To recapitulate, several works on morpho-phonology argue that lexical projections like NP, VP, and AdjP actually consist of category-less lexical roots selected by phasal categorizing heads, respectively n^0, v^0, and a^0 (Marvin 2003; Newell 2008; Embick and Marantz 2008; Embick 2010, a.o.). Such works argue that the phasal status of categorizing heads is revealed by the way that category changing morphemes tend to set boundaries for the application of various morpho-phonological processes like stress assignment and contextual allomorphy.[2] If for this reason NP (or more precisely nP) is a phase, extraction from it must be successive cyclic, thus opening up the possibility of configurations like (3), and the corresponding sentences in (1). In this chapter I will argue in favor of this analysis, and explore a variety of consequences and extensions of it.

1.1   Chapter contents

In section 2, I discuss in detail the justification for the analysis in (3), and establish a number of other basic properties about the sentences in (1). In section 3, I examine an asymmetry for PGs in stacked relative clauses, which mirrors an analogous observation from Nissenbaum (2000) about stacked vP adjuncts. In section 4, I argue that the possible interpretations for sentences like those in (1) corroborate another proposal from Nissenbaum, that overlapping operator movements must

---

[2]See also Bayırlı (2017), who argues that the phasehood of NP affects the distribution of concord.

form crossing paths. In section 5, I identify further evidence for successive cyclic movement from NP, and consider the implications of this finding for the theory that D blocks certain forms of sub-extraction (Bošković 2005, a.o.). In section 6, I consider a remaining puzzle about another variety of PG-containing relative clause, followed by the conclusion. The appendix offers some final thoughts on the distribution of PGs, and in particular the possibility of PG licensing in CP.

2   The syntax and semantics of PGs in relative clauses

Citko (2014), in her brief consideration of the configuration in (1), suggests that sentences of this form constitute potential evidence that extraction from DP passes through its edge. The analysis she outlines relies on the assumption that relative clauses can be adjoined in the projection of D, allowing successive cyclic movement via spec-DP to license a PG in a relative clause. This is illustrated in (4) below, which ignores the vP for convenience:

(4)   *Citko's analysis: PG-bearing relative clause adjoined in DP*
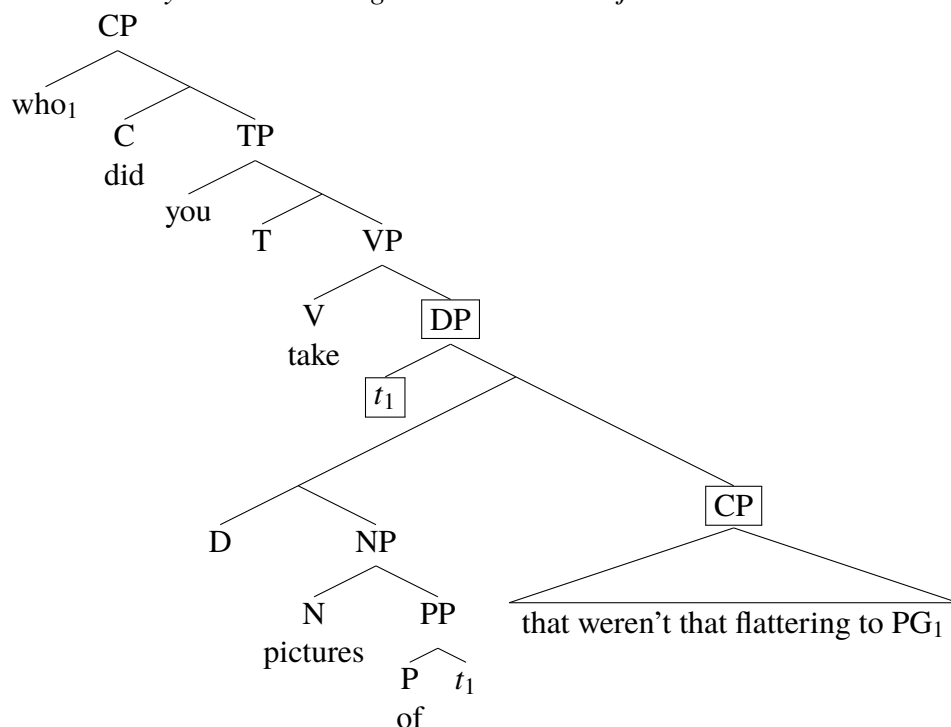
However, restrictive relative clauses like those in (1) have been independently argued to be adjuncts of NP, and thus merged below D (see Heim and Kratzer 1998, p. 88). Under this analysis, NPs are predicates of individuals <e,t>, while relative

clauses are derived predicates of the same type due to movement of a *wh*-operator / relative pronoun within them, enabling the latter to adjoin to the former and be interpreted via Predicate Modification:

(5)   *Structure and interpretation of an NP with a restrictive relative clause*

      a.

      b. $\{x \mid x \text{ is cat}\}_{[\![N']\!]} \cap \{x \mid x \text{ has orange fur}\}_{[\![CP]\!]}$
           $= \{x \mid x \text{ is cat and } x \text{ has orange fur}\}_{[\![NP]\!]}$
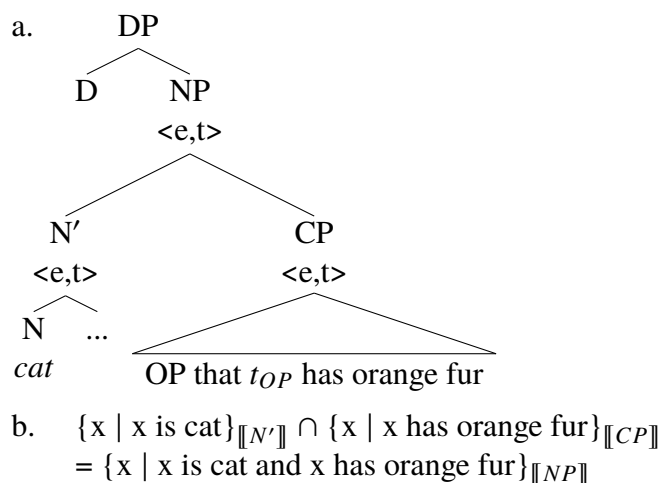
If relative clauses are sisters of NP, then what PG licensing in a relative clause by extraction from NP actually indicates is the possibility of successive cyclic A′-movement through the NP edge, as in (6):

(6)   *This chapter's analysis: PG-bearing relative clause adjoined in NP*
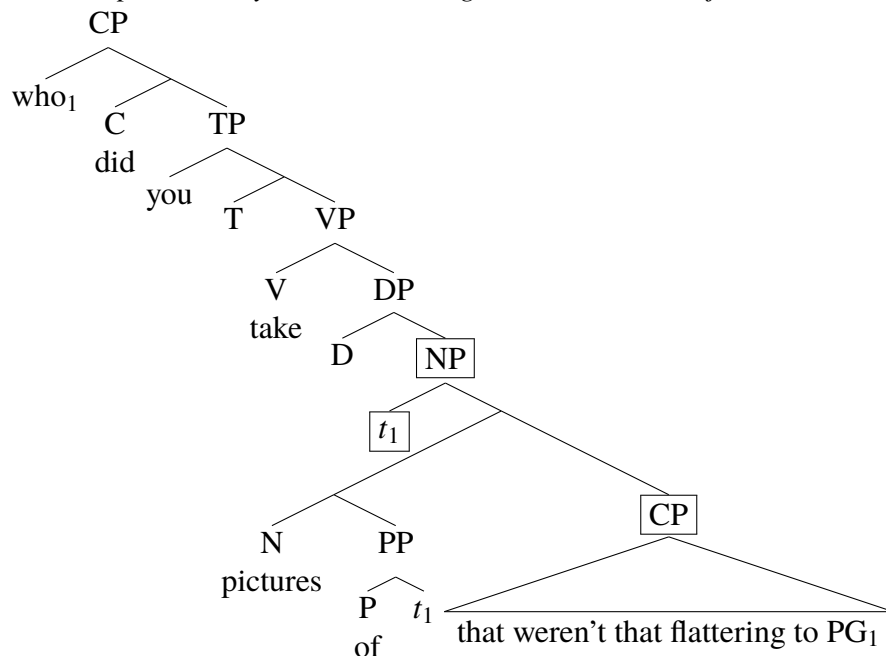
This analysis does not preclude the possibility of movement via the DP edge as well, which will be obligatory if DP is a phase, as assumed in the previous chapter. I consider a consequence of positing successive cyclic movement through both NP and DP in section 5 of this chapter (which, most significantly, challenges the theory of sub-extraction in Bošković (2005)). Until then, however, only movement through the NP edge must be kept in mind. Next, I will establish the basic properties of PGs in relative clauses, before proceeding to a more detailed semantic analysis.

We can confirm that the PGs in question are indeed parasitic on movement from the containing NP by observing that, in the absence of such movement, these gaps become unacceptable:

(7) *Extraction from NP required for PG in relative clause*
    a.   I painted [a portrait of Mary$_4$ [that unfortunately didn't satisfy **her$_4$/*PG$_4$**]].
    b.   I've noticed [an aspect of John$_3$ [that makes me want to avoid **him$_3$/*PG$_3$**]].
    c.   Mary took [pictures of [a person]$_1$ [that weren't that flattering to **them$_1$/*PG$_1$**]].

Given the discussion in chapter 5, we expect PGs in relative clauses to fundamentally be gaps left behind by movement of a null operator to the edge of the containing relative clauses, as for PGs in islands of other varieties. If this is the case, adding another island boundary between the PG and the edge of the relative clause should cause unacceptability. Since these contexts presumably also involve movement of a relativizing operator as well, ensuring that the island test singles out the PG-forming operator requires merging the relevant island below the origination position of the relativizing movement chain. This is done in (8) below, where the relativizing movement chain is distinguished with the marking "REL". The sentences in (8) are unacceptable, as the null operator theory of PGs predicts:

(8) *Additional island in relative clause prevents PG licensing*
    a.   *Relative clause island*
        * Who$_1$ did Mary take [pictures of t_1 [$\varnothing_{REL}$ that t_{REL} were hilarious to everyone **[who has met PG$_1$ before]**]]?
    b.   *Subject island*
        * That's the manager which$_2$ I know [an employee of t_2 [who$_{REL}$ thinks t_{REL} that **[every message from PG$_2$]** causes a problem]]].
    c.   *Complex NP island*
        * This person$_4$, I painted [a portrait of t_4 [$\varnothing_{REL}$ that t_{REL} unfortunately started **[the rumor that I dislike PG$_4$]**]].

Recall from chapter 5 that a PG in an embedded island is possible if the containing

island also hosts a PG, as in (9). This is an automatic expectation of the null operator theory of PGs: the movement that forms the PG in the containing island should itself be able to license another PG in an embedded island.

(9)   *Recursive PG licensing*

      a.   A person who$_1$ I hang out with t_1 [because [friends of PG$_1$] admire PG$_1$]
            (Nissenbaum 2000, p. 26, ex. 13c)

      b.   Guess [which food]$_1$ I can't help but eat t_1 [despite really fearing PG$_1$ [because there are carcinogens in PG$_1$]].

For this reason, we expect the PG-forming operator movement in the sentences in (1) to be able to license a second PG, in an island contained by that relative clause. This is indeed possible, as we see in (10). Example (10a) includes a sentential adjunct with a PG in the relative clause, while (10b) places a PG in the subject of the relative clause.

(10)   *PG in relative clause can license an additional embedded PG*

      a.   Who$_1$ did Mary take [pictures of t_1 [$\varnothing_\text{REL}$ that t_REL weren't very flattering to **PG**$_1$ [because she put an awful wig on **PG**$_1$]]]?

      b.   That's the manager which$_2$ I know [an employee of t_2 [who$_\text{REL}$ t_REL thinks that [every message from **PG**$_2$] reveals the stupidity of **PG**$_2$]]].

    The various facts just discussed are precisely as we expect if PGs in relative clauses are formed by movement of a null operator to their edge, just as for PGs in islands of other sorts. I will therefore assume this in what follows. Next, I will consider the semantic composition required to interpret PGs of this form.

2.1   The semantic composition of relative clauses with PGs

I will assume that relative clauses are derived by movement of a relativizing operator to the edge of a clause of type t, which triggers Predicate Abstraction, yielding the denotation <e,t> for the relative clause (see Heim and Kratzer 1998, p. 96). Under normal circumstances, this constituent would simply adjoin to NP and be interpreted by Predicate Modification, as discussed above. However, for a relative clause containing a PG, more must be said. In such a context, movement of an additional operator in order to form a PG will trigger another instance of Predicate Abstraction, making the relative clause a two-place predicate of type <e,<e,t>>, as

we see in (11) below, which diagrams the relative clause from (1a) above.[3]

(11)   *A PG-containing relative clause is a two-place predicate*

CP

$\langle e,\langle e,t\rangle\rangle$

($\lambda y.\lambda x.x$ weren't that flattering to y)
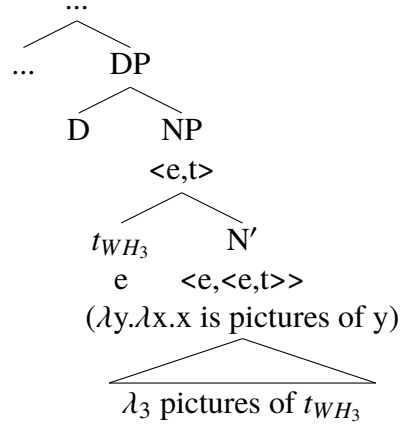
OP_{REL_1}

OP_{PG_2}

$\lambda_2\ \lambda_1$ that t_{REL_1} weren't that flattering to t_{PG_2}
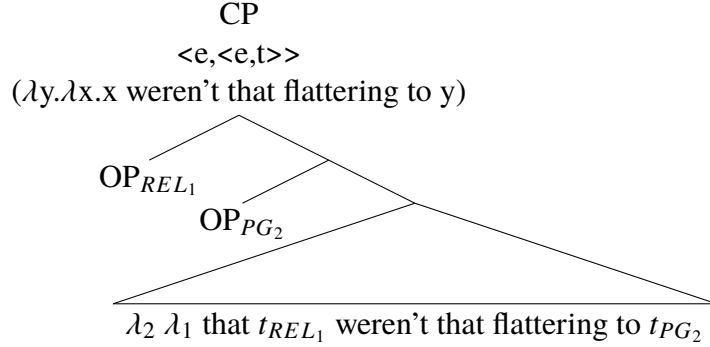
In order for this constituent to be interpreted, it must merge to a position in NP of the same type, carrying over from chapter 5 the assumption that Predicate Modification can unite any two nodes whose semantic types match (Partee and Rooth 1983; Nissenbaum 2000; Nissenbaum and Schwarz 2011). If NPs are generally type $\langle e,t\rangle$ as previously mentioned, such a position will be created by successive cyclic extraction via the NP edge: when movement pauses in the edge of the NP, it will apply Predicate Abstraction to this constituent, making it type $\langle e,\langle e,t\rangle\rangle$. This is diagrammed in in (12) below:

(12)   *Successive-cyclic movement via NP edge*

...

...   DP

D   NP

$\langle e,t\rangle$

t_{WH_3} N′

e $\langle e,\langle e,t\rangle\rangle$

($\lambda y.\lambda x.x$ is pictures of y)

λ_3 pictures of t_{WH_3}

Notice that the relative clause in (11) and the sister of the intermediate trace of

---

[3]In (11) I have represented the relativizing and PG-forming operators as forming crossed paths, with the semantic argument positions created by those movements via Predicate Abstraction appearing in an order opposite to that of the corresponding operators ($OP_{REL_1}\ OP_{PG_2}\ ...\ \lambda_2\ \lambda_1$). This representation follows from the concerns of footnote 11 in chapter 5, and is justified in detail in section 4 of this chapter.

movement through NP in (12) are both type <e,<e,t>>. Consequently, the former can merge to the latter and be successfully interpreted by Predicate Modification. This is illustrated in (13) below, which models the relevant segment of (1a):[4]

(13)   *Licensing of PG in relative clause*

```
                    ...
              ...         DP
                      D        NP
                              <e,t>
                        t_WH₃        N′
                          e      ┌──────────┐
                                 │ <e,<e,t>> │
                                 └──────────┘
        (λy.λx.x is pictures of y and x weren't that flattering to y)

              N′                          CP
          <e,<e,t>>                    <e,<e,t>>
   (λy.λx.x is pictures of y)   (λy.λx.x weren't that flattering to y)

    λ₃ pictures of t_WH₃        OP_REL₁
                                     OP_PG₂

                   λ₂ λ₁ that t_REL₁ weren't that flattering to t_PG₂
```
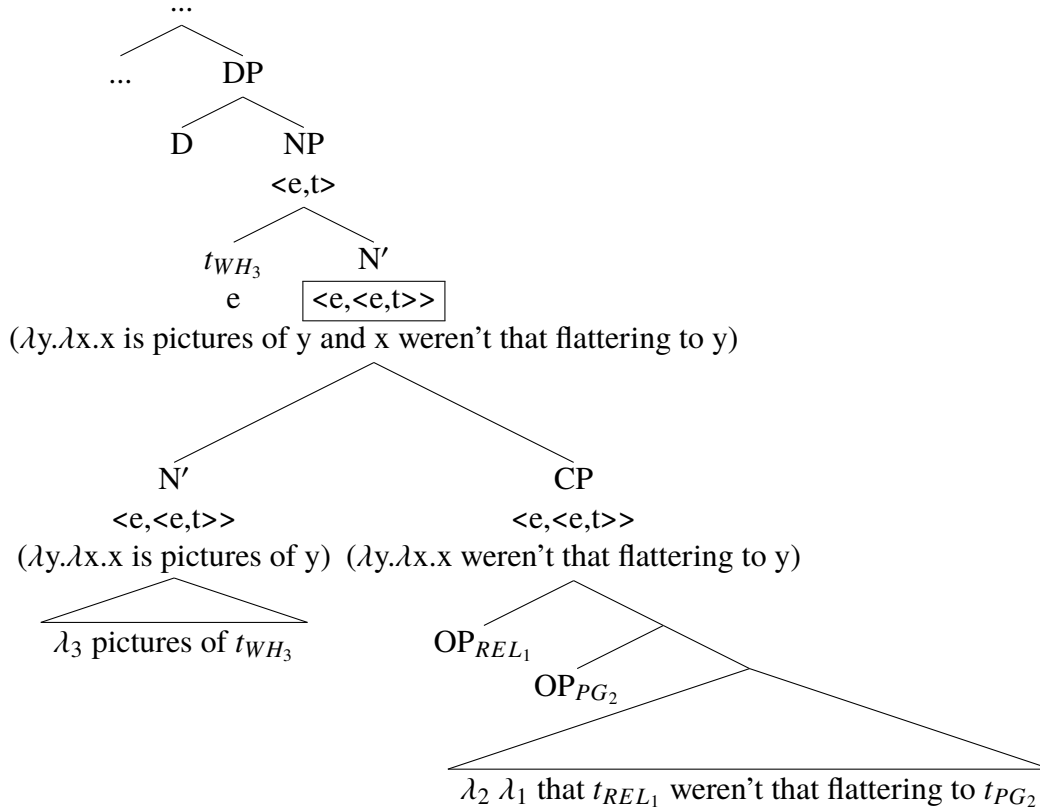
When the first semantic argument position of the two-place predicate created by merger of the relative clause (boxed above) is saturated by the trace of successive cyclic movement from NP, that moved phrase will serve as the binder both for its own original trace, and the PG. When the first semantic argument position of the function denoted by the N′ that merger of the relative clause creates (boxed). After this occurs, the NP node immediately dominating the trace of successive cyclic movement is type <e,t>, as is usual for NPs, and is thus the right type to undergo Functional Application with D/Q if present.

---

[4]The configuration represented by (13) uses the relational noun *picture (of)*. For convenience I do not elaborate here on the internal semantic composition of such noun phrases, since regardless of how this occurs, these must end up with an <e,t> denotation like λx.picture-of(x,John). If this were not the case, such noun phrases would not be the right type to undergo adjunction of adjectives and relative clauses, or to undergo Functional Application with D/Q.

Importantly, the structure in (13) would not be interpretable in the absence of successive cyclic movement from NP, since in this case there would not be a two-place predicate in the NP for the PG-containing relative clause to merge to.[5] [6] The possibility of such movement is consistent with the concept that NP is a phase, as mentioned above. However, if NP is a phase, we expect such movement to not only be possible, but necessary. In this chapter, I will not be able to establish whether this movement is obligatory or not.[7] However, in section 5 we will see several other pieces of evidence that such movement is available.

This concludes the analysis of the basic properties of PG-hosting relative clauses. In the remainder of this chapter, I will consider several further consequences and puzzles that are related to this phenomenon.

3   PGs in stacked relative clauses

In this section, I extend the above analysis to contexts where multiple relative clauses are stacked in one NP. The pattern we will see here directly mirrors an earlier observation from Nissenbaum (2000), who shows that when sentential adjuncts are

---

[5]In my judgment, it is possible for PG-hosting relative clauses to extrapose:

(i) a. Who_1 did Mary take [pictures of t_1 t_2] yesterday [that weren't that flattering to PG_1]$_2$?
     b.   Let me tell you who_3 I've noticed [an aspect of t_3 t_4], just now, [that really makes me want to avoid PG_3]$_4$.

If these PGs depend on the relative clause being interpreted as adjoined to an NP that has been passed through by successive cyclic A′-movement, then this fact may serve as evidence for the theory of adjunct extraposition as late merge after covert movement of the host phrase, which was discussed extensively in the previous chapter. Under such an analysis, the extraposed relative clauses in (1) are interpreted in precisely the same way as usual, the only difference being (covert) movement of the source DP, which contains the PG-licensing trace of extraction from NP.

[6]It is conceivable that an analyst opposed to the proposal of successive cyclic movement from NP might posit that the PG-hosting relative clause extraposes to adjoin to DP, in order for its PG to be licensed by successive cyclic movement through that phrase instead:

(i) $[_{CP}$ WH_4 ... V $\boxed{[_{DP}}$ t_4 D $[_{NP}$ $[_{N''}$ $[_{N'}$ N t_4] t_3]] $[_{RC}$ ... $PG_4]_3$ $\boxed{]}$]

Ignoring the fact that such an analysis is incompatible with the semantic concerns just discussed, this claim leads to the expectation that if an NP contains multiple relative clauses, but only one has a PG, then the PG-containing one would extrapose past the one without a PG in the process of adjoining to DP. We will see in section 3 that the opposite pattern is attested in reality: when only one of two stacked relative clauses in one NP contains a PG, it must be the innermost one.

[7]Note that the edge of NP is not a possible landing site for any independent movement process in English. Thus it is not obvious that such movement can be attributed to anything other than forced successive-cyclicity, though this consideration is theory-internal and therefore not decisive.

stacked within one clause, both adjuncts can have a PG (14a), but if only one has a PG, it must be the inner one (14b), not the outer one (14c):

(14)   *PGs in stacked sentential adjuncts*

    a.   Who$_1$ will you hire t_1 [without interviewing PG$_1$] [if John recommends PG$_1$]?

    b.   Who$_1$ will you hire t_1 [without interviewing PG$_1$] [if John recommends him]?

    c.   * Who$_1$ will you hire t_1 [without interviewing him] [if John recommends PG$_1$]?
        (Nissenbaum 2000, p. 66)

I observe that an analogous contrast applies to PGs in relative clauses stacked in the same NP (15). While both relative clauses can have a PG (15a), if only one does, it must be the inner (15b) and not the outer (15c):
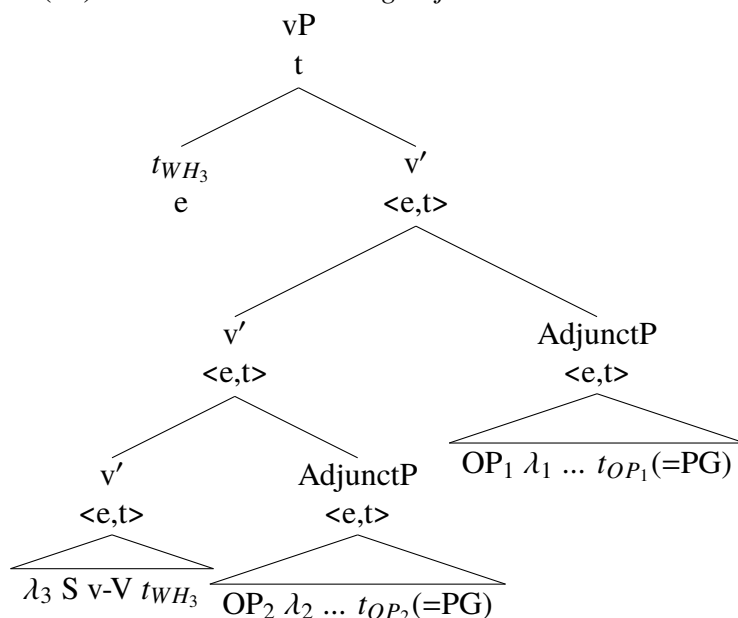
(15)   *PGs in stacked relative clauses*

    a.   Guess [which actor]$_8$ I took pictures of t_8 [that weren't very flattering to PG$_8$] [that unfortunately really embarrassed PG$_8$].

    b.   Guess [which actor]$_8$ I took pictures of t_8 [that weren't very flattering to [him$_8$/**PG$_8$**]] [that unfortunately turned out blurry].

    c.   Guess [which actor]$_8$ I took pictures of t_8 [that unfortunately turned out blurry] [that weren't very flattering to [him$_8$/***PG$_8$**]].

Here I'll argue that Nissenbaum's account for the facts in (14) extends naturally to the pattern in (15), in the context of my proposal that PGs in relative clauses are licensed by successive cyclic movement from NP.

    As mentioned several times throughout this dissertation, for Nissenbaum, PGs in sentential adjuncts are licensed by successive cyclic A′-movement through spec-vP. Such movement triggers Predicate Abstraction in the vP, creating an <e,t> position to which a PG-bearing adjunct can be merged, as illustrated in (2) above. Nissenbaum argues that examples like (14a) simply involve the merger of two such adjuncts, as diagrammed in (16). Since these adjuncts combine with the vP by Predicate Modification, in principle, any number of them could be included in the same way:

(16)   *Two PG containing adjuncts in one vP*

vP
t

t_{WH_3}
e

v′
<e,t>

v′
<e,t>

AdjunctP
<e,t>

v′
<e,t>

AdjunctP
<e,t>

OP_1 λ_1 ... t_{OP_1}(=PG)

λ_3 S v-V t_{WH_3}

OP_2 λ_2 ... t_{OP_2}(=PG)

A sentential adjunct that lacks a PG will simply be a constituent of type t. Since such an adjunct cannot combine with the <e,t> predicate formed by successive cyclic movement from vP, it must adjoin above the landing site of successive cyclic movement, as in (17). This correctly models the sentence in (14b), in which the inner adjunct contains a PG, and the outer one does not:

(17)   *Lower adjunct with PG, higher adjunct without*

vP
t

v′
t

AdjunctP
t

...

t_{WH_3}
e

v′
<e,t>

v′
<e,t>

AdjunctP
<e,t>

λ_3 S v-V t_{WH_3}

OP_2 λ_2 ... t_{OP_2}(=PG)
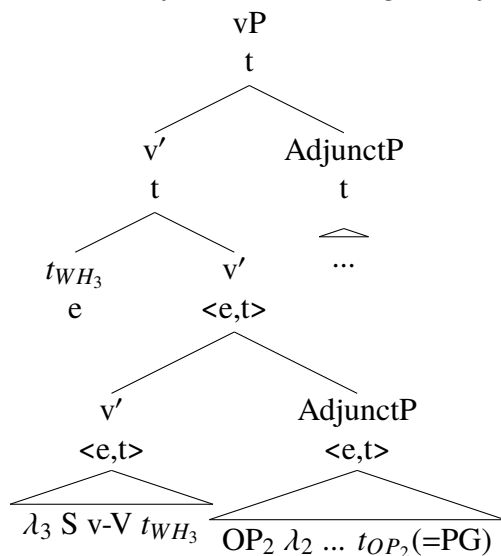
In order to create the structure in (14c), it would be necessary for the PG-less adjunct to merge structurally beneath the PG-containing one, as in (18). Since this region of the vP is a semantic predicate (type <e,t>) due to the effect of Predicate Abstraction, merger of 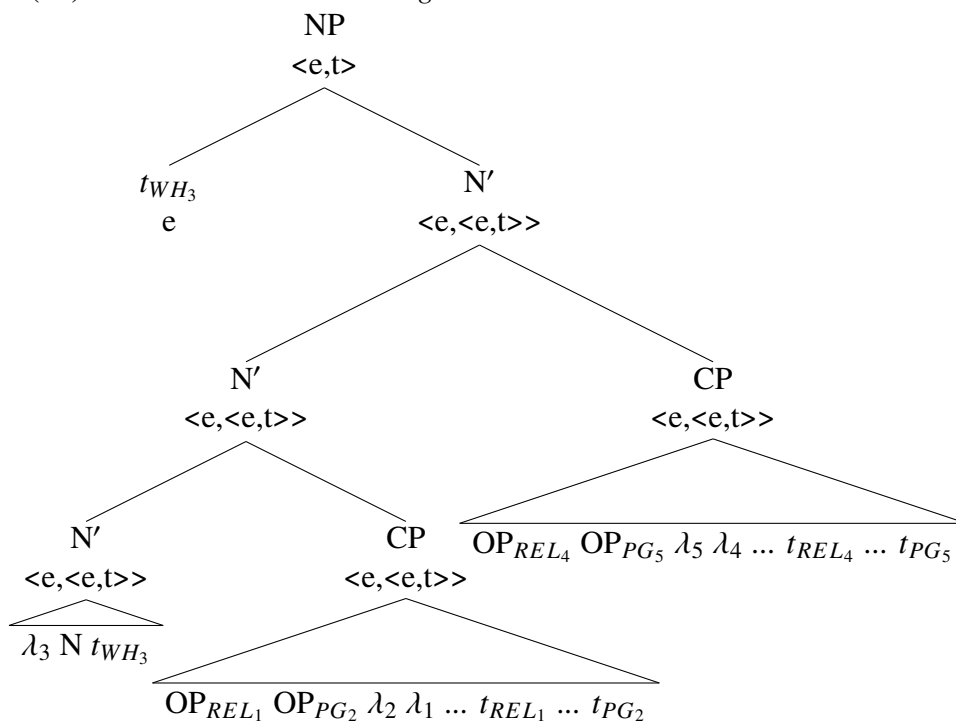a PG-less adjunct (type t) here will result in a type mismatch. Neither Predicate Modification nor Functional Application can yield an interpretation for this configuration, which we have seen in (14c) is indeed unacceptable:

(18)   *No PG-less adjunct below PG-containing one*

$$\begin{array}{c}\text{vP} \\ \text{t}\end{array}$$

Tree diagram:

- vP (t)
  - t_{WH_3} (e)
  - v′ (<e,t>)
    - v′ *
      - v′ (<e,t>): λ_3 S v-V t_{WH_3}
      - AdjunctP (t): ...
    - AdjunctP (<e,t>): OP$_2$ λ_2 ... t_{OP_2}(=PG)
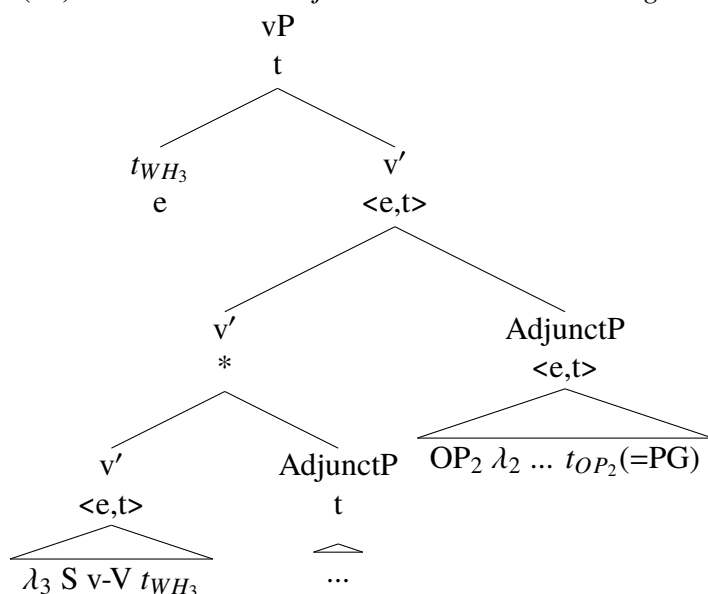
Precisely the same reasoning, involving higher semantic types, applies to the relative clause facts in (15) above. If an NP is usually a predicate from individuals to truth values <e,t>, which will be subject to Predicate Abstraction when successive cyclic extraction from it occurs, then there will be an <e,<e,t>> position in the NP with which any number of PG-containing relative clauses can combine. A structure with two PG-bearing relative clauses in a single NP is shown in (19):

(19) *Stacked PG-containing relative clauses*

```
                        NP
                       <e,t>
              /                    \
          t_WH₃                      N′
            e                   <e,<e,t>>
                          /                  \
                       N′                      CP
                   <e,<e,t>>               <e,<e,t>>
                /          \            /              \
             N′             CP      OP_REL₄ OP_PG₅ λ₅ λ₄ ... t_REL₄ ... t_PG₅
         <e,<e,t>>      <e,<e,t>>
        /    |    \      /      \
      λ₃  N  t_WH₃   OP_REL₁ OP_PG₂ λ₂ λ₁ ... t_REL₁ ... t_PG₂
```

Where:
- NP: $\langle e,t\rangle$
- t_{WH_3} : e
- N′: $\langle e,\langle e,t\rangle\rangle$
- N′: $\langle e,\langle e,t\rangle\rangle$
- CP: $\langle e,\langle e,t\rangle\rangle$
- N′: $\langle e,\langle e,t\rangle\rangle$
- CP: $\langle e,\langle e,t\rangle\rangle$
- λ_3 N t_{WH_3}
- OP_{REL_1} OP_{PG_2} λ_2 λ_1 ... t_{REL_1} ... t_{PG_2}
- OP_{REL_4} OP_{PG_5} λ_5 λ_4 ... t_{REL_4} ... t_{PG_5}

It is also possible to adjoin a PG-containing relative clause below the landing site of extraction from NP, and a PG-less relative clause above it, as (20) shows:

(20) *PG-containing relative clause below PG-less one*

```
                    NP
                   <e,t>
              /             \
            N′                CP
           <e,t>            <e,t>
        /        \         /        \
     t_WH₃        N′    OP_REL₄ λ₄ ... t_REL₄
       e      <e,<e,t>>
            /          \
          N′             CP
      <e,<e,t>>       <e,<e,t>>
      /   |   \       /        \
    λ₃  N  t_WH₃   OP_REL₁ OP_PG₂ λ₂ λ₁ ... t_REL₁ ... t_PG₂
```
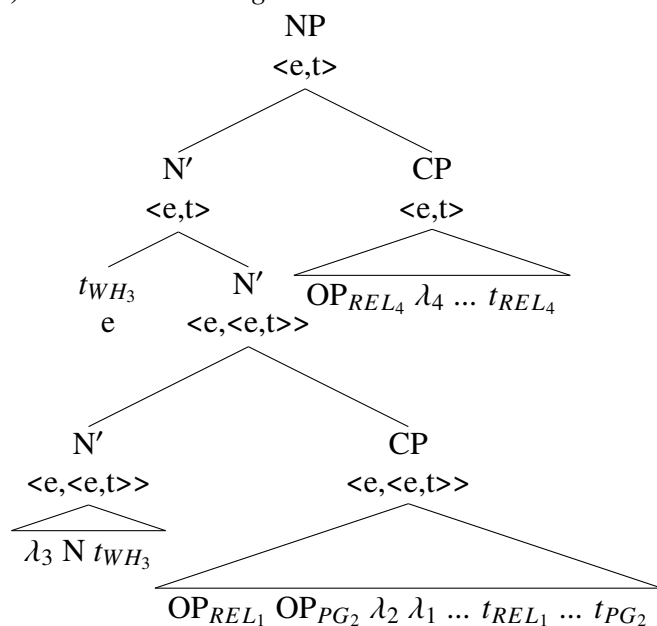
In contrast, a PG-less relative clause cannot be merged beneath a PG-containing one. A PG-containing relative clause is type <e,<e,t>> and thus must be merged below the trace of extraction from NP, but a PG-less relative clause of type <e,t> will simply yield a type mismatch if merged into this region of the NP, as (21) shows:

(21)   *No PG-less relative clause below PG-containing one*
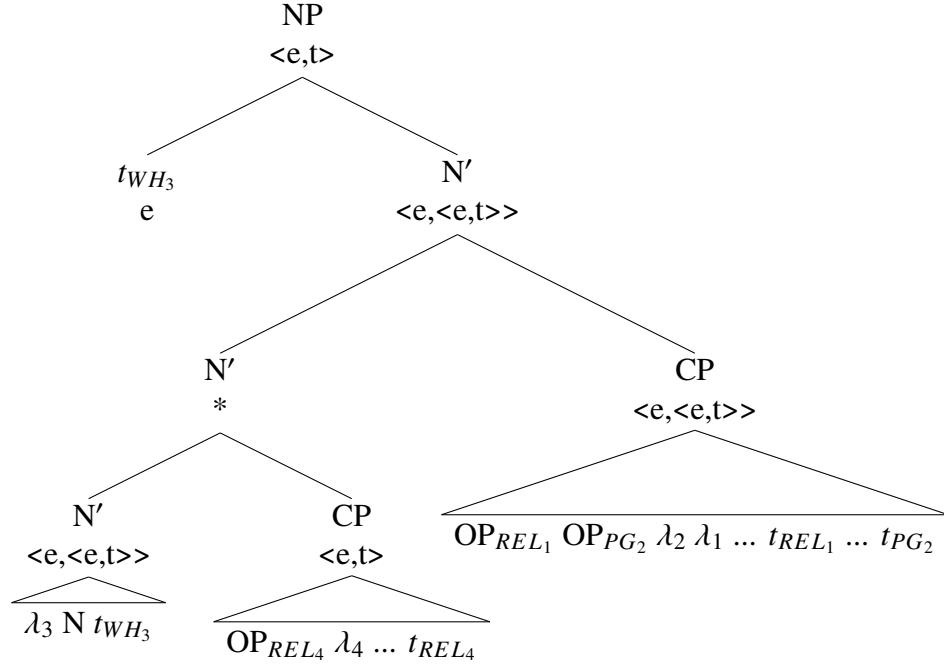
In summary, both sentential adjuncts adjoined to vP, and relative clauses adjoined to NP, show an asymmetry in PG distribution that when stacked within the same domain. This pattern emerges straightforwardly from an understanding of PGs as dependent upon Predicate Abstraction triggered by (successive cyclic) A′-movement. This view opens up the possibility of adjunction either above or below the trace of extraction, except when a semantic conflict would arise, from which the asymmetry just examined stems.[8]

---

[8]It is worth asking why it is not possible to create the illicit configurations in (14c) and (15c) by merging the PG-less constituent even lower, to a segment of vP/NP that is below the region affected by Predicate Abstraction. If as Nissenbaum (2000) suggests Predicate Abstraction is not a mechanism that introduces a distinct λ-node into the syntactic tree (contra Heim and Kratzer 1998) but rather essentially a type-shifting operation, then there is not necessarily a distinct syntactic position where lower merger of the sort just described could actually successfully occur. Alternatively, adjunction too deeply within a given domain, rather than at its periphery, may be independently illicit (see Stepanov 2001 for relevant discussion, in particular his version of *Least Tampering*). This possibility is reminiscent of the discussion at the end of the previous chapter, though not strictly related, since that chapter focused strictly on late adjunction. See Nissenbaum and Schwarz (2011) for further

4 On the order of overlapping operators

Here I will consider in more detail the possible interpretations for the gaps in PG-bearing relative clauses. From this examination we learn, I argue, that overlapping operator movements must form crossing paths—a proposal also made by Nissenbaum (2000), in the context of sentential adjuncts containing multiple PGs. I will first overview Nissenbaum's evidence for this conclusion in order to set the stage for the analysis of the relative clauses in question.

As discussed in section 2 of chapter 5, Nissenbaum (2000) observes that in an English multiple *wh*-question, the seemingly in situ *wh*-phrase can license the second of two PGs contained within the same sentential adjunct, as (22) below shows once more. Based on this fact, Nissenbaum argues that the in situ phrase in such contexts does move, but covertly:

(22) *Wh-in-situ licenses the second of two PGs in one adjunct*
[Which senator]$_1$ did you persuade t_1 to borrow [**which car**]$_2$ [after getting an opponent of PG$_1$ to put a bomb in **PG$_2$**]?
(Nissenbaum 2000, p. 12 ex. 8)

Nissenbaum also observes that in such configurations, the surface order of licensing phrases must match the order of corresponding PGs in the multi-PG adjunct, as shown in (23) below, which contrasts with (22) above:

(23) *Surface order of licensers and PGs cannot be mismatched*
* [Which car]$_1$ did you lend t_1 to [which senator]$_2$ [after getting an opponent of PG$_2$ to put a bomb in PG$_1$]?
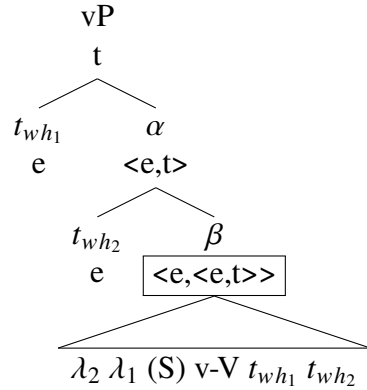(Nissenbaum 2000, p. 110 ex. 27a)

Nissenbaum (p. 128, p. 133 fn. 10) takes this fact as evidence that the two null operators that move in the multi-PG adjunct form crossed paths, precisely as he argues is the case in (non-superiority-violating D-linking) multiple *wh*-questions, as can be seen overtly in languages like Bulgarian and Romanian. Nissenbaum does not make fully explicit why this is the correct conclusion, however. Here I will describe concretely why this is the case, in order to set the stage for further analysis of PG-containing relative clauses.

In order to make the relevant concepts evident, let's begin by considering the interpretation for the vP of a multiple *wh*-question, as shown in (24) below. The node β here (boxed) is a two-place predicate because the two instances of *wh*-movement from vP have both triggered an instance of Predicate Abstraction. As summarized in chapter 5, Nissenbaum argues that the two instances of Predicate Abstraction

---

discussion of this general issue for the null operator theory of PGs.

in such a context must "stack" in this way. If this were not possible, merger of a sentential adjunct with two PGs (which is also a two-place predicate) would not be possible, which would incorrectly rule out examples like (22) above.
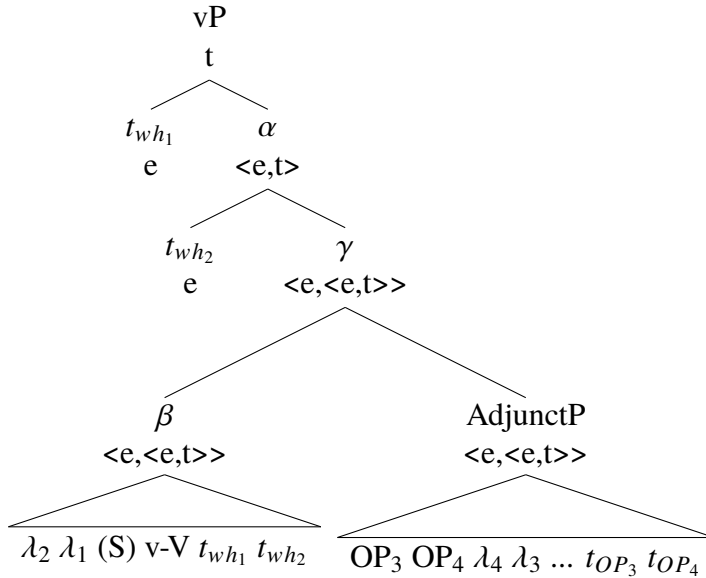
(24)   *vP after two successive-cyclic movements*

$$
\begin{array}{l}
\text{vP} \\
\text{t} \\
\quad t_{wh_1} \qquad \alpha \\
\quad\ \ \text{e} \qquad \langle e,t\rangle \\
\qquad\qquad t_{wh_2} \qquad \beta \\
\qquad\qquad\ \ \text{e} \quad \boxed{\langle e,\langle e,t\rangle\rangle} \\
\qquad\qquad\qquad \lambda_2\ \lambda_1\ (\text{S})\ \text{v-V}\ t_{wh_1}\ t_{wh_2}
\end{array}
$$

Notice that the function denoted by β contains unsaturated argument positions which are in reverse order relative to the specifiers of *wh*-movement that triggered their formation ($t_{wh_1}\ t_{wh_2}\ ...\ \lambda_2\ \lambda_1$). Since this function will combine with the inner specifier of *wh*-movement before the outer one, the ordering $\lambda_2\ \lambda_1$ is necessary for semantic composition here: only this ordering ensures that the moved *wh*-phrases are ultimately interpreted in a way corresponding to their actual traces in VP.

With this fact in mind, next consider a structure like (24) that also includes an adjunct with two PGs, as in (25) below:

(25)   *vP after two successive-cyclic movements with a two PG adjunct*

$$
\begin{array}{l}
\text{vP} \\
\text{t} \\
\quad t_{wh_1} \qquad \alpha \\
\quad\ \ \text{e} \qquad \langle e,t\rangle \\
\qquad\qquad t_{wh_2} \qquad \gamma \\
\qquad\qquad\ \ \text{e} \quad \langle e,\langle e,t\rangle\rangle \\
\qquad\qquad\qquad \beta \qquad\qquad\qquad\qquad \text{AdjunctP} \\
\qquad\qquad \langle e,\langle e,t\rangle\rangle \qquad\qquad \langle e,\langle e,t\rangle\rangle \\
\qquad \lambda_2\ \lambda_1\ (\text{S})\ \text{v-V}\ t_{wh_1}\ t_{wh_2} \quad \text{OP}_3\ \text{OP}_4\ \lambda_4\ \lambda_3\ ...\ t_{OP_3}\ t_{OP_4}
\end{array}
$$

Here the operators in the adjunct also form a predicate whose semantic argument positions are reversed with respect to the phrases that formed them (OP_3 OP_4 λ_4 λ_3). If this is indeed the way that multiple specifiers always interact with Predicate Abstraction (despite the fact that the operators here do not actually saturate the predicate they create), and if multiple operators indeed move with crossing paths as shown here, then we accurately predict the contrast between (22) and (23) above.

In (25), Predicate Modification will unite λ_2, corresponding to t_{wh_2}, with λ_4, corresponding t_{OP_4} (which is the second PG). Predicate Modification will also unite λ_1, corresponding to t_{wh_1}, with λ_3, corresponding to t_{OP_3} (which is the first PG). Consequently, the first λe of the resulting function (labeled γ) binds t_{wh_2} and t_{OP_4}, while the second λe of this function binds t_{wh_1} and t_{OP_3}. This first λe will be saturated by the trace in the inner specifier of vP, while the second λe will ultimately (via the function labeled α here) be saturated by the trace in the outer specifier of vP. As a result, the phrase that formed the inner specifier of vP will bind its trace t_{wh_2} in the VP and the second PG, while the phrase that formed the outer specifier of vP will bind its trace t_{wh_1} as well as the first PG. This accurately corresponds to the readings for the PGs in the sentence in (22) above.

Importantly, the reading for (22) that (25) accurately models depends on the two operators in the adjunct moving with crossing paths, with the corresponding semantic argument positions they trigger via Predicate Abstraction occurring in reversed order. If these operators were able to move with nesting paths, then we would expect the referents of the PGs to potentially be reversed relative to the surface order of their licensers, yielding a sentence like (23). Since (23) is unacceptable, the necessary conclusion is that the operators in such a configuration must form crossing paths, as Nissenbaum suggests.[9] With the rationale for this proposal now made explicit, we are ready to return to the analysis of PGs in relative clauses.

All examples of PG-containing relative clauses shown so far in this chapter have involved the gap of relativization preceding the PG. A further and potentially surprising fact is that reversing the gaps is not acceptable, as (26) shows. Here the intended PG is marked in the usual way, while the gap intended to correspond to relativization is underlined:
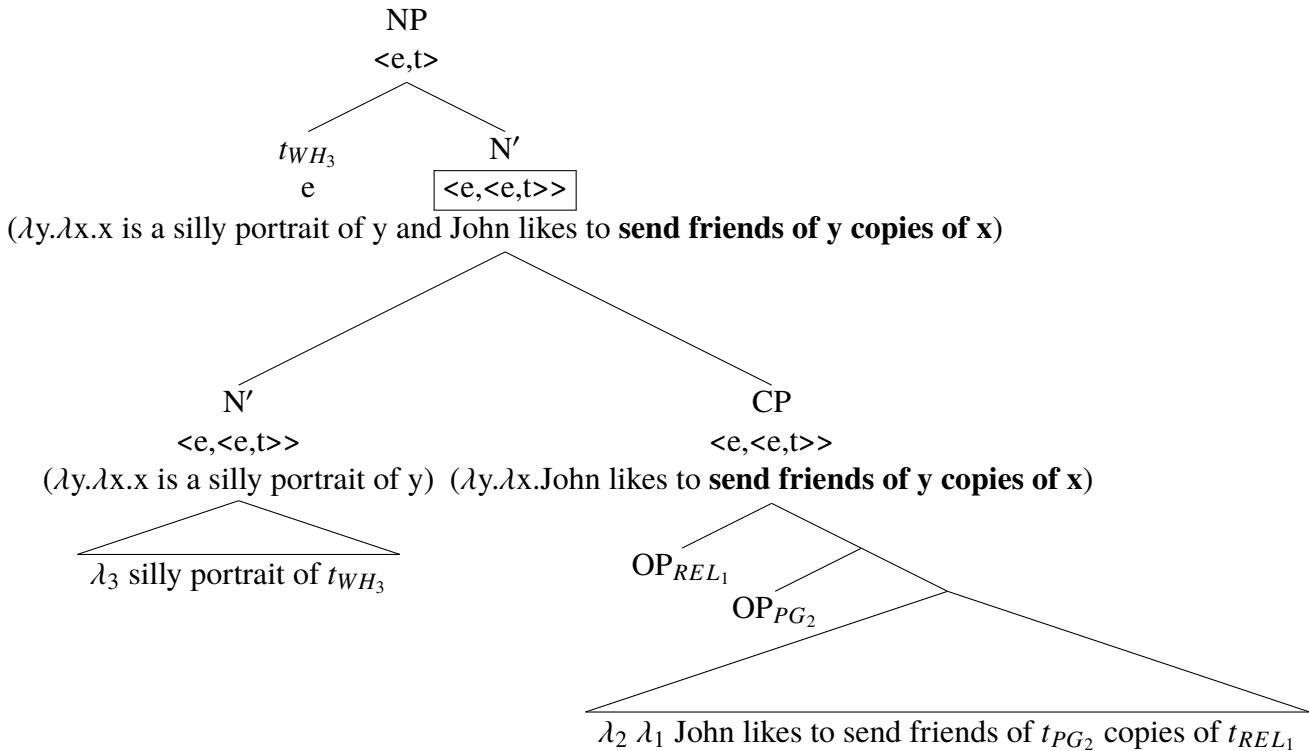
---

[9]This understanding requires the assumption that generalized Predicate Modification of two two-place predicates results in a function whose first semantic argument corresponds to the union of the initial arguments of its two constituent functions, and whose second semantic arguments correspond to the union of those functions' second arguments. If this were not the case, there is no reason why there should be any predictable correspondence between the relative ordering of multiple moved phrases and multiple PGs. I will continue to assume this in what follows.

(26)   *Reltivization gap must precede PG*

      a. Who_1 did you paint [a silly portrait of t_1]$_2$ [that John likes to give copies of __$_2$ to friends of PG_1]?

      b.   *?? Who_1 did you paint [a silly portrait of t_1]$_2$ [that John likes to send friends of PG_1 copies of __$_2$]?

      c.   Let me tell you [which animal]$_1$ I made [a statue of t_1]$_2$ [that I intend to send __$_2$ to a farmer of PG_1].

      d.   *?? Let me tell you [which animal]$_1$ I made [a statue of t_1]$_2$ [that I intend to send a farmer of PG_1 an exact copy of __$_2$].

The proposal that the relativizing operator and the PG forming operator must move with crossing paths correctly predicts the pattern of judgments in (26). To see why, let's consider example (26b) in detail.

     The intended reading for (26b) is unavailable, despite being semantically and pragmatically well-formed. However, this reading would require the moving operators to form nesting paths, as illustrated in the structure in (27):

(27)   *Intended reading of (26b): Nesting operator movements*

NP

$<e,t>$

t_{WH_3} N′

e $<e,<e,t>>$

(λy.λx.x is a silly portrait of y and John likes to **send friends of y copies of x**)

N′   CP

$<e,<e,t>>$ $<e,<e,t>>$

(λy.λx.x is a silly portrait of y)   (λy.λx.John likes to **send friends of y copies of x**)

λ_3 silly portrait of t_{WH_3} OP_{REL_1}

                                  OP_{PG_2}

λ_2 λ_1 John likes to send friends of t_{PG_2} copies of t_{REL_1}

In (27), the order of semantic argument positions in the function that the relative clause denotes will appropriately unite the λe inherent to the denotation of the NP

picture of with that corresponding to the gap of relativization (*copies of __*), and also allow the phrase extracted in the matrix clause (*who*) to license the intended PG (*friends of PG*). However, to the extent that (26b) has an interpretation, the most salient one reverses the order of the gaps, yielding an absurd reading where we are speaking about copies of a person, and friends of a portrait. This is the reading we expect to be forced if the operator movements in the corresponding structure must cross, as opposed to the nesting required for the intended reading that (27) demonstrates. Consequently, the hypothesis that overlapping operator movements must form crossing paths makes the correct prediction.

Note that when examining the sentences in (26), we do not in fact know, prima facie, which is the gap of relativization and which is the PG. Overall, the generalization about sentences of this variety is that the head of the relative clause always co-refers with the first gap, and the extracted phrase with the second gap. If we assume that Predicate Abstraction in NP triggered by successive cyclic extraction from it will always form the outer λe of the resulting type <e,<e,t>> N′ (as shown in the preceding diagrams), and if we maintain that the two operators in the relative clause must always form crossing paths (resulting in Predicate Abstraction in the reverse order as argued above), then this is indeed what we expect. The outer λe of the type <e,<e,t>> relative clause will always correspond to the inner of the two moved operators, which in turn always corresponds to the second gap in the relative clause. Since the outer λe of the relative clause is united with the outer λe of the relevant N′ by Predicate Modification, the fact that the outer λe of N′ is bound by the extracted phrase in the matrix clause means that the second gap in the relative clause will always be bound by that phrase as well. For all intents and purposes, then, we end up with a configuration in which the gap of relativization precedes the PG. As we saw in (26) above, this is indeed the only possibility for such configurations.

Finally, consider the result just established in light of the following prediction of Nissenbaum (2000) about PGs in multiple specifier configurations, which was discussed extensively in chapter 5:

(28)   *Multiple Specifier Single Parasitic Gap Prediction*
        In multiple specifier constructions in which XP_1 is the highest specifier of vP and XP_2 is a low specifier of vP, only XP_1 can license a single parasitic gap in an adjunct to vP.

Given that the gap of relativization always precedes the PG, if the two operators corresponding to those gaps move in crossing-paths fashion, then we expect the outer of the two operators to be the relativizing one. This is as shown in the diagram in (13) above, for instance. This fact, combined with the prediction in (28), leads us to expect that a PG in a sentential adjunct of a PG-containing relative clause will

only be able to be licensed by the relativizing movement chain. This is the correct prediction, as (29) shows. This is indicated by the fact that the additional PG in the sentential adjunct here must be co-referent with the head of the relative clause, rather than the constituent extracted from NP in the matrix clause.

(29)   *Only relativization licenses additional PG*

      a.   That's the guy who$_1$ I painted [a silly portrait of t_1]$_2$ [that John likes to give copies of __$_2$ to friends of PG$_1$ [**in order to make them want to buy PG$_2$**]].

      b.   * That's the guy who$_1$ I painted [a silly portrait of t_1]$_2$ [that John likes to give copies of __$_2$ to friends of PG$_1$ [**in order to introduce them to PG$_1$**]].

Since these triple gap sentences strain the limitations of processing, I will not manipulate them further here.

To conclude this section, the interpretive possibilities for PG-containing relative clauses correctly emerge from the hypothesis that overlapping operators movements form crossing paths—a conclusion proposed by Nissenbaum (2000) based on the behavior of adjuncts containing multiple PGs. Given the conclusions of chapter 5 about about the distribution of crossing and nesting paths, this result indicates that a relativizing operator and PG-forming one are members of the same syntactic natural class, such that any feature that can attract one of them will also attract the other, if present. I will leave a deeper consideration of this finding to other work.

5   More on movement through NP edges

The preceding sections of this chapter have either argued for, or relied on, my proposal that PGs in relative clauses are licensed by successive cyclic A′-movement from NP (and importantly not DP). In this section, I provide further potential evidence for such movement, before then going on to consider a consequence of this result for the theory of left branch extraction in Bošković (2005).

5.1   Further evidence

5.1.1   Variable binding versus principle C

The first piece of potential additional evidence for movement via the NP edge comes from an interaction between variable binding and principle C, discussed previously as a diagnostic for successive cyclicity by at least Lebeaux (1992) and Sauerland (1998). An instance of this interaction consistent with the possibility of movement

via the CP edge is provided in (30) below. In the examples of (30), a relative clause of a *wh*-moved phrase contains a pronoun *he* co-indexed with the quantifier phrase *every student*, as well as an R-expression *Mary* co-referent with the pronoun *she*.

(30)   *Successive cyclicity and variable binding versus principle C*

    a.   * [Which paper that he$_1$ gave to Mary$_2$]$_3$ did she$_2$ think [$_{CP}$ t_3 that [every student]$_1$ would like $\boxed{t_3}$]?

    b.   [Which paper that he$_1$ gave to Mary$_2$]$_3$ did [every student]$_1$ think [$_{CP}$ $\boxed{t_3}$ that she$_2$ would like t_3]?
       (Adapted from Sauerland 1998, ex. 31)

A quantifier phrase must c-command the position in which a co-referent pronoun is interpreted. In contrast, an R-expression must not be interpreted in a position c-commanded by a co-referent element, or else a principle C violation will arise. Given these requirements, in (30a), variable binding must be satisfied by reconstruction of the *wh*-phrase to its base position, which is the only position in the derivation it has occupied that is c-commanded by the quantifier phrase *every boy*. Since this position is also c-commanded by the pronoun *she* co-referent with the R-expression *Mary* in the relative clause, there is an irreconcilable conflict between the needs of variable binding and principle C, resulting in unacceptability. In contrast, in (30), the positions of *every student* and *she* are reversed. In this context, an acceptable result can be achieved by reconstructing the *wh*-phrase to its intermediate landing site in the edge of the embedded CP. In this position, the pronoun *he* in the relative clause is appropriately c-commanded by *every student*, and the R-expression *Mary* is not c-commanded by the co-referent pronoun *she*, which is lower in the embedded clause. The fact that reconstruction to this position is possible is consistent with the hypothesis that A′-movement from CP passes through its edge.

We can manipulate the configuration in (30) in order to provide evidence consistent with the possibility of movement via the NP edge, as in the examples of (31) below. These involve *wh*-movement from a recursive DP structure in which both DPs are possessed. The moved *wh*-phrase contains a relative clause whose content is the same, in the relevant ways, as that in (30) above. While extraction from possessed DPs is widely known to be degraded in English, there is nevertheless a clear contrast between the two sentences here:

(31)   *Variable binding versus principle C in the NP*

    a.   * [Which picture that Mary$_2$ sent him$_1$]$_3$ did you witness [$_{DP1}$ her$_2$ [$_{NP}$ denial of [$_{DP2}$ every boy$_1$'s claims about $\boxed{t_3}$]]]?

    b.   ? [Which picture that Mary$_2$ sent him$_1$]$_3$ did you witness [$_{DP1}$ every boy$_1$'s [$_{NP}$ $\boxed{t_3}$ denial of [$_{DP2}$ her$_2$ claims about t_3]]]?

In (31a), the *wh*-phrase containing this relative clause must reconstruct to its base position in DP2, whose possessor is *every boy*, for the purposes of variable binding. However, this position is c-commanded by DP1's possessor *her* which is co-referent with the R-expression *Mary* in the relative clause. This conflict makes (31a) unacceptable. In contrast, in the modified configuration in (31b), the positions of *her* and *every boy* are reversed. In this context, it is possible to both satisfy variable binding and avoid a principle C violation by reconstructing to a position between above the possessor of DP2, but below the possessor of DP1. Such a position is consistent with the edge of the NP *denial*, as marked by the trace in (31b). Since (31b) is acceptable, reconstruction to a position like this must be available, indicating that successive cyclic movement passed through that region.

A confound for (31b) is that the requirements in question would also be satisfied by reconstruction to the specifier of the PP *of* in the complement of *denial*.[10] While in chapter 2 I argued against the possibility of movement via spec-PP in English, that argument was internal to the theory of that chapter, and thus not necessarily decisive independent of that chapter's concerns.

5.1.2 Successive cyclic anaphor binding

My second piece of potential evidence for movement via the NP edge comes from principle A. As observed by Barss (1986), and further explored by Nissenbaum (2000), while a phrase cannot undergo A′-movement in order to become the binder of an anaphor, A′-movement can serve to bring a constituent containing an anaphor into a domain where binding is possible. We see an example of this in (32) below. In (32a), the anaphor *herself* contained in the object DP of the embedded clause is un-licensed, because its binder *Mary* in the matrix clause is too far away for binding to be established. However, when the embedded object undergoes A′-movement to the edge of the embedded clause (32b), the anaphor within it is thus close enough to its binder to be licensed:

(32)   *Principle A fed by A′-movement*

    a.   * Mary$_1$ told me [$_{CP}$ that we should keep [this picture of herself$_1$]].

    b.   Mary$_1$ told me [$_{CP}$ [which picture of herself$_1$]$_2$ we should keep t_2].

---

[10]An analogous criticism can be made of (30b) above: this example is also predicted to be acceptable under reconstruction to a low position in the matrix clause c-commanded by *every student*, such as the specifier of VP or vP. This diagnostic is thus unfortunately not particularly fine-grained, though nevertheless informative to some extent.

A separate issue is that, as far as I know, a context has not been identified where this diagnostic actually fails to indicate the presence of potential successive cyclic movement. While if true this fact would have wide-reaching implications, I suspect that more thorough investigation is required here.

If binding is evaluated phase-by-phase (Bošković 2016; Charnavel and Sportiche 2016, a.o.), then the fact that an anaphor moved to the edge of an embedded clause can be bound by a phrase in the matrix clause is not surprising. Combined with the hypothesis that movement from CP passes through its edge, this concept automatically predicts the further fact that if the anaphor-containing object in a configuration like (32) moves all the way to the periphery of the matrix clause, the anaphor within it will be licensed, as we see in (33) below. While the anaphor here could not have been bound either in its base position or surface position, given the acceptability of (32b), we would indeed expect binding to be satisfied in (33) at the time when the object DP inhabited the embedded CP edge. The fact that binding here is successful therefore allows us to infer that successive cyclic movement from CP occurred, as diagrammed in (33):

(33)   *Binding under successive cyclic movement*
        [Which picture of herself]$_1$]$_2$ did Mary$_1$ say [$_{CP}$ $\boxed{t_2}$ we should keep t_2]?

   This line of reasoning provides a potential diagnostic for successive cyclic movement that can be applied to other configurations as well. Performing this test in the nominal domain will require identifying a configuration involving extraction of an anaphor from a DP whose possessor is a potentially licit binder. This test is in principle possible, since a possessor can bind into the NP it c-commands (34):

(34)   *Binding into NP by possessor*
        a.   John$_1$'s picture of himself$_1$
             (Pollard and Sag 1992 ex. 94)

        b.   Lucie$_2$'s joke about herself$_2$/*her$_2$
             (Reinhart and Reuland 1993 ex. 8a)

        c.   Please download [[this computer]$_3$'s information about itself$_3$].

This test will also require extraction from a possessed DP, which as acknowledged above, is degraded. In my judgment, however, extraction from a DP with an indefinite possessor is relatively tolerable:

(35)   *Definiteness and extraction from possessed DPs*
        a.   * I want to know who$_1$ you saw [**the bad artist's / John's** awful picture of t_1].

        b.   ? I want to know who$_1$ you saw [**a bad artist's / someone's** awful picture of t_1].

With these concepts in mind, consider the test configuration in (36) below. The object DP in (36a) has an indefinite possessor, as well as an embedded DP in

its complement containing an anaphor bound by the possessor. Example (36b), in contrast, extracts the anaphor-containing DP by *wh*-movement. To the extent that there is a contrast between these two examples, improvement in (36b) can be attributed to extraction here forming a landing site in the NP edge (marked by the boxed trace), directly below the possessor that is co-indexed with the anaphor:[11]

(36)   *Successive cyclic binding by extraction from NP*

      a.   ?? As a computer scientist, I'm glad to have witnessed [$_{DP}$ [**a super-computer**]$_1$'s destruction of [an expensive painting of **itself**$_1$]]].

      b.   Let me tell you [what kind of expensive painting of **itself**$_1$]$_2$ I've just witnessed [$_{DP}$ [**a supercomputer**]$_1$'s $\boxed{t_2}$ destruction of t_2].

5.2   Locality and extraction from DP: Against Bošković (2005)

While the diagnostics just discussed are not without confounds, if the argumentation of the earlier part of this chapter is correct, the possibility of PGs in relative clauses like those in (1) instantiates evidence for the possibility (though perhaps not the necessity) of successive cyclic movement via the edge of NP. This result has consequences for a proposal about the limitations of left branch extraction from Bošković (2005, 2016), and related work.

Bošković argues, extending insights from previous research (Uriagereka 1988; Corver 1990, 1992), that left branch extraction is blocked when D is present. Bošković argues that this is because D demarcates a phase, and that therefore any extraction from it must successive-cyclically pass through its edge. Movement via spec-DP is licit for complements of N, and indeed, extraction from this position is possible in English (a language with D). Many such examples have been shown in this chapter already. In contrast, Bošković argues that the necessary movement through spec-DP is impossible for adjuncts / specifiers of NP. This is because such movement would be too short and thus in violation of anti-locality (Bošković, 1997; Ishii, 1999; Grohmann, 2003; Abels, 2003, a.o.), as diagrammed in (37) below:[12]

(37)   *Anti-local extraction from adjunct / specifier of NP through spec-DP*

      * [$_{CP}$ ... S V [$_{DP}$ D [$_{NP}$ XP N ...]]]

---

[11]Principle A could also be satisfied in (36b) by reconstruction to the specifier of the PP *of* in the complement of the possessed NP *destruction*. As mentioned above, a similar confound applies to example (31b) as well. I do not see a way of independently ruling out the possibility of movement through spec-PP in English, though I am aware of no clear independent evidence for it either.

[12]In particular, the version of anti-locality recently proposed in Erlewine (2016) and Brillman and Hirsch (2016) is parallel to what Bošković proposes here, as mentioned in previous chapters.

According to Bošković, for this reason left branch extraction of an adjective, for instance, is impossible in English (38a), but permitted in Serbo-Croatian (38b), which lacks D. Under this analysis, the necessary conclusion is that DP is always present in English (since this language never allows left branch extraction)[13] but always absent in article-less languages like Serbo-Croatian.[14]

(38)   *Not all languages permit left branch extraction*
a.   * Expensive$_1$ he loves [$_{DP}$ $\varnothing_D$ t_1 cars].
b.   Skupa$_1$ on voli   [$_{NP}$ t_1 kola].
       expensive he loves   cars
       (Bošković 2016, ex. 16-17)

If DP is indeed a phase (as importantly assumed in the previous chapter), then the possibility of extraction via the edge of NP in English is not compatible with Bošković's account of the lack of left branch extraction in this language. While extraction may not obligatorily pass through the NP edge in English, it should do so at least when it is necessary to license a PG in a relative clause. If Bošković were correct, the sentences in (1) should thus be unacceptable, since the extraction that licenses the PG would, due to anti-locality, be trapped within DP after the initial movement step to the NP edge. Consequently, if this chapter's arguments are correct, Bošković's account of the lack of left branch extraction in English cannot also be correct. Since this chapter is not committed to positing that there are never additional projections between NP and DP, the analysis presented here is not incompatible with the version of anti-locality that Bošković's proposes (which I invoked in chapter 4 of this dissertation). However, because Bošković's account of the lack of left branch extraction in English explicitly depends on there being no additional structure between NP and DP, his account cannot be reconciled with the arguments of this chapter.

I will not attempt to resolve this conflict here. The theory in Bošković (2005), though insightful and frequently cited, may in any case be an over-generalization: see, for instance, Fanselow and Féry (2013) and Pankau (2019) for discussion of some potential counterexamples to it.

---

[13]With the exception, of course, of the instance of possessor extraction analyzed in chapter 3. Given the consensus that the English possessor inhabits spec-DP, Bošković (2005) predicts possessor extraction to be possible in principle in English, as he mentions in footnote 5.

[14]Though in chapter 4, I provided several references for the proposal that the article-less language Russian may in fact have D at least some of the time, despite it never being phonologically expressed.

6   Excursus: PGs in relative clauses of subjects

In this section, I consider a puzzle about a different configuration involving a PG in a relative clause. In the examples that this chapter has focused on, extraction from NP licenses a PG in a relative clause adjoined to that same NP. However, much previous literature on PGs reports examples like those in (39) below, in which movement across a subject licenses a PG internal to a relative clause of that subject:

(39)   *PGs in relative clauses of subjects licensed by movement across the subject*
      a.   This is someone [who$_1$ **[people [that talk to PG$_1$]]** usually end up fascinated with t_1].
          (Adapted from Kayne 1983 ex. 15)

      b.   John's the guy $\varnothing_1$ that we invited because **[everyone [who talks to PG$_1$]]** is likely to appreciate t_1.
          (Nissenbaum 1998, ex. 34b)

Both subjects and relative clauses are generally taken to be islands in English. This means that in these sentences, the PG is separated from its licenser by two islands. As discussed in chapter 5, previous work on PGs has observed that they can only be separated from their licenser by one island (Culicover and Postal 2001). If two islands intervene, the PG is unacceptable, as repeated in (40) below.

(40)   *PG licensing across multiple islands normally fails*
      a.   *Relative clause island in adjunct island*
          * Tell me who$_1$ you talked to t_1 [after meeting a person [who likes PG$_1$]].

      b.   *Subject island in adjunct island*
          * Durian is a fruit [which$_1$ I ate t_1 for the first time [after [a fan of PG$_1$] visited me]].

For this reason, the configuration in (39) is exceptional.

In order to solidify the validity of this puzzle, is worth being certain that the dependency that forms the PG in such sentences is indeed movement. We can be sure of this by embedding the PG in an island in the relative clause, though in a position low enough to isolate the PG and exclude the relativizing movement chain. As (41) shows, this configuration is indeed unacceptable, as expected if operator movement forms the PG in this configuration in the usual way:

(41) *PG in relative clause of subject degraded by additional island*

  a. *Additional relative clause island*
  *John's a guy $\varnothing_1$ that [everyone who talks to **[people who know PG$_1$]**] gets a good impression of t_1.

  b. *Complex NP island in relative clause*
  *Tell me [which food]$_1$ [people [who research **[the rumors that the government contaminates PG$_1$]**]]] absolutely never choose to eat t_1.

Consequently, the variety of PG-containing relative clause in (39) appears to involve a PG of the usual sort, making the possibility of this configuration a mystery. While I cannot offer a decisive solution to this puzzle, I have several suggestions.

We have seen both in this chapter and chapter 5 that while a PG cannot be separated from its licenser by two islands, a PG in an embedded island becomes possible when the containing island also has a PG. This is because the operator movement that forms the PG in the outer island can itself serve an appropriately local licenser for the PG in the embedded one:

(42) *Recursive PG licensing in embedded islands*

  a. A person who$_1$ I hang out with t_1 [because [friends of PG$_1$] admire t_1] ([Nissenbaum 2000](), p. 26, ex. 13c)

  b. That's the manager which$_2$ I know [an employee of t_2 [who thinks that [every message from PG$_2$] reveals the stupidity of PG$_2$]]].

This fact raises the possibility that in the exceptional sentences of (39), there is in fact a PG in the subject DP, whose corresponding operator licenses the PG in the relative clause. Under such an analysis, the sentence in (39a) would in fact have the following structure, where the initial PG that licenses the embedded one is bolded:

(43) *Hypothetical recursive PG analysis of (39a)*
  This is someone [who$_1$ [people **PG$_1$** [that talk to PG$_1$]] usually end up fascinated with t_1].

The most natural place for this hypothetical PG in the subject would be in the complement of NP—a position that typically allows PGs in English. If this hypothesis is on the right track, we predict that filling this position with overt non-PG material will degrade the PG in the relative clause. In my judgment, this test reveals the expected contrast, as (44) shows:

(44) *Filled complement of subject degrades PG in relative clause of subject*

  a. That's the professor $\varnothing_1$ that [discussions about physics [that I had with PG$_1$]] usually ended up annoying t_1.

  b. ??* That's the professor $\varnothing_1$ that [friends **of Bill** [who spend too much time with PG_1]] always get annoyed by t_1.

  c. Guess who$_1$ [students of PG_1 [who research Geography with PG_1]] are causing a lot of trouble for t_1.

  d. ??* Guess who$_1$ [students **of mine** [who research Geography with PG_1]] are causing a lot of trouble for t_1.

Since this hypothetical PG in the subject does not obviously have an interpretation, however, an analysis along these lines remains incomplete.

   In chapter 5, I discussed the observation of Nissenbaum (2000) that in a multiple *wh*-question, each moved *wh*-phrase (the lower of which moves covertly) can license a separate PG in a sentential adjunct, as repeated below:

(45)   *Multiple wh-phrases licensing separate PGs in one adjunct*
       [Which senator]$_1$ did you persuade t_1 to borrow [which car]$_2$ [after getting an opponent of PG_1 to put a bomb in PG_2]?
       (Nissenbaum 2000, p. 12 ex. 8)

If a PG in a relative clause of a subject is not dependent on the presence of a PG in the subject as well, then when we unambiguously signal the presence of a PG both in the subject and in the relative clause, we expect to find that the two PGs might be licensed by separated moved phrases. This would be precisely analogous to the sentential adjunct in (45) above, which contains two PGs with distinct licensers. In my judgment, this is not possible, as (46) shows. In the acceptable (46a), one instance of *wh*-movement across the subject licenses a PG in it, which co-occurs with a co-referent PG in the subject's relative clause. Example (46b) differs in involving a multiple *wh*-question, and attempting to have the PG in the subject, and in the relative clause, be licensed by separate *wh*-phrases. Importantly in contrast to the configuration in (45) above, (46b) is unacceptable:

(46)   *A PG in a subject and its relative clause must be co-referent*

  a. [Which kid]$_1$ did [a relative of PG_1 [who really hates PG_1]] send t_1 a package full of dead bugs?

  b. * [Which kid]$_1$ did [a relative of PG_1 [who knows nothing about the ingredients in PG_2]] send t_1 [which cake]$_2$ for her birthday?

This result indicates that a PG in a relative clause of a subject that appears to be licensed by A′-movement across that subject is in fact locally licensed, via a distinct (though necessarily co-referent) PG in the subject. If a PG cannot be separated from its licenser by more than one island, this is precisely what we expect to find. As mentioned above, however, the fact that the hypothesized PG in the subject of

sentences like (39) above is not obviously semantically contentful is puzzling. For now, I must leave this topic for future investigation.[15]

7 Conclusion

In this chapter, I have analyzed configurations involving licensing of a PG in a relative clause by extraction from the same DP in which the relative clause is adjoined. I argued that, given independent semantic conclusions about the nature of (restrictive) relative clauses, that this fact indicates the possibility of successive cyclic movement from NP. This possibility is consistent with the concept that NP is a phase, which has precedent in previous work, though whether or not such movement is obligatory in the way that we expect if NP is a phase remains to be determined. I used the analysis of this construction to argue for several other syntactic conclusions—chiefly that overlapping operators form crossing paths, as proposed by Nissenbaum (2000), and that the theory of left branch extraction in Bošković (2005) is not quite correct. I also described a puzzle about a separate method of including a PG in a relative clause that requires further study.

In the following appendix, I offer some final thoughts about the distribution of PGs that have not had a natural place elsewhere in this work. The concluding remarks in the next chapter are the final component of this dissertation.

8 Appendix: Parasitic gaps beyond vP and NP

Under the theory of PGs in Nissenbaum (2000), PGs are dependent on merging local to a landing site of A′-movement, in order to take advantage of Predicate Abstraction triggered by that movement. In principle, this theory allows PGs to be licensed in either final or intermediate landing sites of A′-movement. In chapter 5, I summarized Nissenbaum's proposal that PGs in sentential adjuncts are licensed by successive cyclic movement through the vP phase. In this chapter, I have argued that PGs in relative clauses can be licensed by successive cyclic movement from NP (which may or may not be a phase). While DP is assumed to be a phase in

---

[15]As far as I know, Frampton (1990) is the only previous work to discuss the puzzle posed by the sentences examined in this section. Frampton does not assume the presence of a distinct D projection, and thus posits that a relative clause is not c-commanded by any nominal-internal material. In the context of this assumption, Frampton hypothesizes that sentences like (39) are possible because a subject NP containing a relative clause does not serve as a barrier for the mechanism that he argues is responsible for uniting the licensing movement chain with the parasitic chain in that relative clause. Since Frampton's analysis relies on a particular notion of *Government*, it is difficult to translate into contemporary assumptions, though it is likely incompatible with the DP hypothesis.

many works, I am not aware of a constituent that can both host a PG, and adjoin to DP edges, such that it would allow us to diagnose successive cyclic movement via spec-DP using a PG. Another constituent that is perhaps the strongest candidate for phasehood is CP, which we also expect successive cyclic movement from to have the ability to license a PG in a constituent adjoined to it.

I am not aware of any substantial discussion of this possibility in previous work. van Urk (2015) briefly mentions this expectation (footnote 40, p. 191) and suggests that PGs in conditionals could instantiate PG licensing under successive cyclic movement from CP, citing the following (modified) examples from Engdahl (1983):

(47)   *PG licensing in conditional adjuncts*
    a.   This is the professor $\varnothing_1$ that Kim says that you must not say hello to t_1 [if you run into PG_1].
    b.   This is the professor $\varnothing_1$ that Kim says that [if you run into PG_1], t_1 won't say hello to you.

We can use principle C to ensure that the PG-containing conditional adjunct is indeed not being interpreted adjoined to vP. If it were, an R-expression in the adjunct co-referent with the subject should be unacceptable, contrary to fact:

(48)   *PG-containing conditional is not interpreted within vP*
    Who_1 did you say (that) [if **Mary**$_2$ talks to PG_1] **she**$_2$'ll definitely throw a tomato at t_1?

However, we see in (47b) and (48) that a conditional in this position can be preceded by a complementizer. This raises the possibility that these PG-bearing constituents are in fact adjoined to TP, rather than CP.

However, as McCloskey (2006) points out, peripheral adjuncts of this variety can be sandwiched between two instances of the complementizer:

(49)   *Doubled complementizer surrounding adjunct*
    They know **that** in general **that** a jury is not going to be financially savvy.
    (McCloskey 2006, ex. 69e, citing BBC World Service News, April 5, 2004)

David Pesetsky (p.c.) points out that in examples like (48), the PG-containing adjunct can similarly appear in between two instances of the complementizer:

(50)   *PG-containing conditional is not interpreted within vP*
    Who_1 did you say **that** [if Mary$_2$ talks to PG_1] **that** she$_2$'ll definitely throw a tomato at t_1?

McCloskey argues that this complementizer doubling is the result of recursion of the

C projection. If this analysis is correct, then it may be the case that a PG in an adjunct in this position is indeed licensed under movement through spec-CP. Whether all or perhaps just some phrases in a layered CP structure count as phases will lead to different predictions about the distribution of successive cyclicity, however. Since I do not see a straightforward way of adjudicating between the different plausible hypotheses about this, I will end the discussion here.

Chapter 8

Final remarks

1 Main results

In this chapter, I will summarize the main concepts and empirical results that this dissertation has argued for, and discuss some additional consequences that the theory of grammar advanced here entails.

The most central component of this dissertation, which I used as a backdrop for exploring a variety of syntactic concepts and phenomena, is the CL theory of phases. CL shares with the theory developed in Chomsky (2000, 2001) and much following work the concept that syntactic derivations are punctuated by phase-by-phase applications of spell-out—the operation that (among other things) subjects the structure built so far to evaluation by the linear, morpho-phonological component of the grammar. CL differs, however, on a number of details about the nature of spell-out, and the laws that govern the information that it establishes. Unlike the phase theory initiated by Chomsky's work, CL hypothesizes that spell-out applies to all material present in a given phase, but does not make the elements it applies to inaccessible for subsequent syntactic operations. This theory thus permits movement of and from previously spelled-out constituents, but not without limitations. The constraining factor is the principle of Order Preservation, which requires the linearization information established by a given instance of spell-out to be preserved for the remainder of the derivation in question. By hypothesis, if a later instance of displacement contradicts a previously established ordering, then the result is an irreconcilable conflict which culminates in a crash at PF.

As we've seen, CL predicts that movement must generally pass through the linear edge of each phase crossed, since doing so appropriately linearizes the moving constituent as preceding the content of all phases it exits. CL thus leads us to expect that movement from phases will generally be successive cyclic. A great deal of

previous research in syntactic theory, as well as this dissertation, has shown that this is indeed the correct conclusion.

(1)   *Successive cyclic movement via linear edge of phase*

a.   * $[_{ZP}$ α Z $[_{YP_{[Phase]}}$ β Y $[_{XP}$ t $]]]$

b.   ✓ $[_{ZP}$ α Z $[_{YP_{[Phase]}}$ t β Y $[_{XP}$ t $]]]$

However, CL also predicts another way to move from a phase. In particular, this theory allows the possibility of a given element exiting a phase from a non-edge position, provided that any material within the relevant phase crossed over by that movement also moves on, to a position above the element that crossed it:

(2)   *Avoiding a linearization problem through order-restoring movement*

a.   *Illicit non-successive-cyclic phase exit...*
      * $[_{XP}$ β $[_{YP_{[Phase]}}$ α β $]]$

b.   *...repaired by order-restoring movement*
      ✓ $[_{XP}$ α β $[_{YP_{[Phase]}}$ α β $]]$

In this dissertation, I have argued that the general predictions of CL illustrated in the schemas in (1) and (2) above yield a wide variety of correct results about displacement and discontinuity in many languages.

However, we have also seen that CL alone is not sufficiently restrictive. Importantly, the majority of the empirical results of this dissertation emerge from combining the CL theory with several independently supported constraints on syntactic movement. Chief among these was the concept that movement from one edge position to another within the same phrase is banned (Ko 2007, 2011, 2014), as the diagrams in (3) below illustrate once more:
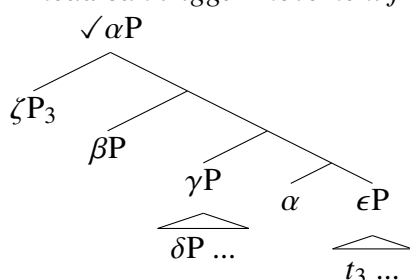
(3)   a.   *No phrase-bounded edge-to-edge movement*

$*\alpha$P
γP$_1$
βP
δP ...
t_1
α
ϵP
ζP ...
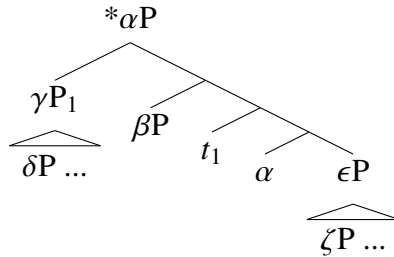
b. *No phrase-bounded extraction from within an edge constituent*

$$*\alpha P$$

$$\delta P_2$$

$$\beta P$$

$$\gamma P \quad \alpha \quad \epsilon P$$

$$t_2 \ldots$$

$$\zeta P \ldots$$
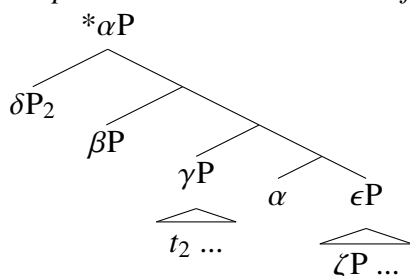
c. *A head can trigger movement from within its complement*

$$\checkmark \alpha P$$

$$\zeta P_3$$

$$\beta P$$

$$\gamma P \quad \alpha \quad \epsilon P$$

$$\delta P \ldots$$

$$t_3 \ldots$$

Following Ko, this constraint emerges from a theory under which movement requires a probing feature on a head to identify and thus attract a relevant goal in its c-command domain. Since a head does not c-command any elements in the edge of its projection, an attracting feature on that head cannot trigger movement of them.

At various points in this dissertation we have also seen the influence of anti-locality—the concept that movement must not be too short (Bošković 2005, 1997; Ishii 1999; Grohmann 2003; Abels 2003, 2012; Erlewine 2016, a.o.). A number of formulations of anti-locality are attested in the literature, some of which subsume the ban on edge-internal movement. While settling on a decisive theory of the set of anti-locality constraints has not been a goal of this dissertation, we have seen that effects of this variety importantly constrain the syntactic derivation in a number of ways that, when combined with CL, yield a wide variety of correct empirical results. I will summarize several of this dissertation's main results next, before considering a few more general implications of the theory developed here.

After the introduction in chapter 1, I argued in chapter 2 that these considerations predict a cross-linguistic generalization about stranding in intermediate positions:

(4) *Intermediate Stranding Generalization*
Leftward movement of a phrase α can only intermediately strand an element β if β is (or can be) ordered to the right of α before stranding occurs.

In chapter 3, I argued that the addition of a phase-level adjacency constraint on the Saxon genitive morpheme [*'s*] in colloquial English predicts the possibility of a form

of possessor extraction, which is subject to a particular but pervasive constraint:

(5)   *Generalization about possessor extraction in colloquial English*
        A possessor must reach the left linear edge of the local CP before extracting from the possessum DP.

In chapter 4, I argued that this perspective on the grammar predicts two generalizations about sub-extraction in Russian:

(6)   a.   *Left branch extraction order preservation generalization*
              Left branch extraction from a Russian nominal phrase containing multiple left branches must result in a relative linear ordering that matches the original order that those left branches had prior to any extraction.

        b.   *Sub-extraction from subject blocking generalization*
              Sub-extraction from a Russian subject (excluding some adjectives and possessors)[1] is banned when other material originating in vP scrambles to a position between the subject and the sub-extracted constituent.

In chapter 5, I considered the predictions of CL in combination with the theory of parasitic gaps in Nissenbaum (2000), based on which I argued for the generalization in (7) about overlapping movement chains. I argued that an examination of this generalization also leads to insights about the distribution of crossing versus nesting paths, and the mechanisms that drive movement:

(7)   *Overlapping Chain Generalization*
        If two A′-moving phrases XP_1 and XP_2 form specifiers of an intermediate phase YP such that XP_1 c-commands XP_2, then XP_1 c-commands XP_2 in all subsequent positions these phrases occupy.

In chapter 6, I argued that CL combined with a theory of extraposition as late merge to a (covertly or overtly) moved phrase (Fox and Nissenbaum 1999; Nissenbaum 2000; Fox 2002, a.o.) derives the generalization in (8) below. I collected evidence for this empirical result both from previous literature (Nissenbaum 2000; Fox and Pesetsky 2009; Jenks 2011, 2013a,b) and from my own recent fieldwork on Wolof and Javanese, and argued for an account of (8) that extends to several other puzzles about the distribution of extraposition.

---

[1]The exceptional status of adjective and possessor extraction described in (6b) does not automatically emerge from the central concepts of this dissertation. This issue is empirically complex and deserves further study, though as discussed in chapter 4, some version of a theory that allows these elements (which are plausibly both adjuncts) to be externally merged more flexibly (such as by late merge, or in a position external to a nominal phrase) can achieve the right result.

(8)   *Source ordering generalization*
        Extraposition of a non-argument constituent α from a source phrase XP must result in α having the same relative linear order with respect to the rest of the content of XP that it would have had in the absence of extraposition.

Finally, in chapter 7 I set aside CL in order to examine several of the concepts from chapter 5 by way of an under-analyzed configuration in English involving parasitic gaps in relative clauses. While the results of chapter 7 do not bear directly on the concerns of CL, they do enrich the understanding of certain concepts from chapter 5 that are tightly intertwined with the nature of phases and linearization. Chapter 7 also has implications for theories of sub-extraction and the phasehood of nominal constituents, both of which are highly relevant to the topics of this dissertation.

Overall, the results of this dissertation amass a new body of evidence for a theory in which the limitations on possible sentences are, to a significant degree, attributable to the interleaving of syntactic processes with the linear component of the grammar. I have argued that the above empirical generalizations also have a number of informative consequences for a variety of more specific phenomena. These are more numerous than can be summarized here, however. With the general results of this dissertation now summarized, next I will discuss some further significant implications of the CL theory that the previous chapters have not touched on.

2   Silent material and linearization

This dissertation has attributed a wide variety of phenomena to the way that linearization constrains the formation of syntactic structures. The concept of linear order is fundamentally about the temporal order in which a set of linguistic elements is expressed by the sensorimotor system—in particular, as phonetic signals produced by the vocal tract, in the case of spoken language. From this perspective, we do not expect the concerns of linearization to be applicable to linguistic elements that have no phonetic correspondent. For this reason, it is plausible that the effects that I have attributed to linearization in this dissertation will disappear when the elements that cause the hypothesized linearization issues in question are silent. As pointed out in several footnotes throughout this dissertation, this does not appear to be the case. Here I will provide some of the relevant evidence for this unexpected result, and discuss what can be concluded from it.

In chapter 2, I argued that the fact that intermediate stranding of prepositions is impossible even in languages that otherwise allow P-stranding (9a) is due to a linearization problem that arises when the phrase that extracts from the preposition crosses over it at a phase edge (9b):

(9)   a. *No intermediate preposition stranding*
            (For) [which dog]$_1$ did they claim [$_{CP}$ (*for) t_1 (that) I cooked an expensive steak **(for)** t_1]?

        b. *No intermediate stranding with crossing at the edge*
            * [$_{ZP}$ α Z [$_{YP[Phase]}$ [β $t_α$] Y [$_{XP}$ $t_{βα}$ X]]]

I argued that this conflict is just one instance of a broader empirical generalization about intermediate stranding, which was stated in (4) above. If this explanation is correct, we might predict intermediate preposition stranding to be possible when the element that strands the preposition is silent. This hypothesis can be tested in a context with movement of a null relative pronoun (10a) or null operator in a parasitic gap configuration (10b). It turns out that while these contexts allow base position preposition stranding, they do not permit intermediate stranding:

(10)   *Movement of a null element cannot intermediately strand a preposition*

        a. That's the person [∅$_1$ that I said [$_{CP}$ (*to) t_1 (that) we should send a nice letter **(to)** t_1]].

        b. Who did you excite t [∅$_1$ because you said [$_{CP}$ (*for) t_1 (that) you planned to buy a birthday present **(for)** t_1]]?

This result is the opposite of what the analysis in chapter 2 predicts, in the context of the hypothesis that silent material is irrelevant for linearization.[2]

---

[2]It is worth noting that full pied-piping is also banned in the scenarios in (10), as shown below:

(i)   a. That's the person [(*to) ∅$_1$ (that) I said [$_{CP}$ (that) we should sent a nice letter **(to)** t_1]].
        b. Who did you excite t [(*for) ∅$_1$ because you said [$_{CP}$ (that) you planned to buy a birthday present **(for)** t_1]]?

While the null operator that forms a parasitic gap is always silent, when we make the relative pronoun in a configuration like (ia) overt, full pied-piping is possible:

(ii)   That's the person [**(to)** which$_1$ I said [$_{CP}$ (that) we should send a nice letter **(to)** t_1]].

Consequently, it could be that intermediate stranding in (10) is ruled out because overt pied-piping with movement of a covert phrase is independently banned. This hypothesis only shifts where the question lies, however, since I cannot say what such a constraint would stem from.

   On this note, recall from section 2.2 of chapter 2 that *wh*-movement in English can facilitate intermediate stranding of certain adjuncts. We saw in this section that an in situ *wh*-phrase in a multiple *wh*-question cannot be construed as modified by such an adjunct which surfaces in an intermediate position, as repeated in (iii):

(iii)   a. [**How much flour**]$_1$ did you say t_1 (**to the nearest pound**) (that) they'll deliver t_1?
        b. * Who said (**to the nearest pound**) that they will deliver [**how much flour**]?

In chapter 3, we saw that possessor extraction in colloquial English cannot strand a non-subject possessum DP in situ, as repeated in (11) below: such a possessum must be pied-piped to at least the edge of the local CP. I argued that an interaction between CL and a condition mandating (phase-level) PF adjacency between the possessor and Saxon genitive [*'s*] predicts this fact.

(11)   *English possessor extraction cannot strand a non-subject in situ*
         a.   * Mary is the person [**who**$_1$ I heard [$_{CP}$ John ate [t_1**'s food**]]].
         b.   ✓ Mary is the person [**who**$_2$ I heard [$_{CP}$ [t_2**'s food**]$_1$ John ate t_1]].

As mentioned in footnote 15 of that chapter, the same constraint holds even when the possessor is not phonologically expressed. To see this, we can simply modify the examples in (11) to use a null relativizing pronoun (12):

(12)   *Extraction of null relative pronoun from non-subject possessum*
         a.   * Mary is the person [∅$_1$ that I heard [$_{CP}$ John ate [t_1**'s food**]]].
         b.   ✓ Mary is the person [∅$_2$ that I heard [$_{CP}$ [t_2**'s food**]$_1$ John ate t_1]].

If PF adjacency to the possessor is indeed a part of the explanation for this effect, then the fact that it applies even when the possessor is silent is unexpected: though as chapter 3 argues the relevant conception of PF adjacency is evaluated phase-by-phase and thus is not satisfied in a surface true way, it is unclear how PF adjacency can have been satisfied in (12b) if the possessor is simply absent at PF.

In chapter 5, I argued that the covertness of tucked-in specifiers in languages like English causes a linearization problem for configurations with overlapping overt movement paths, such that their paths must ultimately nest. This yields the effect described by the path containment condition of Pesetsky (1982):

(13)   *Nesting paths required for overlapping overt movements*
         a.   [This book]$_2$, who$_1$ should we talk to t_1 about t_2?   [2 1 ... 1 2]
         b.   * [This student]$_1$, what$_2$ should we talk to t_1 about t_2?   [1 2 ... 1 2]

Section 4.1 of chapter 5 argued that the linearization problem posed by a crossing paths derivation in such a context stems from the initially lower phrase tucking-

---

         c.   [**How many donuts**]$_1$ did you say t_1 (**to the nearest dozen**) (that) the bakery will give away t_1?
         d.   * [Which bakery] reported (**to the nearest dozen**) that the manager will give away [**how many donuts**]?

Given the evidence from chapter 5 that the in situ *wh*-phrase in such contexts in fact moves, this constraint is unexpected, unless something independent prevents overt pied-piping (and thus any subsequent intermediate stranding) with covert movement.

in below the originally higher moved phrase within vP, and moving on overtly afterwards. As mentioned in footnotes 24 and 28 of that chapter, we would expect this linearization problem to be avoided if the movement of the initially lower phrase could simply be entirely covert. In this case, this phrase could covertly tuck-in beneath the higher phrase prior to covertly moving on to its final landing site. These considerations lead us to expect that when we make the initially lower of two overlapping moved phrases covert, the path containment condition should not hold. In reality, it still does (14):

(14)   *Path containment condition not obviated by covertness*
  a.   This is the saxophone $\varnothing_2$ that I know [precisely which songs]$_1$ to never play t_1 on t_2.
  b.   * This is the saxophone $\varnothing_2$ that I know [precisely which songs]$_1$ to never use t_2 to play t_1.

In summary, several patterns that this dissertation has accounted for using the concepts of linearization persist even when the relevant elements are silent—a result that is surprising, if linearization is by definition a property of phonetically expressed material.[3] There are several ways of resolving this tension, which I discuss next.

One potential hypothesis is that the silent elements in question were originally phonetically present, but silenced at a later point in the derivation via whatever mechanism allows phenomena such as "Free Deletion in Comp" (Chomsky and Lasnik 1977). Such an explanation is conceivable for relative pronouns, which are optionally overt or covert, but not for null operators involved in the formation of parasitic gaps (10b), which are never overt.[4]

Alternatively, the relevance of silent material to patterns that appear to be governed by linear ordering could indicate that having phonetic content is not a prerequisite for being linearized. This proposal is natural in the context of the hypothesis that the linearization of syntactic terminals occurs prior to their assignment of morphophonological form (Embick 2010; Arregi and Nevins 2012; Ostrove 2018, a.o.). This hypothesis about the order of operations at spell-out has been posited, for instance, to explain the fact that linear adjacency is sometimes a prerequisite for morphological

---

[3]Similarly, Erlewine (2017) argues that the *that*-trace effect potentially stems from a conflict between the (anti)locality of subject extraction and CL, which is repaired if the complementizer is simply silent. Erlewine's account also predicts that if the extracted subject is instead silent, that there will be no *that*-trace violation. In my judgment, this prediction is incorrect:

(i)   * That's the person [who$_1$/$\varnothing_1$ I said [that t_1 killed the beavers]]!

[4]Not all theories take parasitic gaps to be the result of null operator movement, however. See Nissenbaum (2000) for a comparison of theories of parasitic gaps.

phenomena like contextual allomorphy, and the formation of portmanteau morphemes. If linearization precedes the assignment of morpho-phonological form, then it cannot be the case that linearization is endemic to pronounced terminals. This is because, under this theory, all terminals lack phonetic information at the time they are linearized. Given this understanding of linearization, the results of this dissertation can be maintained in a way that accurately predicts the relevance of covert elements. It is only necessary to understand the linearization conflicts argued for here as arising from ordering information about abstract syntactic terminals. While this ordering information will presumably serve as the input to decisions about where to insert morpho-phonological content, the operative interactions play out at a representational level that precedes the arrival of such content.[5]

It is also possible that the core patterns investigated in this dissertation stem from more general concerns of information preservation, but not linearization preservation per se. If relative linear order is generally determined based on the relative c-command relations between syntactic nodes (Kayne 1994), it is plausible that many of the effects examined in this dissertation are in fact about the preservation of relative structural relationships, which simply happen to generally correlate with the shape of the corresponding linear representations. An understanding along these lines is directly parallel to proposals in previous literature about structure preservation (Emonds 1970, 1976, a.o.) and shape preservation (Williams 1999, 2003; Müller 2002), though approaches in this vein differ in precisely what information is subject to preservation, and how that information is evaluated. See Fox and Pesetsky (2005a); Williams (2005) and Müller (2007) for relevant discussion. Many of the results of this dissertation would emerge in a way very similar to what I have proposed even if re-cast in terms of a theory in which what is recorded and preserved at the completion of each phase is information about the relative c-command relations of that phase's contents (perhaps in addition to linearization information).

---

[5]This hypothesis is potentially in conflict with the account of extraposition in chapter 6, which rejected the concept that covert material is ordered. However, what is most vital to this chapter is not the proposal that covert syntactic terminals are not linearized, but rather, the idea that the head of covert movement chains has no impact on linearization. As this chapter argues, rejecting the hypothesis of Fox and Nissenbaum (1999) that covert movement is linearized rightward opens the door to an account that determines the relative ordering of an extraposed phrase and its source only indirectly, which facilitates an understanding of some otherwise mysterious phenomena, such as the possibility of extraposition from overtly moved phrases.

The proposal that covert movement operations like QR have no affect whatsoever on linearization is especially natural under a theory of movement as multi-dominance, since under such an approach movement does not create new terminals that we must worry about the linearization of. The initial proposals about CL in Fox and Pesetsky (2005a,b) are in fact framed in terms of a multi-dominance approach, though I have not adopted such a view in this dissertation, since it does not offer an explanatory advantage for the vast majority of topics examined here.

3   On the (un)boundedness of agreement

As discussed at several points in this dissertation, the CL theory of phases differs from that in Chomsky (2000, 2001) and much following work on several details about the nature of spell-out. These theories differ not only in their predictions about movement, which I have focused on here, but also in their expectations for the limitations of other syntactic dependencies like agreement. In this section, I will briefly describe the predictions about agreement for both theories, in order to establish how this phenomenon's constraints can be understood in the context of CL.

For Chomsky (2000, 2001, a.o.), phase-level spell-out applies to the phase head's complement, which is then inaccessible to any later syntactic operations. This constraint is termed the Phase Impenetrability Condition (PIC), which has been proposed in at least two forms, which differ in timing. Under what is termed the *strong* PIC, as soon as a phase is completed, its complement spells-out and becomes inaccessible. This theory predicts that a head external to a given phase cannot, under any circumstances, agree with an element in the complement of that phase. This is diagrammed in (15) below, in which we see that agreement between a head X and a DP in the complement of its phasal sister YP is banned:

(15)   *Strong PIC: No agreement into phase complement*
        $* [_{XP}\ X_{[uF]}\ [_{YP_{[Phase]}}\ Y\ ...\ DP_{[F]}\ ]$

In contrast, under the so-called *weak* PIC, a phase does not spell-out until the next highest phase head is externally merged. Under this theory, the agreement diagrammed in (15) is predicted to be successful, since there is no phase head in this context aside from Y itself. This alternative version of the PIC provides a window between the completion of a phase, and spell-out within it, that opens up the possibility of agreement into (or movement from) a phase's complement.

The weak PIC has been used to account for well-known Icelandic facts like that in (16a) below, in which there is a dative subject and nominative object, and the verbal agreement typically reserved for subjects instead targets the object. Given an analysis under which the agreeing head is T, and the nominative object does not move, such facts are argued to constitute agreement into the complement of the vP phase, as diagrammed in (16b):

(16)   *Agreement with nominative object in Icelandic*
        a.   Henni   leiddust   þeir.
             she.DAT was.bored.by.3PL they.NOM
             'She was bored by them.'
             (Taraldsen 1995, p. 307)

344

b.　　$[_{TP}$　DP_{DAT}　$T_{[u\phi]}$　$[_{vP_{[Phase]}}$　v　...　$DP_{[\phi],NOM}$　$]]$

The occasional necessity of this weakening of the PIC reveals a tension between the actual distribution of agreement, and a strict concept of phase impenetrability. Importantly, this particular issue does not arise in the context of CL, for which there simply is no PIC. However, the fact that spelled-out material is never rendered inaccessible under CL leads to the prediction that there should, in principle, be no limitation on the distance between an agreeing head and its goal.

There are a considerable number of facts reported in the literature that appear to instantiate long distance agreement (Bruening 2001; Polinsky and Potsdam 2001; Bošković 2003; Bhatt 2005; Etxepare 2006; Keine 2019, a.o.). Many of these instances of agreement are even less local than the above Icelandic context, involving, for instance, agreement into an embedded clause. An example of this from Polinsky and Potsdam (2001) is provided in (17) below. Here we see that in the Nakh-Daghestanian language Tsez, a verb with a clausal complement can agree either with the clausal argument itself, or with the object of that embedded clause:

(17)　*Optional long distance agreement in Tsez*

　　a.　*Agreement with clausal argument*

　　　Enir　[užā magalu　　 bāc'ruɫi]　r-iyxo.
　　　mother　boy bread.III.ABS ate　　].**IV IV**-know

　　　'The mother knows the boy ate the bread.'

　　b.　*Agreement with object of embedded clause*

　　　Enir　[užā magalu　　 bāc'ruɫi]　b-iyxo.
　　　mother　boy bread.**III**.ABS ate　　].IV **III**-know

　　　'The mother knows the boy ate the bread.'
　　　(Polinsky & Potsdam 2001, ex. 1)

Importantly, many works on long distance agreement propose that, for one reason or another, the instance of agreement they are concerned with is not as long-distance as it appears to be. Polinsky and Potsdam (2001) argue, for instance, that the possibility of seemingly long distance agreement in the Tsez (17b) is mediated by embedded topicalization of the object, which brings it sufficiently local to the matrix verb. Thus while cross-phasal agreement may be possible to some extent, it is likely incorrect to conclude that agreement is simply unbounded in the way that CL predicts.

What would be compatible with CL is a theory in which agreement is indeed limited, but by independent factors that are unrelated to spell-out. A theory of this variety is proposed by Rackowski and Richards (2005), and several following works (van Urk and Richards 2015; Halpert 2016, 2019; Branan 2018). For research in this

vein, agreement into a phase's complement is possible in principle, but contingent on the possibility of first agreeing with the phase itself:

(18) *Agreement with phase permits agreement into its interior*
$[_{XP}$ X$_{[uF]}$ ... $[_{YP_{[Phase]}}$ Y ... DP$_{[F]}$ $]$
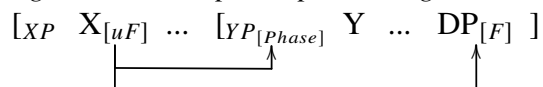
For this approach, phases are general interveners for agreement. Thus when a phase separates a potential goal from a head bearing a feature that would agree with that goal, straightforward locality considerations will force agreement with the phase itself to occur before the lower goal in question can be targeted. If the relevant agreeing probe is not of a variety that is capable of both agreeing with the phase in question and then probing further, then agreement into that phase is precluded.

Importantly, approaches of this variety eschew the PIC, and attribute the locality of agreement (as well as movement) to independent considerations about the interaction between phases and probe-goal relations. Such a perspective is natural in the context of CL. One potential point of tension, however, is that for many of the works just cited, agreement deep into a phase serves to permit extraction from a non-edge position within that phase. This is something that CL generally rules out, in the absence of additional order-restoring movement processes. There are therefore certain details of implementation that must be worked out before CL and the approach to agreement just summarized can be properly united.[6] However, the important point here is that there is the possibility of a theory that divorces the locality constraints of agreement from the nature of spell-out. Given CL, we expect such an approach to agreement to be an accurate one.[7]

---

[6]For some discussion of how these concepts might be unified, see Branan and Davis (2019).

[7]Keine (2016, 2019, 2020) argues that agreement-triggering probes have characteristic limits, dubbed *horizons*, which are independent of the concerns of phase theory. This concept is entirely compatible with the general approach to agreement necessitated by CL.

Keine also argues that while CP systematically blocks agreement, vP never does, and correspondingly argues that vPs are not phases. Given what I have said in this section, another possibility is that vP is a phase that uniquely lacks the trait that would cause it to serve as an intervener for agreement. Rackowski and Richards (2005) argue that CP and DP are both phases which are opaque for agreement unless first agreed-with. Note that CP and DP differ from vP in that the former two are able to serve as arguments. I speculate that if embedded CPs are generally in a sense nominalized, CP and DP may behave as a natural class as far as intervention for agreement is concerned because they are both potential bearers of ϕ-features, whereas this is never the case for vPs.

4   Conclusion

In this chapter, I have summarized the main results of this dissertation, and discussed two more general topics of importance for the CL theory defended here. First, I showed that several phenomena that I have attributed to the concerns of linearization persist even when the relevant elements are covert. I made several suggestions about why this might be so. In particular, a grammar in which linearization precedes the insertion of morpho-phonological information, or in which preservation is evaluated in terms of structural relations perhaps in addition to linear ordering, will essentially achieve the results I have argued for here. Second, I made explicit the predictions about CL for agreement, which under this theory should be in principle unbounded. As I have discussed, there is precedent for a theory in which agreement is not bounded by spell-out, but by other unrelated locality concerns. Both of these topics present a number of potential angles for future work. This must wait for another time, however, since this dissertation is now complete.

4.1   The last final remark

In general, the results of this dissertation constitute evidence for a theory under which derivations proceed in discrete cycles or "chunks", and for which the information established at the end of each cycle must be respected for the remainder of the derivation. Provided that the arguments for such a theory are successful, there is much more that can be asked about the precise characteristics of the relevant cycles, and the nature of the information preservation requirement. I hope to have shown, however, that a cyclic grammar sensitive to the preservation of at least linear order information has a number of advantages.

Bibliography

Abels, Klaus. 2003. *Successive-Cyclicity, Anti-locality, and Adposition Stranding*. Doctoral dissertation, University of Connecticuit.

Abels, Klaus. 2012. *Phases: An Essay on Cyclicity in Syntax*. Berlin: De Gruyter.

Abney, Steven. 1987. *The Noun Phrase in its Sentential Aspect*. Doctoral dissertation, MIT.

Adger, David. 2003. *Core Syntax: A Minimalist Approach*. Oxford University Press.

Aelbrecht, Lobke. 2010. *The syntactic licensing of ellipsis*. Amsterdam: John Benjamins.

Aissen, Judith. 1996. Pied-piping, abstract agreement, and functional projections in Tzotzil. *Natural Language and Linguistic Theory* 14:447–491.

Akmajian, Adrian. 1975. More evidence for an NP cycle. *Linguistic Inquiry* 6:115–129.

Alexiadou, Artemis, Elena Anagnostopoulou, and Susi Wurmbrand. 2014. Movement vs. long distance Agree in raising: Disappearing phases and feature valuation. In *Proceedings of the 43rd meeting of the North Eastern Linguistics Society*, ed. Hsin-Lun Huang, Ethan Poole, and Amanda Rysling, 1–12. University of Massachusetts, GLSA.

Antonenko, Andrei, and Colin Davis. 2020. *Order Preservation in the Russian Nominal Phrase*. Poster presented at the 94th annual meeting of the Linguistic Society of America.

Arregi, Karlos, and Andrew Nevins. 2012. *Morphotactics: Basque Auxiliaries and the Structure of Spellout*. Dordrecht: Springer.

Asakawa, Teruo. 1979. Where does the extraposed element move to? *Linguistic Inquiry* 10:505–508.

Bailyn, John. 2012. *The Syntax of Russian*. Cambridge University Press.

Bailyn, John F. 1995. Underlying phrase structure and" short" verb movement in

Russian. *Journal of Slavic linguistics* 13–58.

Baltin, Mark. 2006. Extraposition. In *The Blackwell Companion to Syntax, Volume I*, ed. Martin Everaert and Henk van Riemsdijk, 237–271. Blackwell.

Baltin, Mark. 2017. Strict bounding. In *The logical problem of language acquisition*, ed. C. L. Baker and John J. McCarthy, 257–295. Cambridge, Mass.: MIT Press.

Barbiers, Sjef. 1995. *The Syntax of Interpretation*. HIL dissertations 14.

Barbiers, Sjef. 2002. Remnant stranding and the theory of Movement. In *Dimensions of movement: From features to remnants*, ed. Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers, and Hans-Martin Gärtner. John Benjamins. Https://doi.org/10.1075/la.48.04bar.

Barss, Andrew. 1986. *Chains and anaphoric dependence: On reconstruction and its implications*. Doctoral dissertation, MIT.

Bašic, Monika. 2008. On Nominal Subextractions in Serbian. *Balkanistika* 21:1–56.

Bašic, Monika. 2009. Left Branch Extraction: A Remnant Movement Approach. In *Novi sad generative syntax workshop 2: Proceeds, august 28-29, 2008*, ed. L. Subotić, 39–51. University of Novi Sad: Novi Sad.

Bayırlı, İsa. 2017. *The universality of concord*. Doctoral dissertation, MIT.

Beermann, Dorothee A. 1997. *Syntactic discontinuity and predicate formation*. Doctoral dissertation, Tilburg University.

Bhatt, Rajesh. 2005. Long Distance Agreement in Hindi-Urdu. *Natural Language & Linguistic Theory* 23:757–807.

Bianchi, Valentina. 2000. The raising analysis of relative clauses: A reply to Borsley. *Linguistic Inquiry* 31:123–140.

Biberaur, Theresa, and Marc Richards. 2005. Explaining *Expl*. In *The function of function words and functional categories*, ed. Marcel den Dikken Christina Tortora, 115–154. Amsterdam: Benjamins. Https://doi.org/10.1075/la.78.06ric.

Billings, Loren, and Catherine Rudin. 1996. Optimality and Superiority: A new approach to overt multiple-*wh* ordering. In *Proceedings of Formal Approaches to Slavic Linguistics 3*, 35–60. Ann Arbor: Michigan Slavic Publications.

Bobaljik, Jonathan, and Susi Wurmbrand. 2005. The domain of agreement. *Natural Language & Linguistic Theory* 23:809–965.

Bobaljik, Jonathan, and Susi Wurmbrand. 2013. Suspension across Domains. In *Distributed Morphology Today: Morphemes for Morris Halle*, ed. Ora Matushansky and Alec Marantz, 289–315. Cambridge, MA: MIT Press.

Https://doi.org/10.7551/mitpress/9780262019675.001.0001.

Bobaljik, Jonathan D. 1995. *Morphosyntax: The Syntax of Verbal Inflection.* Doctoral dissertation, Massachusetts Institute of Technology.

Bondarenko, Tatiana, and Colin Davis. 2019. Parasitic gaps diagnose concealed pied-piping in Russian. In *Proceedings of North East Linguistic Society 49.* UMass Amherst: GLSA Publications.

Bondarenko, Tatiana, and Colin Davis. 2020. *Concealed pied-piping in Russian.* Manuscript under review.

Bondarenko, Tatiana, and Colin Davis. To appear. *Linearization asymmetries in Russian left branch extraction.* Proceedings of Formal Approaches to Slavic Linguistics (FASL) 28.

Borsley, Robert. 1997. Relative Clauses and the Theory of Phrase Structure. *Linguistic Inquiry* 28:629–647.

Borsley, Robert, and Ewa Jaworska. 1998. A Note on Prepositions and Case Marking in Polish. *Linguistic Inquiry* 19:685–691.

Bošković, Željko. 1997. *The Syntax of Nonfinite Complementation: An Economy Approach.* Cambridge, MA: MIT Press.

Bošković, Željko. 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena.* Elsevier Science, Amsterdam.

Bošković, Željko. 2002. On multiple wh-fronting. *Linguistic Inquiry* 33:351–383.

Bošković, Željko. 2003. Agree, phases and intervention effects. *Linguistic Analysis* 33:54–96.

Bošković, Željko. 2004. Be careful where you float your quantifiers. *Natural Language & Linguistic Theory* 22. Https://doi.org/10.1007/s11049-004-2541-z.

Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studica Linguistica* 59.

Bošković, Željko. 2014. Now I'm a Phase, Now I'm Not a Phase: On the Variability of Phases with Extraction and Ellipsis. *Linguistic Inquiry* 45:27–89. Https://doi.org/10.1162/LING_a_00148.

Bošković, Željko. 2015. On multiple left-branch dislocation: Multiple extraction and/or scattered deletion. *Proceedings of Formal Approaches to Slavic Linguistics 23* 20–35.

Bošković, Željko. 2016. Getting really edgy: On the edge of the edge. *Linguistic Inquiry* 45. Https://doi.org/10.1162/LING_a_00203.

Bošković, Željko. 2018. On Movement out of Moved Elements, Labels, and Phases. *Linguistic Inquiry* 49. Https://doi.org/10.1162/LING_a_00273.

Branan, Kenyon. 2017. In-situ Wh-phrases in Superiority Violating Contexts don't have to be In-Situ. In *A Pesky Set: Papers for David Pesetsky*, ed. Claire Halpert, Hadas Kotek, and Coppe van Urk, 353–359. Cambridge, MA: MITWPL.

Branan, Kenyon. 2018. Attraction at a Distance: A′-Movement and Case. *Linguistic Inquiry* 49:409–440.

Branan, Kenyon, and Colin Davis. 2019. Agreement and unlocking at the edge. In *Proceedings of the Linguistic Society of America*, volume 4, 1–15. Linguistic Society of America. Https://doi.org/10.3765/plsa.v4i1.4512.

Bresnan, Joan. 1972. *Theory of complementation in english syntax*. Doctoral dissertation, MIT.

Brillman, Ruth, and Aaron Hirsch. 2016. An anti-locality account of English subject/non-subject asymmetries. In *Proceedings of Chicago Linguistic Society 50*, ed. Ross Burkholder, Carlos Cisneros, and Emily R. Coppess. Chicago Linguistic Society.

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

Bruening, Benjamin. 2001. *Syntax at the edge: Cross-clausal phenomena and the syntax of Passamaquoddy*. Doctoral dissertation, Massachusetts Institute of Technology.

Bruening, Benjamin. 2006. Differences between wh-scope marking and wh-copy constructions in Passamaquoddy. *Linguistic Inquiry* 37:25–49.

Bruening, Benjamin. 2018. CPs Move Rightward, Not Leftward. *Syntax* 21:362–401.

Büring, Daniel, and Katharina Hartmann. 1997. Doing the Right Thing - Extraposition as a Movement Rule. *The Linguistic Review* 14:1–42.

Charnavel, Isabelle, and Dominique Sportiche. 2016. Anaphor Binding: What French Inanimate Anaphors Show. *Linguistic Inquiry* 47:35–87. Https://doi.org/10.1162/LING_a_00204.

Chen, Sherry Yong. 2017. Movement constraints on the relative order of double topics in mandarin chinese. In *Toronto working papers in linguistics*, ed. Ruth Maddeaux, volume 38, 1–21. Linguistics Graduate Course Union (LGCU).

Chomsky, Noam. 1973. Conditions on Transformations. In *A Festschrift for Morris Halle*, ed. Stephen Andersopn and Paul Kiparsky, 232–288. New York: Holt.

Chomsky, Noam. 1977. On wh-movement. In *Formal syntax*, ed. Adriam Akmajian,

Peter Culicover, and Thomas Wasow, 71–132. New York: Academic Press.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press. Linguistic Inquiry Monographs.

Chomsky, Noam. 1992. A minimalist program for linguist theory. In *MIT occasional papers in linguistics #1*. Cambridge, MA: MITWPL.

Chomsky, Noam. 1995. *The minimalist program.* MIT Press.

Chomsky, Noam. 2000. Minimalist Inquiries. In *Step by step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. Roger Martin, David Michales, Juan Urigareka, and Samuel Jay Keyser, 89–155. MIT Press.

Chomsky, Noam. 2001. Derivation by Phase. In *Ken hale: A life in language*, ed. Michael Kenstowicz. MIT Press.

Chomsky, Noam. 2005. Foundational issues in linguistic theory: Essays in honor of john roger vergnaud. In *On phases*, ed. Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 133–166. MIT Press, Cambridge. 10.7551/mitpress/9780262062787.001.0001.

Chomsky, Noam. 2008. On Phases. In *Foundational issues in linguistic theory: essays in honor of jean-roger vergnaud*, ed. Robert Freidin, Carlos P Otero, and Maria Luisa Zubizarreta. MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. Https://doi.org/10.1016/j.lingua.2012.12.003.

Chomsky, Noam, Ángel Gallego, and Dennis Ott. 2019. Generative grammar and the faculty of language: insights, questions, and challenges. *Catalan Journal of Linguistics Special Issue* 229–261.

Chomsky, Noam, and Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8:425–504.

Chung, Sandra. 1991. Functional heads and proper government in Chamorro. *Lingua* 85:85–134.

Chung, Sandra. 2013. Syntactic identity in sluicing: How much and why. *Linguistic Inquiry* 44:1–44.

Citko, Barbara. 2014. *Phase Theory: An Introduction*. Cambridge University Press, Cambridge: UK.

Cole, Peter, and Gabriella Hermon. 2000. Partial wh-movement: Evidence from Malay. In *Wh-Scope Marking*, ed. Gereon Müller and Armin von Stechow, 101–130. Amsterdam: Benjamins.

Collins, Chris. 1994. Economy of Derivation and the Generalized Proper Binding Condition. *Linguistic Inquiry* 25.

Comorovski, Ileana. 1986. Multiple WH Movement in Romanian. *Linguistic Inquiry* 17:171–177.

Coon, Jessica. 2009. Interrogative Possessors and the Problem with Pied-Piping in Chol. *Linguistic Inquiry* 40:165–175.

Corver, Norbert. 1990. *The Syntax of Left Branch Extractions*. Doctoral dissertation, Tilburg University.

Corver, Norbert. 1992. Left branch extraction. In *Proceedings of North East Linguistic Society 22*, ed. K. Broderick. GLSA Publications: UMass Amherst.

Corver, Norbert. 2007. Subextraction. In *The Blackwell Companion to Syntax*, ed. Martin Everaert and Henk van Riemsdijk. Wiley-Blackwell.

Corver, Norbert. 2017. Freezing Effects. In *The Wiley Blackwell Companion to Syntax*, ed. Martin Everaert and Henk van Riemsdijk. Hoboken, New Jersey: John Wiley and Sons, Inc. 10.1002/9781118358733.

Culicover, Peter, and Paul Postal, ed. 2001. *Parasitic Gaps*. Cambridge, MA: MIT Press.

Culicover, Peter, and Michael Rochemont. 1990. Extraposition and the Complement Principle. *Linguistic Inquiry* 21:23–47.

Culicover, Peter, and Michael Rochemont. 1997. Deriving Dependent Right Adjuncts in English. In *Rightward movement*, ed. Dorothee Beermann, David LeBlanc, and Henk van Riemsdijk. Philadelphia: Benjamins.

Davis, Colin. 2019. Possessor extraction in colloquial English: Evidence for successive cyclicity and cyclic linearization. *Linguistic Inquiry* just accepted. Https://doi.org/10.1162/ling_a_00369.

Davis, Colin. 2020. Crossing and stranding at edges: On intermediate stranding and phase theory. *Glossa* 5:1–32. Http://doi.org/10.5334/gjgl.854.

Deal, Amy Rose. 2009. The Origin and Content of Expletives: Evidence from 'Selection'. *Syntax* 12:285–323. Https://doi.org/10.1111/j.1467-9612.2009.00127.x.

Deal, Amy Rose. 2015a. *Interaction and satisfaction in ϕ-agreement*. Handout of talk presented at Interaction of Grammatical Building Blocks colloquium, Universität Leipzig.

Deal, Amy Rose. 2015b. Interaction and satisfaction in ϕ-agreement. In *Proceedings of North East Linguistic Society 45*, ed. Thuy Bui and Deniz Ozyıldız, volume 1, 179–192. Amherst, MA: Graduate Linguistics Student Association.

Deal, Amy Rose. 2016. Cyclicity and Connectivity in Nez Perce Relative Clauses. *Linguistic Inquiry* 47:427–470. Https://doi.org/10.1162/LING_a_00218.

Den Dikken, Marcel. 2007. Phase Extension: Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33:1–41. Https://doi.org/10.1515/TL.2007.001.

Den Dikken, Marcel. 2009. Arguments for successive-cyclic movement through SpecCP: A critical review. *Linguistic Variation Yearbook* 9:89–126.

Despić, Miloje. 2011. *Syntax in the absence of determiner phrase*. Doctoral dissertation, University of Connecticut.

Despić, Miloje. 2013. Binding and the Structure of NP in Serbo-Croatian. *Linguistic Inquiry* 44:239–270. Https://doi.org/10.1162/LING_a_00126.

Doetjes, Jenny Sandra. 1997. *Quantifiers and selection: On the distribution of quantifying expressions in French, Dutch, and English*. Doctoral dissertation, Leiden University.

Doggett, Teal B. 2004. *All Things Being Equal: Locality in Movement*. Doctoral Dissertation, MIT.

Doliana, Aaron. 2019. *Wh-quantifier float in German diagnoses A′-traces and successive cyclicity in vP*. Handout from talk given at NELS 50.

Doron, Edit, and Caroline Heycock. 1999. Filling and licensing multiple specifiers. In *Specifiers: Minimalist perspectives*, ed. D Adger, S Pintzuk, B Plunkett, and G Tsoulas, 69–89. Oxford University Press: Oxford.

Dowty, David, and Belinda Brodie. 1984. A Semantic Analysis of 'Floated' Quantifiers in a Transformationless Grammar. In *Proceedings of WCCFL 3*, ed. Mark Cobler, Susannah MacKaye, and Michael T. Wescoat, 75–90. Stanford: Stanford Linguistics Association.

Du Plessis, Hans. 1977. WH Movement in Afrikaans. *Linguistic Inquiry* 8.

Embick, David. 2010. *Localism versus Globalism in Morphology and Phonology*. MIT Press, Cambridge.

Embick, David, and Alec Marantz. 2008. Architecture and Blocking. *Linguistic Inquiry* 39:1–53. Https://doi.org/10.1162/ling.2008.39.1.1.

Emonds, Joseph. 1970. *Root and structure-preserving transformations*. Doctoral dissertation, MIT.

Emonds, Joseph. 1976. *A Transformational Approach to English Syntax: Root, Structure Preserving, and Local Transformations*. Academic Press Inc.

Engdahl, E. 1983. Parasitic Gaps. *Linguistics and Philosophy* 6.

Engelhardt, M, and H Trugman. 1986. D as a source of adnominal genitive in Russian. In *Proceedings of Formal Approaches to Slavic Linguistics 6*, ed. Ž Bošković, S Franks, and W Snyder, 114–133. Ann Arbor: Michigan Slavic Publications.

Epstein, Samuel David. 1992. Derivational constraints on A′-chain formation. *Linguistic Inquiry* 23:235–259.

Erlewine, Michael Y., and Hadas Kotek. 2018. Focus association by movement: Evidence from binding and parasitic gaps. In *Proceedings of Sinn und Bedeutung*, 1–9.

Erlewine, Michael Yoshitaka. 2016. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language & Linguistic Theory* 34. Https://doi.org/10.1007/s11049-015-9310-z.

Erlewine, Michael Yoshitaka. 2017. Why the Null Complementizer is Special in Complementizer-Trace Effects. In *A Pesky Set: Papers for David Pesetsky*, ed. Claire Halpert, Hadas Kotek, and Coppe van Urk, 232–288. Cambridge, MA: MITWPL.

Etxepare, Ricardo. 2006. Number long distance agreement in (substandard) Basque. In *Supplements of the anuario del seminario de filología vasca julio de urquijo*, ed. Joseba A. Lakarra and José I. Hualde, volume 40, 303–350. International Journal of Basque Linguistics and Philology.

Faneslow, Gisbert, and Damir Ćavar. 2002. Distributed deletion. In *Theoretical approaches to universals*, ed. Artemis Alexiadou. Amsterdam: John Benjamins. Https://doi.org/10.1075/la.49.05fan.

Fanselow, G., and D. Lenertová. 2011. Left peripheral focus: mismatches between syntax and information structure. *Natural Language & Linguistic Theory* 29:169–209. Https://doi.org/10.1007/s11049-010-9109-x.

Fanselow, Gisbert, and Caroline Féry. 2013. A comparative perspective on intervention effects on left branch extractions in Slavic. In *Non Progredi Est Regredi. Festschrift fur Alla Paslwaska*, ed. Wolodymyr Sulym, Mychajlo Smolij, and Chrystyna Djakiw, 266–295. Pais: Lviv.

von Fintel, Kai, and Sabine Iatridou. 2019. Since ~~since~~. In *The semantics of plurals, focus, degrees, and times: Essays in honor of Roger Schwarzschild*, ed. Daniel Altshuler and Jessica Rett, 305–333. Springer. Https://doi.org/10.1007/978-3-030-04438-1_15.

Fitzgibbons, Natalia. 2012. Pied-pipe your preposition and strand it too. *Snippets*

26. Https://doi.org/10.7358/snip-2012-026-fitz.

Fitzpatrick, Justin Michael. 2006. *The Syntactic and Semantic Roots of Floating Quantification*. Doctoral dissertation, MIT.

Fox, Danny. 2002. Antecedent-Contained Deletion and the Copy Theory of Movement. *Linguistic Inquiry* 33:63–96.

Fox, Danny. 2017. *Embedded Late Merge and the Theory of Displacement*. Manuscript, MIT.

Fox, Danny, and John Nissenbaum. 1999. Extraposition and scope: A case for overt QR. In *Proceedings of WCCFL 18*, ed. Sonya Bird, Andrew Carnie, Jason D. Haugen, , and Peter Norquest, 132–144. Somerville, Mass.: Cascadilla Press.

Fox, Danny, and John Nissenbaum. 2018. *Pied-piping and covert movement: Evidence from parasitic gap licensing*. Talk given at Workshop on Altaic Formal Linguistics 14: Workshop in honor of Shigeru Miyagawa.

Fox, Danny, and David Pesetsky. 2005a. Cyclic Linearization and its interaction with other aspects of the grammar. *Theoretical Linguistics* 31. Https://doi.org/10.1515/thli.2005.31.1-2.235.

Fox, Danny, and David Pesetsky. 2005b. Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics* 31. Https://doi.org/10.1515/thli.2005.31.1-2.1.

Fox, Danny, and David Pesetsky. 2009. *Rightward Movement, Covert Movement, and Cyclic Linearization*. Handout from presentation at Ben Gurion University.

Frampton, John. 1990. Parasitic gaps and the theory of wh-chains. *Linguistic Inquiry* 21:49–77.

Frampton, John. 2001. *Copies, traces, occurrences, and all that: Evidence from Bulgarian multiple wh-phenomena*. Manuscript, Northeastern University. Http://mathserver.neu.edu/ ling/pdf/CopiesTraces.pdf.

Franks, Steven. 1995. *Parameters of Slavic morphosyntax*. Oxford Universty Press.

Franks, Steven, and Asya Pereltsvaig. 2004. D as a source of adnominal genitive in Russian. In *Proceedings of Formal Approaches to Slavic Linguistics 12*, ed. O. Arnaudova, 109–128. Ann Arbor: University of Michigan Press.

Franks, Steven, and Ljiljana Progovac. 1994. On the placement of Serbo-Croatian clitics. *Indiana Linguistic Studies* 7:69–78.

Gallego, Ángel J. 2010. *Phase theory*. John Benjamins. Https://doi.org/10.1075/la.152.

Gavruseva, Elena. 2000. On the Syntax of Possessor Extraction. *Lingua* 110.

Gavruseva, Elena, and Rosalind Thornton. 2001. Getting It Right: Acquisition of Whose-Questions in Child English. *Language Acquisition* 9.

Goncharov, Julie. 2015. P-doubling in split PPs and information structure. *Linguistic Inquiry* 46:731–742. Https://doi.org/10.1162/LING_a_00199.

Graschchenkov, P.V. 2016. On the Syntactic Structure of Shifted Attributes (k voprosu o sintaksičeskoj strukture obsoblennyx atributov). In *Proceedings of the V.V. Vinogradov Russian Language Institute 2016, No. 10 Materials of International Scientific Conference (Grammatical Processes and Systems in Synchrony and Diachrnony)*, 70–81. Sholokhov Moscow State University for Humanities.

Graschchenkov, P.V. 2018. *Grammar of the adjective: The typology of adjective-ness and attributivity (grammatika progatel'nogo: Tipologia adëktivnosti i atributivnosti)*. LRC Publishing House Moscow.

Gribanova, Vera. 2013. Verb-stranding verb phrase ellipsis and the structure of the russian verbal complex. *Natural Language & Linguistic Theory* 31:91–136.

Gribanova, Vera. 2017. Head movement and ellipsis in the expression of russian polarity focus. *Natural Language & Linguistic Theory* 35:1079–1121.

Gribanova, Vera, and Boris Harizanov. 2018. Wither head movement? *Natural Language & Linguistic Theory* .

Groat, Erich, and John O'Neil. 1996. Spell-out at the LF interface. In *Minimal ideas*, ed. Werner Abraham, Samuel D. Epstein, Höskuldur Thráinsson, and Jan-Wouter Zwart. Amsterdam: John Benjamins.

Grohmann, Kleanthes. 2003. *Prolific domains: On the anti-locality of movement dependencies*. Amsterdam: John Benjamins.

Grohmann, Kleanthes, Markus Pöchtrager, Tobias Scheer, Michael Schiffmann, and Neven Wenger. 2017. The Apex Paradox. *Snippets* 31.

Guéron, Jacqueline. 1980. On the syntax and semantics of PP extraposition. *Linguistic Inquiry* 11:637–678.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The View From Building 20*, ed. Ken Hale and Samuel Jay Keyser, 1–52. MIT Press.

Halpert, Claire. 2016. *Argument licensing and agreement*. Oxford University Press, New York.

Halpert, Claire. 2019. Raising, unphased. *Natural Language & Linguistic Theory* 37:123–165.

Harley, Heidi, and Rolf Noyer. 1999. Distributed morphology. *Glot International*

4:3–9.

Harwood, William. 2015. Being progressive is just a phase: Celebrating the uniqueness of progressive aspect under a phase-based analysis. *Natural Language & Linguistic Theory* 33:523–573. Https://doi.org/10.1007/s11049-014-9267-3.

Heck, Fabian. 2009. On Certain Properties of Pied-Piping. *Linguistic Inquiry* 40:75–111.

Heck, Fabian, and Malte Zimmermann. 2004. *DPs and Phases*. Ms., Universitat Leipzig and HU Berlin.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Henry, Alison. 2012. Phase Edges, Quantifier Float, and the Nature of (Micro-) Variation. *Iberia: An International Journal of Theoretical Linguistics* 4:23–39.

Henry, Alsion. 1995. *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford,UK: Oxford University Press.

Holmberg, Anders. 1986. *Word order and syntactic features in the scandinavian languages and english*. Doctoral dissertation, University of Stockholm.

Holmberg, Anders, Michelle Sheehan, and Jenneke van der Wal. 2019. Movement from the Double Object Construction Is Not Fully Symmetrical. *Linguistic Inquiry* 50:677–721. Https://doi.org/10.1162/ling_a_00322.

Honcoop, Martinus F. 1998. *Dynamic excursions on weak islands*. Doctoral dissertation, Leiden University.

Hooper, Joan, and Sandra Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4:465–497.

Horvath, Julia. 2006. Pied-Piping. In *The Blackwell Companion to Syntax, Volume I*, ed. Martin Everaert and Henk van Riemsdijk. Blackwell Publishing.

Huck, Geoffrey J., and Younghee Na. 1990. Extraposition and focus. *Language* 66:55–77.

Ishii, Toru. 1999. Cyclic spell-out and the *that*-trace effect. In *Proceedings of WCCFL 18*, ed. A. Carnie, J. D. Haugen, and P. Norquest, 220–231. Somerville, MA: Cascadilla Press.

Jenks, Peter. 2011. *The Hidden Structure of Thai Noun Phrases*. Doctoral dissertation, Harvard University.

Jenks, Peter. 2013a. *Accounting for a Generalization about Quantifier Float and Word Order in Classifier Languages*. Handout from presentation at SEALS 23,

May 30.

Jenks, Peter. 2013b. Quantifier Float, Focus, and Scope in Thai. In *Proceedings of Berkeley Linguistics Society 39*, 90–107. Linguistic Society of America. Http://dx.doi.org/10.3765/bls.v39i1.3872.

Johnson, Kyle. 2012. Towards deriving differences in how Wh-Movement and QR are pronounced. *Lingua* 122:529–533. Https://doi.org/10.1016/j.lingua.2010.11.010.

Kayne, Richard. 1983. Connectedness. *Linguistic Inquiry* 14:223–249.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press Linguistic Inquiry Monographs.

Kayne, Richard. 2002. On some prepositions that look DP-internal: English *of* and French *de*. *Catalan Journal of Linguistics* 1.

Keine, Stefan. 2016. *Probes and their Horizons*. Doctoral Dissertaion, University of Massachusetts Amherst.

Keine, Stefan. 2019. Selective Opacity. *Linguistic Inquiry* 50:13–62. Https://doi.org/10.1162/ling_a_00299.

Keine, Stefan. 2020. *Probes and Their Horizons*. MIT Press Linguistic Inquiry Monographs.

Kitahara, Hisatsugu. 1994. Restricting Ambiguous Rule-Application. In *Formal Approaches to Japanese Linguistics 1*, ed. Masatoshi Koizumi and Hiroyuki Ura, 179–209. Cambridge Mass.: MITWPL.

Kitahara, Hitatsugu. 1991. *Target-α: A unified thecky of movement and structure-building*. Doctoral Dissertaion, Harvard University.

Kitahara, Hitatsugu. 1997. *Elementary operations and optimal derivations*. Cambridge, MA: MIT Press.

Ko, Heejeong. 2007. Asymmetries in Scrambling and Cyclic Linearization. *Linguistic Inquiry* 38. Https://doi.org/10.1162/ling.2007.38.1.49.

Ko, Heejeong. 2011. Predication and edge effects. *Natural Language & Linguistic Theory* 29. Https://doi.org/10.1007/s11049-011-9143-3.

Ko, Heejeong. 2014. *Edges in Syntax: Scrambling and Cyclic Linearization*. Oxford University Press.

Koizumi, Masatoshi. 1995. *Phrase structure in minimalist syntax*. Doctoral dissertation, MIT.

Korsah, Sampson, and Andrew Murphy. 2019. Tonal reflexes of movement in Asante

Twi. *Natural Language & Linguistic Theory* .

Koster, Jan. 2000. *Extraposition as parallel construal*. Ms., University of Groningen.

Kratzer, Angelika. 1996. Severing the External Argument from its Verb. In *Phrase structure and the lexicon*, ed. Johan Rooryck and Laurie Zaring, 109–137. Springer. Https://doi.org/10.1007/978-94-015-8617-7_5.

Kroch, Anthony. 1989. *Amount quantification, referentiality and long wh-movement*. Manuscript, University of Pennsylvania.

Kuno, Susumu, and Ken-ichi Takami. 1997. Remarks on negative islands. *Linguistic Inquiry* 28:553–576.

Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19:335–391.

Lasnik, Howard, and Mamoru Saito. 1999. *Move α*. Cambridge, MA: MIT Press.

Lebeaux, David. 1988. *Language Acquisition and the Form of the Grammar*. Doctoral dissertation, University of Massachusetts, Amherst: GLSA.

Lebeaux, David. 1992. Relative clauses, licensing, and the nature of the derivation. In *Perspectives on phrase structure: Heads and licensing*, ed. Susan Rothstein and Margaret Speas, volume Syntax and Semantics volume 25, 209–239. New York: Academic Press.

Legate, Julie Anne. 2003. Some Interface Properties of the Phase. *Linguistic Inquiry* 34. Https://doi.org/10.1162/ling.2003.34.3.506.

Legate, Julie Anne. 2014. *Voice and v: Lessons from Acehnese*. Linguistic Inquiry Monographs.

Linebarger, Christine. 1980. *The grammar of negative polarity*. Doctoral dissertation, MIT.

Longenbaugh, Nicholas. 2019. *On expletives and the agreement-movement correlation*. Doctoral dissertation, MIT.

Longenbaugh, Nicholas. To appear. The agreement/movement correlation: evidence from participle agreement. *Glossa* .

Longobardi, Giuseppe. 1984. Connectedness, scope and c-command. *Linguistic Inquiry* 16:163–192.

Lyutikova, E.A. 2012. On two types of inversion in the Russian noun phrase (o dvux tipax inversii v russkoj imennoj gruppe). *Russian Language and Linguistic Theory (Russkij yazyk v nauchnom osveshchenii)* 2:65–106.

Lyutikova, E.A. 2014. Russian genitive possessor and formal models of the noun

phrase (Russkij genitivnyj posessor i formal'nyje modeli imennoj gruppy). In *Typology of morphosyntactic parameters. Materials of the international conference. (tmp 2014)*, ed. E.A. Lyutokova, 120–145. Sholokhov Moscow State University for Humanities.

Marantz, Alec. 2007. Phases and words. In *Phases in the theory of grammar*, ed. S. H. Choe et al. Dong In Publisher, Seoul.

Martinović, Martina. 2019. Interleaving syntax and postsyntax: Spellout before syntactic movement. *Syntax* Https://doi.org/10.1111/synt.12169.

Marvin, Tatjana. 2003. *Topics in the Stress and Syntax of Words*. Doctoral dissertation, MIT.

Matushansky, Ora. 2005. Going through a phase. In *Perspectives on phases*, ed. Martha McGinnis and Norvin Richards, 157–81. Cambridge, MA: MIT Working Papers in Linguistics.

McCarthy, John J., and Alan S. Prince. 1993. Generalized alignment. In *Yearbook of Morphology*, ed. Geert Booij and Jaap van Marle, 79–153. Dordrecht: Kluwer.

McCloskey, James. 2000. Quantifier Float and Wh-movement in an Irish English. *Linguistic Inquiry* 51. Https://doi.org/10.1162/002438900554299.

McCloskey, James. 2001. The morphosyntax of wh-extraction in Irish. *Journal of Linguistics* 37:67–100.

McCloskey, James. 2002. Resumption, successive cyclicity, and the locality of operations. In *Derivation and explanation in the minimalist program*, ed. SD Epstein and TD Seeley, 184–226. Oxford, UK: Blackwell.

McCloskey, James. 2006. Questions and Questioning in a Local English. In *Crosslinguistic research in syntax and semantics: Negation, tense, and clausal architecture*, ed. Raffaella Zanuttini, Héctor Campos, Elena Herburger, and Paul Portner, 87–126. Georgetown University Press.

McDaniel, Dana, Cecile McKee, and Judy B. Bernstein. 1998. How children's relatives solve a problem for minimalism. *Language* 74:308–334.

McDaniel, Dana, Helen E. Stickney, Sadie Fowler, and Cecile McKee. 2002. What's thats? *English Studies* 83:53–69.

McGinnis, Martha. 1998. *Locality in A-movement*. Doctoral dissertation, MIT.

Medeiros, David. 2013. Hawaiian VP-remnant movement: A cyclic linearization approach. *Lingua* 127:72–97. Https://doi.org/10.1016/j.lingua.2013.01.008.

Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.

Merchant, Jason. 2002. Swiping in Germanic. In *Studies in Comparative Germanic Syntax: Proceedings from the 15th Workshop on Comparative Germanic Syntax*, ed. C. Jan-Wouter Zwart and Werner Abraham, 289–315. Amsterdam: John Benjamins. Https://doi.org/10.1075/la.53.18mer.

Merchant, Jason. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27:661–738.

Miyagawa, Shigeru. 1989. *Structure and Case Marking in Japanese*. San Diego Calif. Academic Press.

Miyagawa, Shigeru. 2017. Numeral Quantifiers. In *Handbook of japanese syntax*, ed. Masayoshi Shibatani, Shigeru Miyagawa, and Hisashi Noda. Mouton de Gruyter.

Müller, Gereon. 1998. *Incomplete Category Fronting: A Derivational Approach to Remnant Movement in German*. Springer: Dordrecht.

Müller, Gereon. 2002. Two types of remnant movement. In *Dimensions of Movement: From features to remnants*, ed. Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers, and Hans-Martin Gärtner, 209–241. Amsterdam: John Benjamins. Https://doi.org/10.1075/la.48.04bar.

Müller, Gereon. 2007. Towards a Relativized Concept of Cyclic Linearization. In *Interfaces + recursion = language?*, ed. Uli Sauerland and Hans-Martin Gärtner, 61–114. Berlin: Walter de Gruyter. Https://doi.org/10.1515/9783110207552.61.

Müller, Gereon. 2010. CED Effects from the PIC. *Linguistic Inquiry* 41:35–82.

Munn, Alan. 1995. The Possessor that Stayed Close to Home. In *Proceedings of the Western Conference on Linguistics 24*, ed. V. Samiian, 181–195.

Murphy, Andrew. 2020a. *Left-Branch Extraction and Barss' Generalization: Against a remnant movement approach*. Talk presented at West Coast Conference on Formal Linguistics 38.

Murphy, Andrew. 2020b. *Left-branch extraction and remnant movement*. Unpublished manuscript, University of Chicago. Http://www.andrewmurphy.de/articles/lberem.pdf.

Newell, Heather. 2008. *Aspects in the morphology and phonology of phases*. Doctoral dissertation, McGill University.

Nissenbaum, John. 1998. Movement and Derived Predicates: Evidence from Parasitic Gaps. In *The Interpretive Tract*, 247–295. Cambridge, MA: MIT Working Papers in Linguistics.

Nissenbaum, John. 2000. *Investigations of covert phrase movement*. Doctoral dissertation, MIT.

Nissenbaum, John, and Bernhard Schwarz. 2011. Parasitic degree phrases. *Natural Langauage Semantics* 19:1–38. Https://doi.org/10.1007/s11050-010-9061-7.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Cambridge, MA: MIT Press Linguistic Inquiry Monographs.

Obenauer, Hans-Georg. 1984. On the identification of empty categories. *Linguistic Review* 4:153–202.

Ostrove, Jason. 2018. Stretching, spanning, and linear adjacency in Vocabulary Insertion. *Natural Language & Linguistic Theory* 36:1263–1289.

Ott, Dennis. 2011. *Local instability: The syntax of split topics*. Doctoral dissertation, Harvard University.

Overfelt, Jason. 2015. Extraposition of NPIs from NP. *Lingua* 164:25–44. Http://dx.doi.org/10.1016/j.lingua.2015.06.006.

Overfelt, Jason. 2016. Rightward DP movement licenses parasitic gaps: A reply to Postal 1994. *Linguistic Inquiry* 47:127–146. Https://doi.org/10.1162/LING_a_00206.

Pankau, Andreas. 2019. *Left branch extraction in Lower Sorbian: A problem for the DP/NP analysis*. Poster presented at Formal Approaches to Slavic Linguistics 28.

Partee, Barbara, and Mats Rooth. 1983. Generalized conjunction and type ambiguity. In *Meaning, use, and interpretation of language*, ed. Rainer Baüerle, Christoph Schwarze, and Arnim von Stechow, 361–383. Berlin: de Gruyter.

Pereltsvaig, Asya. 2006. Small Nominals. *Natural Language & Linguistic Theory* 24:433–500.

Pereltsvaig, Asya. 2007. The Universality of DP: A View From Russian. *Studica Linguistica* 61:59–94.

Pereltsvaig, Asya. 2008. Split phrases in colloquial Russian. *Studica Linguistica* 62:5–38.

Pesetsky, David. 1982. *Paths and categories*. Doctoral dissertation, MIT.

Pesetsky, David. 1989. *Language-particular processes and the earliness principle*. Unpublished manuscript, MIT.

Pesetsky, David. 2000. *Phrasal Movement and Its Kin*. Cambridge: MIT press.

Pesetsky, David. 2006. *Cyclic Linearization of Syntactic Structure: Evidence from Scandinavian, Korean and perhaps Russian*. Talk given at Moscow State University.

Pesetsky, David. 2013. *Russian case morphology and the syntactic categories*. Cambridge: MIT press.

Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 335–426. Cambridge, MA: MIT Press.

Podobryaev, Sasha. 2009. Postposition Stranding and related phenomena in Russian. In *Studies in formal Slavic Phonology, Morphology, Syntax, Semantics and Information Structure*, ed. G Zybatow, U Junghanns, D Lenertová, and P Biskup. Frankfurt am Main: Peter Lang.

Polinsky, Maria, and Eric Potsdam. 2001. Long-Distance Agreement and Topic in Tsez. *Natural Language & Linguistic Theory* 19:583–646.

Pollard, Carl, and Ivan Sag. 1992. Anaphors in English and the scope of binding theory. *Linguistic Inquiry* 23:261–303.

Postal, Paul. 1972. On Some Rules That Are Not Successive Cyclic. *Linguistic Inquiry* 3.

Postal, Paul. 1974. *On Raising*. MIT Press.

Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford University Press, New York.

Preminger, Omer. 2019. What the PCC tells us about abstract agreement, head movement, and locality. *Glossa: a journal of general linguistics* 4:1–42. Https://doi.org/10.5334/gjgl.315.

Progovac, Ljiljana. 1998. Determiner phrase in a language without determiners. *Journal of Linguistics* 34:165–179.

Rackowski, Andrea. 2002. *The structure of Tagalog: Specificity, voice, and the distribution of arguments*. Doctoral dissertation, MIT.

Rackowski, Andrea, and Norvin Richards. 2005. Phase Edge and Extraction: A Tagalog Case Study. *Linguistic Inquiry* 36. Https://doi.org/10.1162/002438905774464368.

Radford, A. 1997. *Syntax: A Minimalist Introduction*. Cambridge University Press, Cambridge: UK.

Reeve, Matthew, and Glyn Hicks. 2017. Adjunct extraposition: Base generation or movement? *Syntax* 20:215–248. Https://doi.org/10.1111/synt.12136.

Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. Chicago: University of Chicago Press.

Reinhart, Tanya. 2006. *Interface Strategies*. The MIT Press, Cambridge, MA.

Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657–720.

Richards, Norvin. 1997. *What moves where in which language?*. Doctoral dissertation, MIT.

Richards, Norvin. 1998. The principle of minimal compliance. *Linguistic Inquiry* 29:599–629.

Richards, Norvin. 1999. Featural cyclicity and ordering of multiple specifiers. In *Working minimalism*, ed. Samuel D. Epstein and Norbert Hornstein, 127–158. Cambridge, MA: MIT Press.

Richards, Norvin. 2004. Against bans on lowering. *Linguistic Inquiry* 35. Https://doi.org/10.1162/0024389041402643.

Richards, Norvin. 2010. *Uttering trees*. Cambridge, MA: MIT Press.

Richards, Norvin. 2016. *Contiguity Theory*. MIT Press.

van Riemsdijk, Henk. 1978. *A case study in syntactic markedness: The Binding Nature of Prepositional Phrases*. Lisse: The Peter de Ridder Press.

van Riemsdijk, Henk. 1997. Push chains and drag chains: Complex predicate split in Dutch. In *Scrambling*, ed. Shigeo Tonoike, 7–33. Tokyo: Kurosio.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar: A handbook of generative syntax*, 281–337. Dordrecht: Kluwer.

Rizzi, Luigi. 2006. On the form of chains: Criterial positions and ECP effects. In *Wh-movement*, ed. Lisa Lai-Shen and Norbert Corver, 97–33. Cambridgem MA: MIT Press.

Ross, John. 1967. *Constraints on variables in syntax*. Doctoral dissertation, MIT.

Ross, John. 1969. Guess who? In *Papers from the 5th regional meeting of the Chicago Linguistics Society*, ed. Robert Binnick, Alice Davison, Georgia Green, and Jerry Morgan, 252–286. Chicago Linguistic Society: Chicago, Ill.

Rudin, Catherine. 1988. On Multiple Questions and Multiple WH Fronting. *Natural Language & Linguistic Theory* 6.

Sabbagh, Joseph. 2007. Ordering and linearizing rightward movement. *Natural Language & Linguistic Theory* 25. Https://doi.org/10.1007/s11049-006-9011-8.

Safir, Ken. 1986. Relative clauses in a theory of binding and levels. *Linguistic Inquiry* 17:661–689.

Safir, Ken. 2018. The A/A′ Distinction as an Epiphenomenon. *Linguistic Inquiry*

50:285–336. Https://doi.org/10.1162/ling_a_00305.

Saito, Mamoru. 1985. *Some asymmetries in Japanese and their theoretical implications*. Doctoral dissertation, MIT.

Sauerland, Uli. 1998. *On the making and meaning of chains*. Doctoral dissertation, MIT.

Sauerland, Uli. 1999. Erasability and interpretation. *Syntax* 2:161–188.

Sauerland, Uli. 2003. Intermediate adjunction with A-movement. *Linguistic Inquiry* 34.

Seppänen, Aimo, and Göran Kjellmer. 1995. The dog that's leg was run over: On the genitive of the relative pronoun. *English Studies* 4:389–400.

Sheehan, Michelle. 2010. Extraposition and antisymmetry. In *Linguistic variation yearbook 10*, ed. Jeroen van Craenenbroeck, 10:201–251. Amsterdam: John Benjamins.

Sportiche, Dominique. 1988. A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure. *Linguistic Inquiry* 19:425–499.

Stanton, Juliet. 2016. Wholesale Late Merger in A′-movement: Evidence from Preposition Stranding. *Linguistic Inquiry* 47:86–126. Https://doi.org/10.1162/ling_a_00205.

Starke, Michal. 2001. *Move dissolves into Merge: A theory of Locality*. Doctoral dissertation, University of Geneva.

Stepanov, Arthur. 2001. Late adjunction and minimalist phrase structure. *Syntax* 4:94–125.

Stjepanović, Sandra. 2010. Left Branch Extraction in Multiple Wh-Questions: A Surprise for Question Interpretation. In *Proceedings of Formal Approaches to Slavic Linguistics 18*, ed. W. Browne, A. Cooper, A. Fisher, E. Kesici, N. Predolac, and D. Zec, 502–517. Michigan Slavic Publications: Ann Arbor MI.

Stowell, Timothy. 1981. *Origins of Phrase Structure*. Doctoral dissertation, MIT.

Stroik, Thomas. 2009. *Locality in Minimalist Syntax*. Cambridge, MA: MIT Press. Linguistic Inquiry Monographs. Vol. 51.

Syed, Saurov, and Andrew Simpson. 2017. On the DP/NP status of nominal projections in Bangla: Consequences for the theory of phases. *Glossa: a journal of general linguistics* 2:1–24. Https://doi.org/10.5334/gjgl.186.

Szabolcsi, Anna. 1984. The possessor that ran away from home. *The Linguistic Review* 3.

Szabolcsi, Anna, and Terje Lohndal. 2017. Strong vs. Weak Islands. In *The Wiley Blackwell Companion to Syntax, Second Edition*, ed. Martin Everaert and Henk van Riemsdijk. John Wiley and Sons, Inc.

Tada, Hiroaki. 1993. *A/A-bar partition in derivation*. Doctoral dissertation, MIT.

Takahashi, Shoichi. 2004. Pseudogapping and Cyclic Linearization. In *Proceedings of North East Linguistic Society 34*, ed. Keir Moulton and Matthew Wolf. UMass Amherst: GLSA Publications.

Takahashi, Shoichi. 2018. *Stranding in Disguise and Locality*. Handout from Given!: A Workshop to Celebrate the Life and Work of Michael Rochemont, University of British Columbia, Canada, December 5-6.

Takahashi, Shoichi, and Akane Ohtaka. 2017. A Lacuna in Adjunct Extraposition. In *Proceedings of North East Linguistic Society 47*, ed. Andrew Lamont and Katerina Tetzloff, 155–164. Amherst, Mass.: GLSA Publications.

Takano, Yuji. 1993. *Minimalism and Proper Binding*. Ms., University of California, Irvine.

Takita, Kensuke. 2010. *Cyclic Linearization and Constraints on Movement and Ellipsis*. Doctoral dissertation, Nanzan University.

Taraldsen, Knut. 1981. The theoretical interpretation of a class of marked extractions. In *Theory of Markedness in Generative Grammar: Proceedings of the 1979 GLOW conference*, ed. L. Brandi A. Belletti and L. Rizzi, 476–516. Scuola Normale Superiore di Pisa.

Taraldsen, Knut. 1995. On agreement and nominative objects in Icelandic. In *Studies in comparative germanic syntax*, ed. Hubert Haider, Susan Olsen, and Sten Vikner, 307–327. Dordrecht: Kluwer.

Torrence, Harold. 2012. The morpho-syntax of silent wh-expressions in Wolof. *Natural Language & Linguistic Theory* 30:1147–1184.

Torrence, Harold. 2013. *The Clause Structure of Wolof: Insights into the left periphery*. Amsterdam/Philadelphia: John Benjamins.

Torrence, Harold. 2018. *Left Edge Agreeing Elements inWolof Wh-Questions*. MS, UCLA.

Travis, Lisa. 1984. *Parameters and Effects of Word Order Variation*. Doctoral dissertation, MIT.

Ura, Hiroyuki. 1996. *Multiple feature-checking: A theory of grammatical function splitting*. Doctoral dissertation, MIT.

Urban, Emily. 1999. *Exactly stranding*. Ms., University of California, Santa Cruz.

Uriagereka, Juan. 1988. *On Government*. Doctoral dissertation, University of Connecticuit.

van Urk, Coppe. 2015. *A uniform syntax for phrasal movement: A case study of Dinka Bor*. Doctoral dissertation, MIT.

van Urk, Coppe, and Norvin Richards. 2015. Two Components of Long-Distance Extraction: Successive Cyclicity in Dinka. *Linguistic Inquiry* 46:113–155.

de Vries, Mark. 2002. *The Syntax of Relativization*. Doctoral dissertation, Universiteit van Amsterdam.

de Vries, Mark. 2007. Invisible constituents? Parentheticals as b-merged adverbial phrases. In *Parentheticals*, ed. Nicole Dehé and Yordanka Kavalova, 203–234. Amsterdam: John Benjamins. Https://doi.org/10.1075/la.106.

Watanabe, Akira. 1986. *On some root/non-root asymmetries*. Unpublished manuscript, University of Tokyo.

Wiland, Bartosz. 2010. Overt Evidence from Left-Branch Extraction in Polish for Punctuated Paths. *Linguistic Inquiry* 41. Https://doi.org/10.1162/ling.2010.41.2.335.

Williams, Edwin. 1974. *Rule ordering in syntax*. Doctoral dissertation, MIT.

Williams, Edwin. 1999. *Economy as shape conservation*. Unpublished manuscript, Princeton University.

Williams, Edwin. 2003. *Representation Theory*. Cambridge, MA: MIT Press.

Williams, Edwin. 2005. A Note on "Cyclic Linearization". *Theoretical Linguistics* 31:229–234. Https://doi.org/10.1515/thli.2005.31.1-2.229.

Wurmbrand, Susi. 2013. QR and Selection: Covert Evidence for Phasehood. In *Proceedings of North East Linguistic Society 42*, ed. Stefan Keine and Shayne Sloggett. GLSA Publications: UMass Amherst.

Wurmbrand, Susi. 2017. Stripping and topless complements. *Linguistic Inquiry* 48:341–366.

Zanon, Ksenia. 2015. Russian anaphoric possessive in context. In *Proceedings of Formal Approaches to Slavic Linguistics 23*, ed. M. Szajbel-Keck, R. Burns, and D. Kavitskaya, 362–381. Ann Arbor: Michigan Slavic Publications.

Zwart, Jan-Wouter. 1997. Transitive expletive constructions and the evidence supporting the multiple specifier hypothesis. In *German: Syntactic problems*, ed. W Abraham and van Gelderen, 105–134. Niemeyer: Tubingen.

Zwicky, Arnold. 1971. In a manner of speaking. *Linguistic Inquiry* 2:223–233.

Zwicky, Arnold. 1987. Suppressing the Zs. *Journal of Linguistics* 23:133–148.

Zyman, Erik. 2019. *Antilocality at the Phase Edge*. Manuscript, University of Chicago.

Zyman, Erik. To appear. *Phase Constrained Obligatory Late Adjunction*. Syntax.