Tough Constructions in the Context of English Infinitives

by

Ruth Brillman

B.A., New York University (2012)

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

Doctor of Philosophy in Linguistics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2017

Author ..... Signature redacted .....
Department of Linguistics and Philosophy
August 25 2017

Certified by ..... Signature redacted .....
David Pesetsky
Ferrari P. Ward Professor of Linguistics
Thesis Supervisor

Accepted by ..... Signature redacted .....
David Pesetsky
Chair, Department of Linguistics and Philosophy

Tough Constructions in the Context of English Infinitives

by

Ruth Brillman

Submitted to the Department of Linguistics and Philosophy
on August 25 2017, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Linguistics

Abstract

The dissertation was inspired by the question of why subjects cannot undergo *tough* movement (1).

(1) a. Jonathan Franzen is easy for Anneke to criticize _.
b. *Anneke is easy _ to criticize Jonathan Franzen.

To answer this question, this dissertation proposes that a spec-to-spec anti-locality constraint (in the spirit of Erlewine 2016 and Brillman & Hirsh *to appear*) limits subject *tough* movement because the subject *tough* movement chain is "too short." Brillman & Hirsh's spec-to-spec anti-locality constraint is given in (2). Spec-to-spec anti-locality bans subject *tough* movement because subject *tough* movement would need to involve Ā movement from the embedded spec-TP to the immediately adjacent spec-CP. This movement chain is banned by spec-to-spec anti-locality.

(2) **Spec-to-Spec Anti-Locality**
Ā-movement of a phrase from the specifier of XP must **cross** a specifier projected by a maximal projection other than XP. Movement from position α to β **crosses** γ if and only if γ dominates α but does not dominate β

A spec-to-spec anti-locality analysis of the ban on subject *tough* movement also provides an explanation for why gapped degree phrases, a syntactic structure that shows many parallels to *tough* constructions (Lasnik & Fiengo 1974) can license subject movement (3). This dissertation will show that, compared to *tough* constructions, gapped degree phrases have a structurally "larger" embedded clause, with a DegP layer dominating the CP layer. This DegP layer contains a syntactically and semantically active (but optionally silent) evaluator argument in it's specifier; this argument situates the threshold of the degree predicate, allowing the degree word to be compared to a pre-determined standard. Subject movement within a gapped degree phrase would involve Ā-movement from spec-TP to spec-DegP, across CP and the evaluator argument in spec-DegP. This movement chain is not banned by spec-to-spec anti-locality.

(3) a. Jonathan Franzen is banal [OP EVAL enough for Anneke to criticize _]
b. Anneke is intelligent [OP EVAL enough _ to criticize Jonathan Franzen]

This dissertation will argue that spec-to-spec anti-locality can explain much more than the contrast between how *tough* constructions and gapped degree phrases treat subject extraction. Particularly, it will propose that anti-locality can explain a wide range of subject/non-subject asymmetries, both within English and cross-linguistically. These include complementizer trace effects, *do*-support asymmetries in English subject *wh*-questions, as well as specific subject/non-subject Ā alternations in Imbabura Quechua, Hebrew and Berber.

Finally, this dissertation is also interested in the related question of whether or not *tough* construction subject/non-subject asymmetries are represented identically across both the adult and child grammars. To that end, this dissertation presents the results of a novel acquisition experiment which shows that children do not represent *tough* construction extraction asymmetries the same way that adults do. Specifically, these results show that—while adults find subject *tough* constructions ungrammatical and object *tough* construction grammatical—children find both subject and object *tough* constructions ungrammatical. Interestingly, this experiment also shows that children do not have an adult-like representation of argument extraction asymmetries in raising constructions. While the adult grammar only allows for subjects to raise, the child grammar allows both subjects and object to raise. This dissertation will discuss what these results mean, both in terms of how these results relate to previous work on the acquisition of *tough* and raising and in terms of what these results can tell us about the syntax of *tough* movement.

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

Acknowledgments

Allow me a minute (well, two and half pages) of self-indulgence: Grad school was hard. At least, grad school was hard *for me*. For the past three years, about once every six months, I've asked myself (with varying degrees of sincerity) whether or not I should just take a masters degree and quit this whole thing. Don't get me wrong. Now that I'm out the other side, it's obvious that leaving MIT without a PhD would have been a huge mistake. But if I was at another department—and, especially, if I had a less committed advisor and if I had not been able to form a small, chosen family among my peers—it's a mistake that I would have made. I tend to shy away from declarations of feeling, so bear with me. I finally have a good excuse, and there's a lot to say:

I never would have gotten here—to grad school, to MIT—without Stephanie Harves. If not for Stephanie, I never would have majored in linguistics, much less pursued it after college. Stephanie is the beating heart of undergraduate linguistics at NYU, and I was incredibly fortunate to be able to work with her so early in my career. The ultimate complement in our small, weird, wonderful world, Stephanie is often referred to as "the David Pesetsky of undergrads." She's the best teacher (and the best mentor) that I've ever had, and I've spent most of my adult life trying to live up to her impression of me.

My graduate advisor, of course, was the *actual* David Pesetsky—I might be the first person who's ever been lucky enough to get them both. David doesn't just live up the rumors; he exceeds them—a feat that I didn't think was possible. I didn't always make myself an easy advisee (would *you* want to advise someone who kept wanting to just take a masters degree and quit this whole thing?), but David pushed me to do my best work, even when I was content to let my work be mediocre. In a very real sense, my PhD would not have happened without him. At least two decades of students have written about the intellectual benefits of working with David, so I'm not sure there's anything new for me to say, but here goes: David has taught me how to question my assumptions; he never let me skate (and I *tried*); he forced my writing to be straightforward and clear; he showed me that it's ok (and often helpful) to admit when I didn't know the answer[1]. Because of David, I'm a stronger researcher, a more transparent writer, and a sharper critical thinker. I'm profoundly grateful for all of these things. I might have gotten here because of Stephanie, but I'll owe David for wherever I go next.

I'm also extremely for the input of my thesis committee members, Norvin Richards and Danny Fox. It was a privilege to work with both them over the course of the dissertation, as well as on previous projects. Danny and Norvin are both exceptional linguists, and they make a wonderful and complementary team.

[1]He has failed, despite his best efforts, to curb my tendency to confuse *its* and *it's*.

I am deeply appreciative of the time and care they put into the (substantial) feedback that they offered me over the course of this work. This dissertation is much stronger for their comments and input. What's more, I can't recall ever having a meeting or conversation with either of them that wasn't helpful, enjoyable, and productive.

Aron Hirsh was my co-author on our original paper on anti-locality—which the third chapter of this dissertation is largely based on—and I am incredibly grateful for his contributions. Aron's thoughtfulness and carefulness are evident throughout the anti-locality chapter of this dissertation, both within and outside of the sections adapted from our joint work. The development of our theory of anti-locality was very much a joint effort, and any successes of the anti-locality chapter are owed as much to Aron as they are to me (any errors, unfortunately, are my responsibility).

Mingshi Yang was an invaluable contributor to my language acquisition work. She helped develop the scenarios and stimuli used in the experiment (including drawing all images by hand), and ran the experiment in my absence when I spent a summer working in industry. She put tremendous effort into making this project work, and I honestly believe that I would not have been able to finish it without her.

I've also enormously benefitted from the MIT Linguistics faculty. Martin Hackl has met with me about almost every project I've ever worked on, from grant applications in my first year to experimental design development for my dissertation. Ken Wexler also helped me design my first language acquisition experiment, and I'm extremely happy to have gotten to work with such a giant of that field. Adam Albright has very patiently explained the same statistics concepts to me at least seven times, and somehow never made me feel stupid in the process; what little I know about statistical analysis I owe to him. If any of the semantics in this dissertation makes sense, you can thank Kai von Fintel and Irene Heim (and probably Martin, again, too).

The MIT Linguistics community, of course, extends beyond the faculty. Without my officemates Gretchen Kern, Mia Nussbaum, Carrie Spadine, and TC Chen, I would have spent even more time languishing at home (and missed out on so much free food). I'm grateful to my ling-12 cohort: Sam Zukoff, Aron Hirsch (let me thank him again, and again, and *again*), Chris O'Brien (particularly for the halcyon era of pi-lambda-phi), Ben Storme, Ishani Guta, Lilla Magyar, Juliet Stanton and, particularly, Anthony Brohan, whose move to greener pastures I never quite recovered from. I'm so happy that grad school gave me the chance to meet Michelle Yuan (and #Copernikisses), Kenyon Brannan, Michelle Fullwood, Athulya Aravind, Sam Steddy, Verena Hehl, and Snejana Iovtcheva, among many, many others.

I was especially lucky to have the support of a non-academic community over the

past five years. Sarah Schoengold, Kate Lemay, Nicole Torres and Tessa Hite all helped keep me sane, and made sure I never missed the forest for the trees. I'm grateful to my many colleagues at Akamai, Amazon, and Spotify, who helped me understand exactly how important this kind of work can be, even when it seemed too esoteric and obscure. Jenn Thom, Sarah Mennicken, Henriette Kramer and the rest of my team at Spotify made it possible for me balance the mammoth tasks of working and writing (they've even managed to make that balancing act fun). I'm also very grateful to Anneke Schwob—and her profound intolerance for white male exceptionalism—without whom the examples in this dissertation would be much more milquetoast (much like the writing of Jonathan Franzen).

Finally, I'm so thankful for my family and their unending love and support. In particular, I'd like to thank my mother, Judith Brillman and my step-father, Bart Wohl. My mother has never let me settle for anything; where some mothers urge their daughters to marry a doctor, mine pushed me to become one[2]. I'd also like to thank my husband, Chris Davis. Since I've known him, Chris has served as my north star example of what a brilliant, thorough and fastidious scientist should look like. I couldn't imagine doing this—any of this—without his patience, sense of humor, and endless compassion. Or, I could imagine it. But that version's black and white, and I'd so much rather live in color.

[2]Mission accomplished on both fronts, Mom, so take *that*.

To my family

Contents

Chapter 1

Background

1.1 Overview

This dissertation is about subject/non-subject Ā extraction asymmetries from infinitival and embedded clauses. Particularly, it is interested in when subjects can and cannot be extracted from an infinitive and why these extraction asymmetries occur in the grammar. Some subject/non-subject extraction restrictions and asymmetries are understood better than others. For example, subject/non-subject A-movement asymmetries are well known; (1) shows a classic asymmetry where embedded subjects are allowed to raise to a matrix subject position while embedded objects are not. This asymmetry is due to relativized minimality: the object cannot extract to the matrix subject position without first A-moving across the embedded subject. However, other extraction asymmetries—particularly Ā extraction asymmetries—are less well understood. For example, while there is a substantial literature surrounding *tough* constructions, it is not understood why subjects cannot undergo *tough* movement, while objects can (2).

(1) Raising

a. Anneke is unlikely _ to read about Jonathan Franzen.
(subject raising ✓)

b. *Jonathan Franzen is unlikely for Anneke to read about _.
(object raising ✗)

(2) *Tough* movement

a. *Anneke is tough _ to think about Jonathan Franzen[1].
(subject *tough* ✗)

b. Jonathan Franzen is tough for Anneke to think about _.
(object *tough* ✓)

Perhaps the most well known English subject/non-subject Ā extraction asymmetry is the *that*-trace effect, where subject extraction from an embedded clause introduced with an overt complementizer is disallowed, while subject extraction from an embedded clause without an overt *that* is allowed (3).

(3) *That*-trace effects

a. Who does Bill think that _ saw John? (subject extraction ✗)
b. Who does Bill think that John saw _? (object extraction ✓)
c. Who does Bill think _ saw John? (subject extraction ✓)
d. Who does Bill think John saw _? (object extraction ✓)

Most broadly, this dissertation is interested in why these asymmetries are part of the grammar. To answer this question, however, this dissertation will primarily focus on *tough* constructions. By closely investigating a specific embedded clause construction that displays a subject/non-subject extraction asymmetry, this dissertation seeks to understand the distribution of subject/non-subject Ā extraction asymmetries across embedded clause constructions.

1.1.1 Core questions and goals

The main goal of this dissertation is to provide a theory of subject/non-subject extraction asymmetries by looking at the relationship between *tough* constructions and other infinitival and embedded clause constructions. This relationship is investigated across three main research chapters.

The first of these (the second chapter of this dissertation) will look at the relationship between *tough* constructions and gapped degree phrases[2]. Gapped

[1]Subject *tough* constructions are marginally grammatical on a reading where the matrix subject receives a Θ-role directly from the *tough* predicate and the embedded clause is interpreted as an adjunct that does not operate within the modal scope of the *tough* predicate. This reading is approximately "Anneke is a tough person because she is willing to think about Jonathan Franzen" for (2a). This is not a *tough* construction interpretation, and will be ignored for the purposes of this dissertation. I leave the analysis of these marginally grammatical constructions and any question of their relationship to *tough* constructions open for future research.

[2]Gapped degree phrases are also called *parasitic degree phrases* by Nissenbaum and Schwarz. However, this name implies a particular analysis—namely, that the gap in these constructions is

degree phrases have previously been observed to be relatively mysterious: since Lasnik & Fiengo (1978) gapped degree phrases have been known to be, in some sense, syntactically like *tough* constructions and they have also been argued to be semantically parallel to parasitic gap constructions (Nissenbaum & Scwharz 2011). Chapter 2 provides a more detailed inquiry into the similarities and differences between *tough* constructions and gapped degree phrases. This chapter provides further evidence for the claim that the two constructions are, in fact, very much alike. The core of this "likeness" is an argument that gapped degree phrases and *tough* construction contain the same type of movement chain: an improper movement chain that involves an Ā-movement step, followed by an A-movement step. However, despite this similarity, gapped degree phrases and *tough* constructions show one key difference: while subject *tough* movement is banned, gapped degree phrases allow both subject and object gaps, as in (4).

(4) Gapped degree phrases

a. Anneke is *too smart* _ to read about Jonathan Franzen.
(subject gap ✓)

b. Jonathan Franzen is *too inconsequential* for Anneke to read about _.
(object gap ✓)

Chapter 2 also argues that another key difference between *tough* constructions and gapped degree phrases is that, while *tough* constructions and gapped degree phrases both involve an adjectival predicate that embeds a subordinate clause, the two constructions contain embedded clauses of different sizes. Specifically, *tough* constructions contain an embedded CP clause, while gapped degree phrases contain an embedded Deg(ree)P clause, with a DegP layer immediately above a CP layer. In this way, gapped degree phrases contain an embedded clause that is one layer "larger" than the embedded clause in a corresponding *tough* construction.

Chapter 3, looks at why gapped degree phrases allow for both subject and object gaps while *tough* constructions only allow for object gaps. This chapter argues that a spec(ifier)-to-spec(ifer) anti-locality constraint, like the one proposed by Erlewine (2016) or Brillman & Hirsch (*to appear*) underlies this difference. This

syntactically and semantically like the gaps in parasitic gap constructions. Although this work adopts and relies on the analysis of gapped degree phrases put forth by Nissenbaum & Schwarz, this dissertation will use a more theory-neutral name for the construction. Outside of Nissenbaum & Schwarz, this work, and my own previous work (Brillman 2015, Brillman & Hirsch *to appear*), I do not believe that gapped degree phrases have a standardized name across the literature, and are often referred to very descriptively (e.g., "gapped infinitival clauses with a degree word like "too").

chapter discusses spec-to-spec anti-locality more generally: a syntactic constraint that limits Ā-movement operations between the specifiers of adjacent nodes, including phase-driven successive-cyclic subject Ā-movement from spec-TP to an immediately adjacent spec-CP. Because *tough* constructions have an embedded clause that only extends to a CP layer, successive-cyclic subject movement within a *tough* construction would need to progress from the embedded spec-TP to the embedded spec-CP, violating anti-locality. However, because gapped degree phrases involve structurally "larger" embedded clauses (with a DegP layer directly above the CP layer), Chapter 3 argues that subject movement within a gapped degree phrase can proceed from the embedded spec-TP to spec-DegP (crossing the embedded CP layer), respecting anti-locality.

This analysis involves a somewhat non-standard view of phasehood—as movement in gapped degree phrases is able to cross an embedded CP layer without causing a Phase Impenetrability Condition (PIC, see Chomsky 2000) violation—and a non-traditional definition of clausehood—as the DegP layer, not the CP layer, is argued to be the embedded clause boundary in gapped degree phrases. However, these are not ad hoc proposals. Particularly, these assumptions follow directly form an independently proposed definition of phasehood put forth by Wurmbrand (2017) to explain a set of ellipsis facts. Wumbrand's theory of phasehood, and the data the led to her proposal, are discussed in Section 3 of this chapter.

Following the introduction and definition of spec-to-spec anti-locality, Chapter 3 lays out examples of how spec-to-spec anti-locality is active in the grammar and shows that spec-to-spec anti-locality can account for some of the differences between gapped degree phrases and *tough* constructions. Additionally, Chapter 3 proposes that spec-to-spec anti-locality has the power to unify a wide set of subject/non-subject asymmetries—both within English and cross-linguistically—that were previously believed to be unrelated (though see Rizzi and Shlonsky 2007 for an alternative analysis) and discusses how anti-locality and its interactions with other components of the grammar can capture these asymmetries.

Finally, Chapter 4 discusses the acquisition of subject and object extraction asymmetries in raising and *tough* constructions. The acquisition literature has long shown that children are delayed in their acquisition of *tough* constructions, and that they do not demonstrate a full mastery of *tough* constructions until around age eight (see **?** 1969, Cromer 1970, among many others). The raising

literature is mixed. Some studies show that children demonstrate an adult-like comprehension of raising constructions by age four (Becker 2006) while other studies argue that children do not display a mastery of raising until age eight, similar to with *tough* constructions (Hirsch et al. 2008). However, both constructions are controversial in the acquisition literature. The *tough* debate surrounds the question of *why* children are delayed in their acquisition of *tough* constructions while the raising debate surrounds *whether* children are delayed in their acquisition of raising.

I argue that both debates stem from the fact that much of the data produced from traditional acquisition experiments is inherently ambiguous. Traditional acquisition studies are focused on performance data, which can only show whether or not child and adults respond to a particular construction the same way. When child and adult performance differ, traditional acquisition studies cannot provide greater insight into how a child represents a particular construction, or even whether or not the child finds the construction grammatical. Chapter 4 presents an experiment that solicits grammaticality judgments directly from children. Child grammaticality judgment are a richer data source than pure performance data and so can present a clearer picture of how raising and *tough* movement are situated within the child grammar.

The results of this chapter's experiment provide evidence that the child grammar does not contain the same subject/non-subject asymmetries as the adult grammar. The results of this study show that children find both subject and object raising constructions grammatical and also believe that both subject and object *tough* movement is ungrammatical. This chapter questions why the adult and child grammars might differ in which subject/non-subject asymmetries they allow and, to explain the raising facts, argues that the child grammar does not involve the same relativized minimality restrictions as the adult grammar. This chapter also proposes that the *tough* movement data is best explained by an improper-movement analysis of *tough* movement in the adult grammar, as well as a theory of the child grammar where children have not learned that *tough* movement is an exception to a general ban on improper movement.

1.1.2 Necessary background and theoretical contributions

The syntax of *tough* movement

The syntax of *tough* constructions is a hotly debated issue (? 1964, 1977, 1981, Rosenbaum 1967, Ross 1967, Postal 1971, Lasnik and Fiengo 1974, Hornstein 2001, Hicks 2009, and Hartman 2012, among others). The *tough* construction debate essentially boils down to the question of whether *tough* constructions are null operator (e.g., Chomsky 1977, 1981) or overt DP movement (e.g., Hartman 2012) constructions, or a hybrid account (Hicks 2009). While this dissertation does not directly probe the syntax of *tough* constructions, the representation of *tough* constructions is still very relevant to this work. To this end, this chapter provides an overview of the competing syntactic theories of *tough* movement in Section 2.

To a certain extent, the contributions of this dissertation are prediction neutral across competing theories of *tough* movement. Take, for example, the discussion of why subject *tough* constructions are ungrammatical while subject gapped degree phrases are grammatical (Chapter 3). This analysis depends on the fact that *tough* constructions contain an embedded CP phase, while gapped degree phrases have an embedded DegP phase. Whether *tough* constructions involve moving a null operator or overt DP within their embedded clauses (to the edge of a phase boundary) does not influence the analysis.

However, this dissertation also contributes new information to the *tough* syntax debate, particularly the results of the acquisition experiment in Chapter 4. While these results are far from definitive, they are also not neutral. The results of the acquisition experiment in Chapter 4 are more compatible with the overt DP theory of *tough* movement than they are with the null operator theory of *tough* movement. In this respect, this dissertation can also be taken as evidence—albeit tentative and preliminary evidence—in favor of an overt DP theory of *tough* syntax. For this reason, when *tough* constructions are represented in this dissertation, unless otherwise noted (e.g., to illustrate the parallels between *tough* movement and a null operator construction), they will be represented with a syntax that involves the movement of an overt DP.

Definitions of phases and clauses

This dissertation can also be understood as an argument for a theory of syntax where phasehood is dynamically determined and where multiple embedded clause sizes and structures (e.g., finite bare TPs, non-finite CPs) are possible.

Particularly with respect to the comparison of *tough* constructions and gapped degree phrases, this dissertation requires a theory of syntax where there is no one-size-fits-all analysis of embedded clause structure. This kind of syntactic theory would allow for *tough* constructions to have a CP as their embedded clause, while allowing gapped degree phrases to have a DegP as their embedded clause. In the analysis of *that*-trace effects and other English subject/non-subject asymmetries in Chapter 3, additionally, we will show that there is reason to believe that some finite clauses which lack an overt complementizer are best analyzed as bare TPs.

This theory of clause size is due to Wurmbrand (2017), who argues that phasehood is not static but rather dynamically determined based on context. Rather than having phasehood be a property of certain syntactic heads (e.g., C, agentive v), Wurmbrand argues that phasehood is really a property of structure. Particularly, she argues (Wurmbrand 2017:346) that the "highest projection of a cyclic domain" constitutes a phase. Here, the notion of phasehood is also linked to the notion of clause size. Wurmbrand's (2017) theory provides the grammar with the flexibility to allow for a range of embedded clause types. Wurmbrand's proposal and supporting arguments are presented in Section 3 of this chapter.

1.2 Two (plus) theories of *tough* movement

This section very briefly outlines the debate surrounding the syntax of *tough* movement, focusing on the two main competing hypotheses: the improper movement analysis of *tough* constructions and the null operator analysis of *tough* constructions. Particularly, this section is devoted to giving a brief overview of the pros and cons of each approach, without taking a definitive stance on which analysis best captures the distributions and properties of *tough* constructions.

1.2.1 The improper movement analysis

The improper movement analysis (sometimes called the overt DP movement or the long movement analysis) proposes that *tough* constructions have the structure outlined in (5). There, the embedded object DP first undergoes successive-cyclic Ā-movement to the edge of the embedded clause (represented here as a CP) before A-moving into the matrix subject position.

(5) Improper movement *tough* movement structure
[$_{DP}$ Jonathan Franzen]$_i$ is tough [$_{CP}$ <DP$_i$> for Anneke to [$_{vP}$ <DP$_i$> think about <DP$_i$>]]

The fact that an Ā-movement chain exists within the embedded clauses of *tough* constructions is non-controversial, and the analysis of movement within the embedded clause is essentially the same in both an improper movement and a null operator analysis of *tough* constructions (in improper movement approaches, an overt DP undergoes Ā-movement; in null operator theories, a covert null operator Ā undergoes Ā-movement). For this reason, this summary does not provide evidence for an embedded Ā-movement step in *tough* constructions, and will simply assume that all *tough* constructions involve some kind of embedded clause Ā-movement chain (though see Chapter 2 of this dissertation for evidence in support of this assumption). However, the final movement step depicted in (5)—the A-movement step from the edge of the embedded clause into the matrix subject position—is the locus of controversy within the *tough* construction literature.

Arguments for an improper movement analysis

The improper movement theory of *tough* constructions capitalizes on the fact that, in many ways, *tough* constructions look like A-movement constructions. For example, *tough* constructions share many similarities with raising constructions. In (6), for example, we see that the matrix subject of a *tough* construction does not appear to receive a Θ-role from the matrix *tough* predicate. If *Jonathan Franzen* received a Θ-role from *tough* in (6a) it would be infelicitous to also describe *Jonathan Franzen* as weak and pathetic, as illustrated in (6b).

(6) a. Jonathan Franzen is *tough* for Anneke to think about _, because he is so weak and pathetic.
b. #Anneke thinks that Jonathan Franzen is tough, because he is so weak and pathetic.

The claim that *tough* predicates do not assign Θ-roles to their matrix subjects is also supported by the fact that (some) idioms retain their non-literal interpretation across *tough* movement, as in (7).

(7) The hatchet is hard to bury after long years of war. (Berman 1973)

Additionally, like raising predicates, *tough* predicate appears to assign a Θ-role directly to their embedded clause. This is demonstrated in (8a), where the embedded clause moves to the matrix subject position, parallel to the raising example in (8b).

(8) a. [$_{CP}$ For Anneke to tolerate Jonathan Franzen] is tough _.
b. [$_{CP}$ For Anneke to tolerate Jonathan Franzen] is unlikely _.

As noted by **?** and Chomsky (1970), *tough* constructions also resist nominalizations in a way that parallels raising constructions, but not control constructions (9).

(9) a. *John's easiness/difficulty to please. (*tough* ✗)
b. *John's certainty/likelihood to win. (raising ✗)
c. John's eagerness to please. (control ✓)

Additionally, as noted by Hartman (2012), *tough* constructions also appear sensitive to defective intervention effects, a hallmark of A-movement operations. Briefly, arguments cannot A-move across other arguments. This is illustrated in (10) for exceptional case marking verbs, where the embedded subject cannot A-move across an oblique experiencer argument.

(10) Defective intervention effects in ECM constructions
a. *The prosecutor proved the defendant$_i$ [$_{PP}$ **to the jury**] [$_{TP}$ $_i$ to be guilty]
b. The prosecutor proved the defendant$_i$ [$_{TP}$ $_i$ to be guilty]

Like ECM predicates, (some) *tough* predicate can also optionally license an oblique experiencer argument. In these cases, the embedded object cannot *tough* move across the oblique argument. The example in (11) shows that it is the

presence of the oblique experiencer argument that restricts *tough* movement. This distribution looks like a hallmark of A-movement (for a more detailed explanation of this analysis, see Chapter 2, Section 4; though see Keine and Poole 2016 for an argument against this approach), suggesting that *tough* construction involve an A-movement step in their derivation.

(11) Defective intervention effects in *tough* constructions
   a. It is enjoyable ([PP to me]) [$_{CP}$ for John to eat strawberries].
   b. Strawberries are enjoyable [$_{CP}$ for John to eat _]
   c. *Strawberries are enjoyable [$_{PP}$ to me] [$_{CP}$ for John to eat _].

Additionally, binding and reconstruction facts suggest that the matrix subject of *tough* constructions must be c-commanded by the embedded subject at some point in the derivation, as in examples like (12) (first proposed by Pesetsky 1987, here modified from Hicks 2009), which suggest that the matrix subject is derived from the embedded object position.

(12) Pictures of themself$_i$ are hard for any photographer$_i$ to ignore.

These facts show that *tough* constructions, in some ways, have the same distribution of as raising construction, a canonical A-movement construction. Additionally, Hartman's (2012) defective intervention facts show that, above the embedded clause, *tough* constructions contain A-movement chain that patterns with other A-movement chains. All these facts suggest that *tough* constructions involve A-movement chain where the embedded object first Ā-moves to the left-edge of the embedded clause and then A-moves to the matrix subject position.

Arguments against an improper movement analysis

There are two main arguments against an improper movement analysis of *tough* constructions[3]. The first argument is a question for Case theory. Under an improper movement theory of *tough* movement, a DP moves from the embedded object position to the matrix subject position. In the embedded object position, this element receives accusative case from the embedded verb. That the verb embedded by *tough* constructions can assign accusative case to its object is illustrated in (13a). However, when the embedded object DP moves to the matrix subject position, it also receives nominative case. That the matrix subject of *tough*

[3]This section will distinguish between arguments against an improper movement analysis of *tough* construction and arguments for a null operator analysis of *tough* constructions

constructions cannot be accusative-marked is illustrated in (13b). Hicks notes that this raises the question of how the embedded object is able to "shed" its accusative case in order to later receive nominative case (or, alternatively, why the embedded object can receive case twice), though this may only be a problem under certain assumptions about case theory and case marking.

(13) *Tough* movement and case
   a. It is difficult [$_{CP}$ for Mary to see *he/him]
   b. He/*Him is difficult [$_{CP}$ for Mary to see _].

The second argument involves the nature of improper movement. Recall that this theory of *tough* constructions proposes that *tough* movement involves a two part movement chain. This movement chain contains an instance of Ā and A-movement, in that order. This order of movement operations is not normally grammatical, hence why it is labeled as "improper." The ungrammaticality of a (non-*tough*) improper movement chain is illustrated in (14).

To be clear, improper movement is not a restriction on having an element undergo both A and Ā-movement. Instead, it is a restriction on A-moving an element that has already been Ā-moved, as in (14), where the DP *which novel* first Ā-moves to the left edge of the embedded clause before being raised into the matrix clause. In (14), a solid line references A-movement, while a dashed line references Ā-movement.

(14) ✗A ← Ā ← DP
*Which novel is possible [*t* Anneke will critique *t*]]

The improper movement chain proposed for *tough* movement looks very similar to the ungrammatical movement chain in (14). If *tough* constructions are improper movement constructions, there needs to be some independent reason why *tough* constructions can allow for improper movement chains while constructions like (14) cannot. Failure to account for why *tough* constructions can allow an improper movement chain incorrectly predicts that a wider number of improper movement constructions will also be grammatical.

1.2.2 The null operator analysis

An alternative approach proposes that *tough* constructions do not involve an A-movement step and instead have the structure outlined in (15). In (15), the

object of the embedded clause is not a phonologically overt DP but rather a silent null operator. This null operator undergoes Ā-movement to the edge of the embedded clause. However, unlike in an improper movement analysis, this null operator remains in the embedded clause and does not move into the matrix clause. In a null operator analysis of *tough* movement, the matrix subject of the *tough* predicate is base generated in the matrix clause. The matrix subject is base generated in the matrix clause and binds with the embedded object null operator, explaining why the object gap and the matrix subject obligatorily share their referent.

(15) Null operator *tough* movement structure
[$_{DP}$ Jonathan Franzen]$_i$ is tough [$_{CP}$ OP$_i$ for Anneke to [$_{vP}$ <OP$_i$> think about <OP$_i$>]]

Again, the idea that *tough* constructions involve an Ā-movement step within the embedded clause is not a controversial claim in any analysis of *tough* constructions. Instead, the controversy surrounding the null operator approach is the fact that this analysis involves an overt DP that is base-generated as the subject of the *tough* predicate, and that a silent null operator (not an overt DP) is the element proposed to undergo Ā-movement in the embedded clause.

Arguments for a null operator analysis

The primary theoretical argument for a null operator theory of *tough* movement is the fact that a null operator analysis avoids requiring an improper movement chain. This style of analysis, therefore, also avoids falsely predicting the grammaticality of other improper movement constructions. Proponents of the null operator analysis (sometimes called the base generation analysis, e.g., Keine & Poole 2017) include Ross (1967), Akmajian (1972) Williams (1983) and Fleisher (2013, 2015), among others.

Additional support for the null operator analysis comes from its compatibility with certain empirical facts. For example, Chomsky (1981) notes that, while expletive *it* can be the subject of a *tough* predicate, expletive *there* cannot be (16a). This is in contrast to raising constructions (16b). Additionally, while examples like (7) suggest that *tough* predicates do not assign a Θ-role to their matrix subject, this fact is not necessarily clear cut. Other idioms no not retain their non-literal interpretation across *tough* movement, suggesting that a *tough* construction might contain a base generated matrix subject that receives a Θ-role from the *tough*

predicate (17).

(16) a. *There is hard to believe to have been a crime committed
b. (%)There seems like there's gonna be a riot (Potsdam & Runner 2001)

(17) *Heed is important to pay to such warning. (* on idiomatic reading)

The contrast in (16) and the obligatorily non-idiomatic interpretation of (17) are difficult to explain if the matrix subject of a *tough* construction is derived. However, these facts follow naturally from a theory where the subject of *tough* constructions is base generated in the matrix clause.

Additional evidence in favor for a null operator analysis comes from weak crossover tests. As noted by Lasnik and Stowell (1991), *tough* constructions do not give rise to weak crossover effects in examples like (18).

(18) John$_i$ should be easy for his$_i$ wife to love

Weak crossover is used as a test to distinguish between Ā-movement (which is sensitive to weak crossover) and A-movement (which is not sensitive to weak crossover). The data in (18) actually suggest two things. First, it supports a null operator analysis of *tough* constructions. Second, it supports a structural analysis as in (19) (modified from Chomsky 1995). In (19) the DP *for his$_i$ wife* is not analyzed as the embedded subject, but instead as an oblique experiencer licensed by the *tough* predicate. Crucially, this oblique experiencer is proposed to be licensed above the embedded CP.

(19) John$_i$ should be easy [$_{PP}$ for [$_{DP}$ his$_i$ wife]$_j$] [$_{CP}$ OP$_i$ [$_{TP}$ PRO$_j$ to love <OP$_i$>]]

In (19), OP$_i$ moves to the embedded spec-CP, where it is bound to the matrix subject. The matrix subject is base-generated in the matrix clause. Because there is no movement of the matrix subject, there is no weak crossover effect. Under an improper movement approach, (19) should be ungrammatical: if the embedded object DP *John* A-moved across the oblique experiencer *his wife*, there would be a relativized minimality violation and the example would be ungrammatical, as in the schema given in (20a).

In (20), it's difficult to analyze the oblique experiencer as the subject of the embedded clause, on either a null operator or improper movement analysis of *tough* constructions. If the embedded object DP Ā-moved to the embedded

spec-CP, it would Ā move across the embedded subject *his wife*. In this case, we would expect a weak crossover effect. Note that an analysis of (19) where *his wife* is interpreted as the embedded subject would also predict a weak crossover effect on a null operator theory of *tough* movement.

(20) John should be easy for his wife to love.

a. $John_i$ should be easy $[_{PP}$ for $[_{DP}$ his_i wife$]_j]$ $[_{CP}$ <$John_i$> $[_{TP}$ PRO_j to love <$John_i$>]]
(relativized minimality violation: should be ✗)

b. $John_i$ should be easy $[_{CP}$ <$John_i$> for $[_{DP}$ his_i wife$]_j$] to love <$John_i$>
(weak crossover violation: should be ✗)

While these facts do not represent the full range of arguments in support of the null operator analysis to *tough* movement, they do suggest that at least some aspects of the distribution of *tough* constructions are difficult to capture under an improper movement analysis. These facts provide reason to believe that an alternative theory may be needed to fully capture the *tough* movement facts.

Arguments against a null operator analysis

Under a null operator analysis of *tough* constructions, there is no movement relationship between expletive *tough* predicate constructions and *tough* movement constructions, as in (21). Because of this, there is no way to capture the fact that these constructions appear synonymous with each other. In fact, if the matrix subject of (21b) is base generated (a requirement of the null operator analysis), we would need a theory where *easy* (and all other *tough* predicates) had two forms: an intransitive form that could only assign a Θ-role to the clause it embeds (as in (21a)) and a transitive form that assigns a Θ-role to both its matrix subject and the clause that it embeds, as in (21b). Under this approach, we would predict that (21a) and (21b) would not be synonymous with each other, as (21b) would contain an additional Θ-role, relative to (21a).

(21) a. It is easy for Anneke to criticize Jonathan Franzen.
b. Jonathan Franzen is easy for Anneke to criticize.

Obviously, this is not the only argument against a null operator theory of *tough* movement—a null operator theory of *tough* movement would also struggle to capture the raising/*tough* movement parallels as well the defective intervention facts discussed in Section 2.1.1. However, hopefully this section that shown that

there are arguments both in favor and against a null operator theory of *tough* movement.

1.2.3 Complications to both approaches

One complication to the *tough* syntax debate is the fact that many of the *tough* movement facts are compatible with both an improper movement and a null operator analysis, while other facts appear incompatible with both proposals. Take, for example, the binding facts from (12), repeated in (22). Hicks (2009) uses the example in (22) to argue against a null operator *tough* movement analysis, it shows that the matrix subject must be able to reconstruct below the embedded subject at some point in the derivation. These facts are, of course, compatible with an improper movement approach, as discussed in Section 1.2.1. However, these facts are also compatible with a null operator analysis, at least assuming the theory of null operator binding put forth by Nissenbaum (2000), which allows for null operators to reconstruct: if the null operator, which is bound to the matrix subject, Ā-moves to the left-edge of the embedded clause, it should still be able to reconstruct below the embedded subject.

(22) Pictures of themself$_i$ are hard for any photographer$_i$ to ignore

Additionally, some syntactic tests yield contradictory results when applied to *tough* constructions. Idiom tests, which tie into the question of whether or not the *tough* predicate assigns a Θ-role to its matrix subject (and whether or not the matrix subject should be understood as derived or base generated), are perhaps the clearest example of this. As set of idiom facts can be used by proponents of an improper movement approach to show that *tough* predicates do not assign a Θ-role to their matrix subjects, while a different set of idiom facts can be used by proponents of a null operator approach to show that *tough* predicates do assign a Θ-role to their matrix subjects. Neither side can explain the full range of the data. It is, in fact, difficult to imagine an alternative analysis that could capture this apparent contradiction.

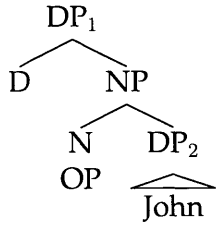
This dissertation, when it needs to reference the structure of *tough* constructions or discuss aspects of the Ā-movement chain within their embedded clauses, will represent *tough* constructions as improper movement constructions. Largely, this is an arbitrary choice, and should not necessarily be taken as evidence for or against an improper movement theory of *tough* movement. That

being said, Chapter 4 of this dissertation does contributes experimental evidence that, I believe, is not compatible with a null operator theory of *tough* constructions, and so favors an improper movement approach. In this way, one portion of this dissertation does take a stand within the *tough* movement debate. The remaining chapters of this dissertation, however, are prediction neutral across both a null operator and improper movement analysis.

1.2.4 An hybrid improper movement/OP approach

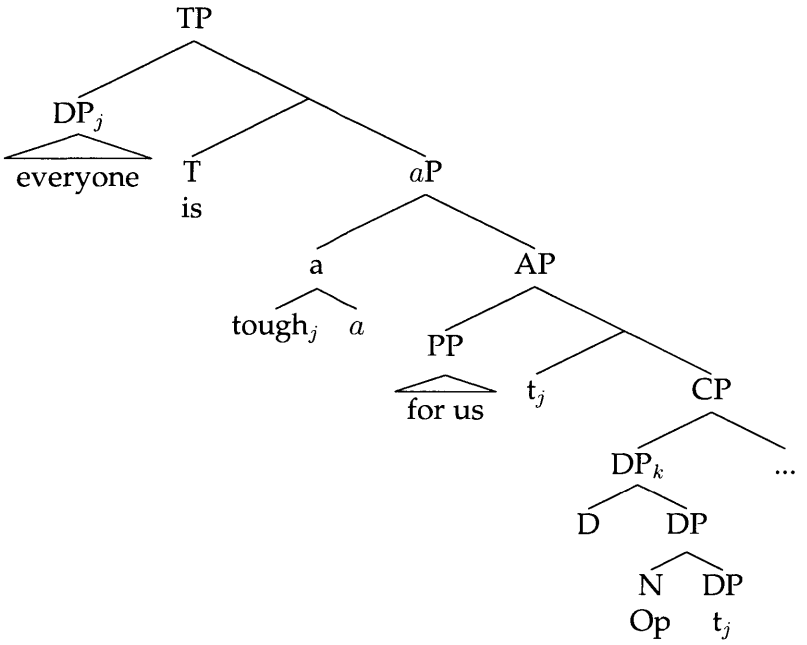
Recently, Hicks (2003, 2009) has proposed a hybrid (improper) movement/null operator analysis of the syntax of *tough* constructions. In Hicks' approach, null operators can be thought of as syntactically complex elements, as in (23).

(23) Hick's null operator structure:



Hicks proposes that *tough* constructions contain a null operator that is different from the kind of null operator found in, for example, parasitic gap constructions. One way that Hick's *tough* operator is different from standard null operators is that is not necessarily null. In Hick's analysis of *tough* construction, object of the embedded clause is a syntactically complex null operator like (23). This operator Ā-moves to the edge of the embedded clause. After Ā-movement, the DP embedded within the complex operator (DP_2 in (23)) sub-extracts from within the complex operator (DP_1 in (23)) and A-moves to the matrix subject position, as in (24). Because the entire DP_2 was the argument that underwent Ā-movement, DP_1 has not technically been Ā-moved. To Hicks, this means that A-moving DP_2, technically does not violate a ban on improper movement. Hicks' full *tough* movement derivation is given in (24).

(24) Hicks' *tough* movement syntax
Everyone is tough for us to please.



This dissertation will not adopt Hicks' analysis of *tough* movement. First, while Hicks proposes that this analysis is able to capture the best parts of the movement and null operator approaches, it is not clear that his analysis has any more predictive power than either the null operator or movement approach. For example, it captures the Θ-role facts from the improper movement camp, but fails to capture the Θ-role facts from the null operator camp. Additionally, unlike a null operator representation of (24), relativized minimality predicts (24) to be ungrammatical, as DP$_j$, *everyone*, must A-move across *for us* (analyzing *for us* as the embedded subject would falsely predict that *tough* constructions are sensitive to weak crossover facts).

Obviously, these are real problems for any analysis of *tough* movement. However, Hicks' approach does not appear to capture a wider range of facts than either the null operator approach or the improper movement approach, and so it is not clear that it should be favored over any existing approach. This is particularly true given the uniqueness of Hicks' null operator, which he does not claim appears elsewhere in the grammar[4]. However, Hicks also does not propose a mechanism for limiting the distribution of the *tough* construction null operator; as this section will later discuss, allowing this type of null operator to appear

[4]Though Hicks does liken this null operator to certain analyses of clitics as complex elements.

outside of *tough* constructions potentially makes a wide range of false predictions.

Hicks' main argument for his proposal is that it technically avoids an improper movement chain. While this is, technically, true, it stands to ask whether or not the distinction between what, technically, qualifies an improper movement and the movement depicted in Hicks' proposal is meaningful. For example, Hicks' proposal predicts that the (25) should be grammatical since, even though (25) involves A-moving a sub-constituent of a DP that has undergone Ā-movement, it is not technically improper movement. Additionally, it is not clear that any part of Hicks' analysis limits the same kind of complex null operator movement from occurring in improper movement constructions like (26). In this way, Hicks' analysis may also predict the grammaticality of ungrammatical improper movement constructions, as well.

(25) a. *[$_{DP_i}$ Which novelist] is possible [[$_{DP}$ pictures of <DP$_i$>]$_j$ Anneke will critique $\__j$?
b. *[$_{DP_i}$ Which novelist] is possible [[$_{DP}$ OP <DP$_i$>]$_j$ Anneke will critique [$_{DP}$ picture of $\__j$]]

(26) *[$_{DP_i}$ Which novel] is possible [[$_{DP}$ OP <DP$_i$>]$_j$ Anneke will critique $\__j$]

Additionally, I believe that Hicks' proposed complex null operator structure—which allows for a covert null operator to contain a fully referential and overt DP—makes incorrect predictions for other types of null operator constructions. Null operators enter the derivation as semantically vacuous elements which achieve their obligatory interpretation through semantic binding with a higher predicate (c.f. Nissenbaum 2000). In Hick's system, the full DP internal to the complex null operator determines the reference of the entire complex operator. First, this predicts that null operators should be able to occur in contexts where they do not have to bind to a higher operator. This incorrectly predicts, for example, that there should be a counterpart to parasitic gaps that are not "parasitic" on a previous instance of Ā-movement, and simply serve as adjunct clauses, as in (27).

(27) Null operator structures predicted to be grammatical by Hicks (2009)
*Anneke burned Jonathan Franzen's latest novel [$_{CP}$ after reading OP]

Another consequence of this structure is the prediction that null operators can refer to DPs other than the arguments of a higher clause. For example, this null operator structure predicts that the parasitic gap in (28a) should have any of the

interpretations given in (28b)-(28d).

(28) a. Which book did Anneke burn _ [$_{CP}$ after reading *pg*]
b. Which book did Anneke burn _ [$_{CP}$ [$_{OP}$ OP [$_{DP}$ which magazine]] after reading _]
c. Which book did Anneke burn _ [$_{CP}$ [$_{OP}$ OP [$_{DP}$ which newspaper]] after reading _]
d. Which book did Anneke burn _ [$_{CP}$ [$_{OP}$ OP [$_{DP}$ which article]] after reading _]

1.2.5 *Tough* movement summary

This section has outlined the debate between the improper movement and null operator theories of *tough* syntax. This debate surrounds a difficult and complicated dataset, and it is not clear if either approach is totally suited to capture the full distribution of *tough* constructions. Because the experimental chapter of this dissertation presents soft evidence in favor of an improper movement approach to *tough* movement, this dissertation will represent *tough* constructions as improper movement constructions. However, this is by no means the final word in this debate, and determining exactly how to analyze the syntax of *tough* constructions is an important question for future research. This section has also outlined Hicks' alternative theory of *tough* movement, which this dissertation will not adopt.

1.3 Phases and embedded clause sized are determined dynamically

This dissertation adopts a theory of syntax where phase boundaries are dynamically determined. This section will discuss the ways in which stripping and fragment answers provide empirical support for the contextually specified theory of phasehood that this dissertation will adopt.

In biclausal constructions, this theory has consequences for determining embedded clause size, as well. For instance, if the embedded phase is a bare TP, the embedded clause must only extend up to the TP layer, as well. In particularly, this dissertation adopts the dynamic theory of phasehood advocated for in Wurmbrand (2017), following Željko Bošković (2014). Wurmbrand's proposal was originally used to explain the stripping and fragment answer asymmetries given

in (29)-(30), where the inclusion of the complementizer in an ellipsis construction leads to ungrammaticality (as in (29a-c) and (30a). However, in both examples, when the overt complementizer is absent, the constructions are grammatical (29b)-(30b).

(29) Stripping
  a. *Abby claimed (that) Ben would ask her out, but she didn't think that Bill (too).
  b. Abby claimed (that) Ben would ask her out, but she didn't think Bill (too).
  c. *Jane loves to study rocks, and John says that geography too.
  d. Jane loves to study rocks, and John says geography too.

(30) Fragment answers (Wurmbrand 2017:351, citing Morgan 1973:732)
Q: How does Nixon eat his tapioca?
  a. *I think that with a fork.
  b. I think with a fork.

This section will discuss how a dynamic approach to phasehood (as well a theory of ellipsis where elided constituents are unpronounced Spell-Out Domains) can account for the alternations in (29)-(30). First, this section will detail why a static theory of phasehood is theoretically undesirable when compared with a dynamic theory of phasehood. This section will then discuss Wurmbrand's assumptions about ellipsis constructions more generally. Finally, this section will move to Wurmbrand's (2017) analysis of stripping constructions and fragment answers.

1.3.1 Static approaches to phasehood and the "list problem"

Traditional theories of syntax propose a tight connection between certain syntactic categories—particularly CP and (many instances of) vP—and phasehood. This static approach to phasehood, dating back to Chomsky (2000), relies on a certain arbitrariness: it is not immediately clear why vP and CP are phases, to the exclusion of, for instance, TP. In matrix clauses (where the CP marks the sentence as propositional, and the TP is part of the cyclic domain that extends up to the CP) this kind of distinction can seem intuitive. However, this line blurs when other kinds of syntactic structures are considered. Wurmbrand (2017:7) states that "in relative clause CPs it is particularly difficult to see in what way the CP is propositional but the TP is not–semantically exactly the opposite would hold."

This is part of what Wurmbrand, citing Grano and Lasnik (2015:7), describes

as the "list problem" of phasehood: "the set of phase types has to be stipulated rather than following from something more general". In this way, a static theory of phasehood can reduce to a description of the distribution of phases, not a theory that is able to predict where phases will (and will not) occur. This problem is additionally complicated when a wider rage of syntactic structures is considered. Take, for example, the set of non-agentive verbal constructions, including passives, unaccusatives and raising verbs (discussed in Wurmbrand 2017:346). These constructions all lack the phasal agentive vP. However, their vP/VP layer also appears to function as a phase. Evidence of this is given in (31), from Sauerland (2003)[5] for raising verbs. The example in (31a) shows that a raised subject can be interpreted below the scope of negation. For this reading to be possible in the grammatical example in (31a), the raised subject must be able to be interpreted below the scope of negation, but above the bound pronoun *his*$_i$. The only position that meets these requirements is the labeled spec-vP, and *every child*$_i$ would only have reason to move through this spec-vP position if it were a phase edge, as illustrated in (31b). If the raising verb contains a phrasal v, we predict that the subject will move successive cyclically through spec-vP during movement to the canonical subject position. However, if v were not phrasal, we would not predict this movement to take place.

(31) a. Every child$_i$ doesn't seem to his$_i$ father to be smart. $\forall > \neg; \neg > \forall$
'It's not the case for every child that the child seems to his father to be smart'
b. Every child$_i$ doesn't [$_{vP}$ ✓ seem to his$_i$ father [* to be smart]]

Wurmbrand, (following Wurmbrand and Haddad 2016, among a number of other proposals, including Grohmann et al. 2000), also shows that there are constructions where a bare TP must be phasal, including the bare TP embedded by a raising predicate. The example in (32a) illustrates four phase boundaries in raising constructions, while (32b) illustrates the three positions that a subject must successive-cyclically move through before arriving in spec-TP, the canonical subject position (labeled ① in (32)).

(32) a. [$_{CP=phase}$ John [$_{vP=phase}$ seems [$_{TP=phase}$ to have [$_{vP=phase}$ left]]]]
b. [$_{CP=phase}$ [$_{TP_1}$ ① T [$_{vP_1=phase}$ ② [$_{TP_2=phase}$ ③ [$_{vP_2=phase}$ ④]]]]]

[5]See Sauerland 2003 for additional arguments that non-agentive verbal constructions are still phasal.

The most controversial phase position here is spec-TP_2 (see example (31) for a discussion of why the raising v should be understood as phasal). Wurmbrand notes that there is not evidence from English that subjects must move successive-cyclically through spec-TP_2. However, citing Wurmbrand & Haddad (2016), Wurmbrand (2017) claims that Arabic shows evidence of this kind of movement, as illustrated in (33)[6]. In the Arabic example in (33a), the subject *l-tˤa:liba:t-i* occurs in spec-TP_2. Evidence that the subject occurs in this position comes, in part, from the fact that the subject can only occur with accusative case (and is ungrammatical with nominative case, as in (33b)). This case asymmetry, Wurmbrand claims, shows that the subject is spelled out together with the matrix VP once the matrix vP phase is complete. This would only be possible if the subject were in position (3) (spec-TP_2). If the subject were in position (2), it would need to be spelled out with the matrix TP and so would require nominative case. For a closer examination of these examples, and additional arguments that Arabic allows subjects to be pronounced in the phasal spec-TP_2 position, see Wurmbrand (2017) and Wurmbrand & Haddad (2016).

(33) a. jumkinu [$_{TP_2}$ **l-tˤa:liba:t-i** ʔan
be.possible.3.MASC.SG [$_{TP_2}$ the-students.FEM-ACC to
jaʒaħna]
succeed.3.FEM.PL]
'It is possible for the female students to succeed.'

b. *jumkinu [$_{TP_2}$ **l-tˤa:liba:t-u** ʔan
be.possible.3.MASC.SG [$_{TP_2}$ **the-students.FEM-NOM** to
jaʒaħna]
succeed.3.FEM.PL]

Given this type of evidence, Wurmbrand (2017) proposes a modified theory where phases are contextually (or dynamically) determined, dependent on the structure of the sentence in question. Wurmbrand's proposal, which has its roots in Chomsky (2001), says that phases do not occur due to the presence of a single head but rather emerge as a result of a clausal domain being "'complete' by some linguistic measure." I will take this linguistic measure to mean that the domain is (i) propositional (ii) does not contain a syntactic head with any active Θ-roles left to discharge and (iii) is able to be embedded as an argument of a higher predicate. This theory of phasehood additionally solves the so-called "list problem." Under a dynamic theory of phasehood, phase heads do not need to be listed. Instead

[6]Similar evidence can be found in discussions of English quantifier float, see McCloskey 2000.

phase heads can be contextually determined given the structure of the construction in question. Wurmbrand (2017) provides a dynamic definition of phasehood in (34)[7].

(34) Contextually determined phasehood (Wurmbrand 2017:346)
- a. The highest projection of a cyclic domain constitutes a phase
- b. The cyclic domains are:
  - (i) the extended thematic domain of V[8]
  - (ii) the combined T and C domains

This analysis can also be thought of as a re-framing of Grimsahw's (2000) concept of the extended projection: in Wurmbrand's theory the grammar allows for two extended projections[9]: the verbal domain where arguments are introduced and connected and the combined TP/CP domain, where the verbal proposition is interpreted with respect to tense, aspect, modality and information structure. Every sentence includes at least one instance of each of these modified extended projections, and the highest head in the modified extended projection is the phase head. In cases where the combined TP/CP projection is also embedded as the argument of some higher predicate, the phase edge of the TP/CP projection also signifies the boundary of the embedded clause.

1.3.2 Zero Spell-Out and ellipsis

Wurmbrand's (2017) theory of stripping also depends on an analysis of ellipsis as a form of Spell-Out. Wurmbrand (2017) assumes, following Gengel (2006, 2009), Gallego (2010) and Boškovič (2014), among others, that phase heads trigger the Spell-Out of their complements and that ellipsis constructions constitute unpronounced Spell-Out domains. One consequence of this analysis is that only the complements of phases can be elided.

In non-ellipsis constructions, the entire complement of the phase head is

[7]This definition excludes categories like DPs, which are also argued to be phases (c.f., Szabolsci 1994). Wurmbrand suggests seeing Boškovič 2004 for a definition of cyclic domain which includes these projections

[8]One potential consequence of my definition of "linguistic completeness" (given on the previous page) is the possibility that both V and v might emerge as phases in the same clause. While this does not necessarily make false predictions, we also do not have empirical evidence to support this claim. Wurmbrand (2017) avoids this issue by bifurcating the clause into two extended domains: the extended vP/VP domain and the extended TP/CP domain. This is, fundamentally, a reframing of Grohmann's (2000) notion of extended projections. This dissertation will adopt this theory.

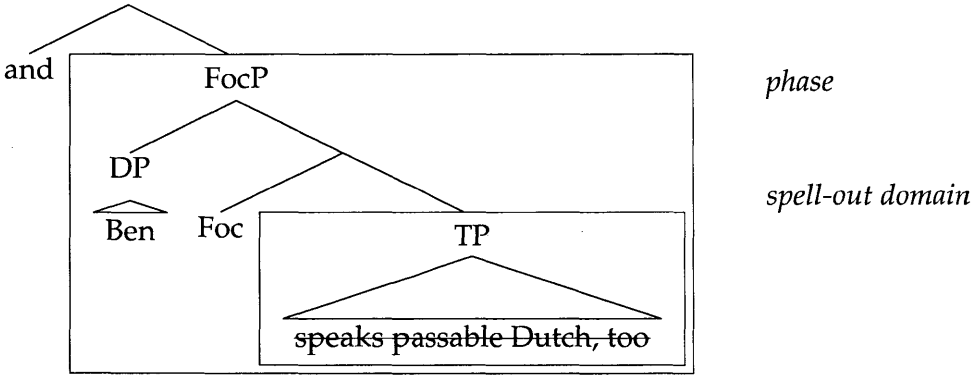
[9]Three, if you're including DP phases: the projection of the argument itself.

Spelled-Out. The specifier of the phase head and the phase head itself, however, are not Spelled-Out. This fact underlies operations like successive-cyclic movement. This fact also makes predictions about ellipsis. Particularly, if the complement of a phase head can be elided (and therefore unpronounced), the phase heads and its specifier will still be pronounced (i.e., they will not be elided) in an ellipsis construction.

1.3.3 Stripping data & topless complementizers

Following Merchant (2003), Wurmbrand proposes that the remnant of stripping occurs in a focus position in the left periphery. Here, a dynamic theory of phasehood and an analysis of ellipsis as unpronounced Spell-Out domains straightforwardly explains the stripping data. Under this theory, stripping involves a bare FocP. Here, Foc^0 is the phase head and so the complement of Foc^0 is the spell out domain. This straightforwardly explains why the canonical stripping examples are grammatical: the top projection of the cyclic domain is the FocP, thus the TP complement of Foc^0 constitutes a spell-out domain that can be elided, as in (35). Note that the English Foc^0 is phonologically null.

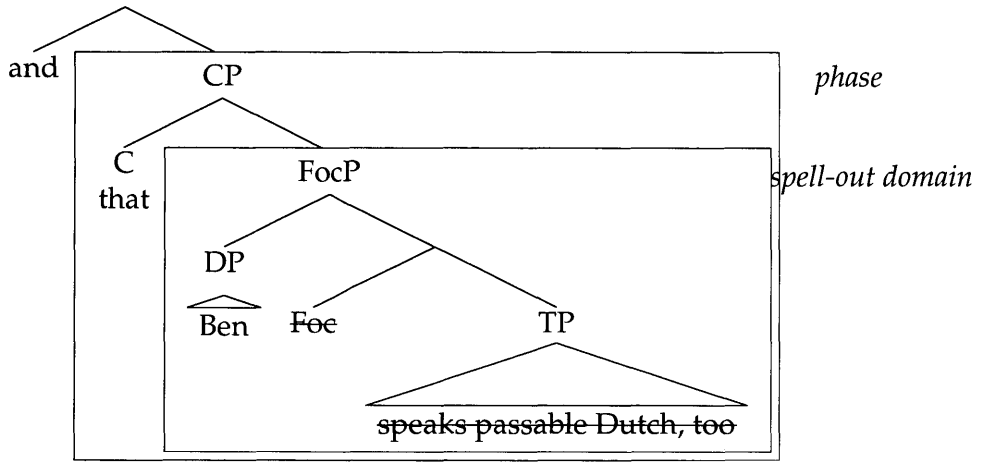
(35) Abby can speak passable Dutch, and BEN too.



This analysis also explains why stripping cannot occur when the embedded clause contains an overt complementizer *that*, as in (36). In this case, the top projection of the cyclic domain would be CP, above FocP. As a result, the entire FocP would constitute the SOD targeted by ellipsis. It would be impossible to elide only the TP in (36), as the TP is only part of the SOD—the full spell-out domain is instead the FocP containing both *Ben* and the TP. If we can assume that *that*-less embedded clauses truly lack a CP layer, the stripping facts given earlier

in (29) can be easily accounted for[10].

(36) *Abby can speak passable Dutch, and that BEN too.



The same analysis can also explain parallel contrasts with fragment answers. Following Merchant (2003), Wurmbrand proposes an ellipsis account of fragment answers that involve a FocP with a remnant argument in their specifier, as illustrated in (37). If this FocP is both the highest projection and the phase boundary, then all material below the Foc^0 can be completely elided. This also correctly predicts that fragment answers cannot occur when a CP embeds the FocP, as illustrated in (37b)-(37c).

(37) Q: How does Nixon eat his tapioca?

a. I think [$_{FocP}$ [$_{PP}$ with a fork] ~~[$_{SOD}$Nixon eats his tapioca _]~~]

b. *I think [$_{CP}$ that ~~[$_{SOD}$ [PP with a fork] Nixon eats his tapioca _]~~]

c. *I think [$_{CP}$ that [$_{FocP}$ [$_{PP}$ with a fork] ~~Nixon eats his tapioca~~]]

1.4 Structure of this dissertation

The remainder of this dissertation is structured as follows:

Chapter 2 directly compares *tough* constructions and gapped degree phrases, and argues that the two constructions are very similar. Particularly, Chapter 2

[10]See Wurmbrand (2017:349) for discussions of why this analysis, when combined with the existing ellipsis literature, does not predict the grammaticality of examples like "Abby claimed that Ben would ask her out, but she didn't think [$_{CP=phase}$ that ~~[$_{FocP=SOD}$ Bill would ask her out]~~]." Discussing how and why Wurmbrand (2017) makes the correct predictions for other types of ellipsis constructions involves a thorough discussion of the existing ellipsis literature, and is beyond the scope of this dissertation.

argues that *tough* constructions and gapped degree phrases contain the same movement chain. The difference between the constructions, however, lies in the size of their embedded clauses: while *tough* constructions contain an embedded CP, the embedded clause in a gapped degree phrase is a DegP.

Chapter 3 explores why gapped degree phrases allow for subject gaps, while subject *tough* movement is ungrammatical. This chapter proposes that this asymmetry is causes by a spec-to-spec anti-locality constraint, like that proposed in Erlewine (2016) or Brillman and Hirsh (pear). Chapter 3 also argues that this anti-locality constraint can unify a wide range of subject/non-subject asymmetries, both within English and cross-linguistically.

Chapter 4 asks whether the subject/non-subjects that occur in the adult grammar are replicated in the child grammar. To do this, this chapter presents the results of a language acquisition experiment that seeks to understand whether children find subject and object *tough* constructions and subject and object raising constructions grammatical. The results of this experiment show that the distribution of subject and object gaps is not the same across the child grammar and the adult grammar. Unlike adults, children find both subject and object raising constructions grammatical and subject and object *tough* constructions ungrammatical. This chapter discusses this experiment and what it might mean for both theories of acquisition and syntax.

Finally, Chapter 5 concludes the dissertation.

Chapter 2

Gapped Degree Phrases and Tough Constructions

2.1 Introduction

This chapter focuses on an unusual embedded clause construction: gapped degree phrases (1). The unique behaviors of gapped degree phrases have been noted since Lasnik & Fiengo (1974), and the literature has since investigated both their syntax (e.g., Chomsky 1977) and semantics (Nissenbaum & Schwarz 2011).

(1) a. Jonathan Franzen is too self-absorbed for Anneke to read _.
(object GDP)

b. Anneke is too intelligent _ to read Jonathan Franzen.
(subject GDP)

Syntactically, GDP are mysterious because of their parallels to *tough*-constructions, a "notorious" construction that displays a series of "apparently contradictory empirical properties" (Hicks 2009, see the Introduction to this dissertation for further discussion). Similar to *tough*-constructions, gapped degree phrases appear to display an improper movement chain within their derivations (see Chomsky 1977 for evidence of an Ā-movement step in gapped degree phrases; Brillman 2015, for evidence, following Hartman 2009, 2012, of an A-movement step in gapped degree phrases). Semantically, gapped degree phrases are mysterious because they appear closely related to parasitic gaps, though the gap found in gapped degree phrases is not obviously parasitic on any previous movement operation (Nissenbaum & Schwarz 2011).

This chapter argues that the unusual behaviors of gapped degree

phrases—across both the syntax and the semantics—can be straightforwardly explained by assuming that gapped degree phrases have *a larger-than-usual embedded clause structure*. This can be understood as meaning that the embedded clause in gapped degree phrases has an additional layer above the CP. This layer is a DegP layer that contains the degree word (e.g., *too* or *enough*). Thus, gapped degree phrases can be understood as biclausal constructions where everything embedded under the AP is part of the embedded clause, similar to other adjectival embedding constructions (e.g., control and raising constructions). Because the DegP is the beginning of the embedded clause, Deg^0 acts as a phase head within the syntax (following Wurmbrand's 2017 theory of contextually determined phasehood, outlined in Chapter 1).

This chapter will focus primarily on the movement chain that occurs within the embedded clause in gapped degree phrases. Analyzing gapped degree as having an embedded clause that extends beyond CP to DegP allows gapped degree phrases to be understood, informally, as involving an instance of *tough*-(like-)movement entirely within their embedded clauses. The addition of the CP layer also explains why gapped degree phrases appear similar to parasitic gaps. This requires assuming Nissenbaum's (2000) analysis of parasitic gaps—which proposes that parasitic gaps are parasitic on a particular type configuration within the semantics, not necessarily a previous instance of movement (though the two often overlap). In a gapped degree phrases, the DegP is the only position in the LF where the semantic operation COMPOSE can take place. This operation binds the matrix subject to the embedded gap site.

Put another way: this chapter will show that the presence of a CP-dominating DegP layer is crucial for explaining both the movement patterns of gapped degree phrases, as well as the obligatory co-reference between the matrix subject (*Jonathan Franzen*) and the embedded gap site (*to read* __).

As discussed in Chapter 1, this dissertation argues for a theory of syntax which includes a more flexible understanding of embedded clause size (where there is no "one size fits all" embedded clause size) and allows for embedded clauses that are (at least) larger than the traditional CP/TP structure. This chapter contributes to this goal by showing that, while gapped degree phrases initially appear to exhibit a variety of seemingly exotic behaviors, these behaviors can be explained by assuming that gapped degree phrases have an embedded clause headed by DegP, above CP. This chapter will primarily focus on object gapped degree phrases, which bear the closest surface similarity to *tough*-constructions.

Object gapped degree phrases are also more obviously parallel to a variety of parasitic gap constructions. The movement chain involved in subject gapped degree phrases will prove to be somewhat more complex than that of object degree phrases, and one behavior of subject gapped degree phrases—the fact that subject extraction is licensed at all in gapped degree phrases, compared with *tough*-constructions, which prohibit subject extraction entirely—will require more than an extended embedded clause structure to explain. Don't despair. Subject gapped degree phrases will be re-examined in Chapter 3.

The structure of this chapter is as follows: Section 2 provides a brief discussion of previous semantic analyses for gapped degree phrases, with a focus on Nissenbaum & Schwarz (2011), whose analysis this chapter will adopt. Section 2 will also devote some attention to the semantics necessity for a structural layer that is simultaneously above CP and still a part of the embedded clause. Section 3 provides a brief discussion of the syntax and semantics of degree words themselves. This section will also discuss the fact that degree words can introduce an optionally overty evaluator argument into the derivation, and discuss where these arguments must occur (Meier 2003). Section 4 gives evidence in favor of analyzing the movement chain within the gapped degree phrase embedded clause as improper movement, meaning that it involves both an $\bar{A}$ bar movement (section 4.1) and an A-movement (Section 4.2) step. This section will pay special attention to how the extended embedded clause structure allows for this movement chain without predicting the existence of improper movement chains across other $\bar{A}$ constructions. Section 5 provides a brief discussion of subject gapped degree phrases. Section 6 concludes the chapter.

2.2 Background: the semantic structure of gapped degree phrases

One concern of this chapter is the relationship between gapped degree phrases and *tough*-constructions. While *tough*-constructions and gapped degree phrases share many similarities (see Section 4 for a full description of the similarities between *tough*-constructions and gapped degree phrases), the structures differ in terms of Θ-role assignment. While *tough*-constructions crucially *do not* assign a Θ-role to the subject of their higher clause, gapped degree phrases crucially *do*. Thus, while (unmoved) *tough*-constructions can occur with an expletive *it* in

matrix subject position, there is no parallel possibility for gapped degree phrases (2).

(2) a. It's tough for Jonathan Franzen to talk to women.
expletivie subject possible in TC

b. *It's too arrogant for Jonathan Franzen to talk to women.
expletive subject not possible GDP

Additionally, *tough*-constructions cannot contain a non-expletive subject if they contain a *for* phrase without a gap. This is illustrated in (3a), which appears to fail precisely because no Θ-role is assigned to *Olivia*, the matrix subject. Gapless degree phrases, like (3b), can occur with a matrix subject and an embedded clause, further evidence that gapped degree phrases do assign a Θ-role to their matrix subject.

(3) a. *Anneke is tough for anyone to ignore her critiques of Jonathan Franzen.
syntactic subject impossible in TC

b. Anneke is too smart for anyone to ignore her critiques of Jonathan Franzen.
syntactic subject possible GDP

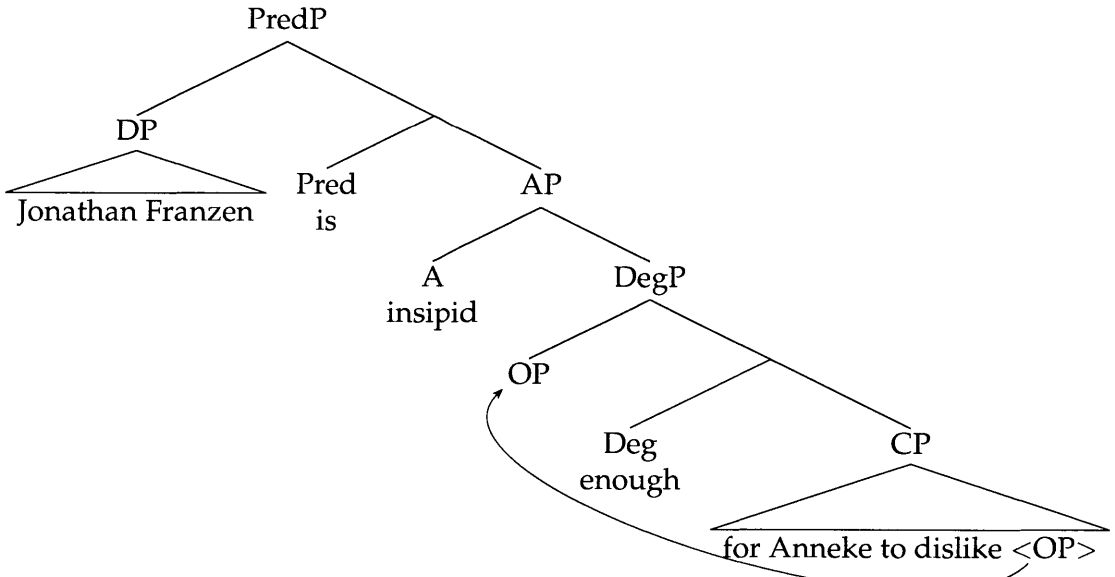
This Θ-role asymmetry can be accounted for by claiming that gapped degree phrases, but *not tough*-constructions, are *null operator* constructions. This paper will adopt an analysis of *tough*-constructions where an overt DP undergoes movement to the higher clause and is a assigned a Θ-role in its base position, not its derived position. This is, of course, not the only analysis of *tough*-constructions proposed within the literature. See the introduction to this dissertation for an overview of the two most popular approaches to *tough* movement within the literature.

In *tough*-constructions, then (under an overt DP interpretation), an overt DP undergoes movement and is assigned a Θ-role only in its base position. In gapped degree phrases, it is not the matrix subject itself but rather a *null operator* that moves. Thus, the subject of a *tough* construction can be understood as a derived subject, while the subject of a gapped degree phrases can be understood as base generated. In gapped degree phrases, the null operator movement step is entirely internal to the embedded clause; the null operator never moves into the matrix clause. After the null operator movement chain is completed, the null operator is

semantically linked to its antecedent in the matrix clause[1].

A null operator movement account of gapped degree phrases requires a particular constituent structure, given in (4) and taken from Nissenbaum & Schwarz. This constituent structure corresponds to the word-order of *enough*-type degree phrases, but not the word-order of *too*-type degree phrases[2]. The structure for (4) leads to the following (very informal) semantics: 'Jonathan Franzen is insipid to a degree, *d*. This degree is greater than or equal to the degree of insipidness that allows Anneke to dislike an author's collected works."

(4) Jonathan Franzen is insipid enough for Anneke to dislike.



The exact specifications of the movement chain in (4) are left intentionally vague for the moment—for a full description of DegP-internal movement, see Section 4. Nissenbaum & Schwarz's (2011) semantic analysis of gapped degree phrases[3] requires exactly this movement chain for two interconnected reasons.

[1]Another way to think of this is in terms of *semantic types*. *Tough*-constructions involve an adjective that relates a situation (or a set of truth conditions) to degrees (e.g., of toughness, ease, difficulty); these adjectives are of type <st<d,st>>. Gapped degree phrases involve an adjective that relates an individual to degrees (e.g., of shyness, happiness, intelligence); these adjectives are of type <e<d,st>>.

[2]This structure is minimally modified from Nissenbaum & Schwarz (2011). Their original structure had the DegP as a *leftward* complement to the AP. In the structure provided here, the DegP is a *rightward* complement of the AP. This change is made largely for clarity, as it better reflects English word-order and does not (to the best of my knowledge) make any predictions that diverge from the original structure proposed in Nissenbaum & Schwarz 2011

[3]Nissenbaum & Schwarz call gapped degree phrases *parasitic degree phrases*; this terminology is tied to their specific analysis of gapped degree phrases: as a structure that utilizes the same technology and semantic composition mechanism as parasitic gaps. This dissertation adopts a

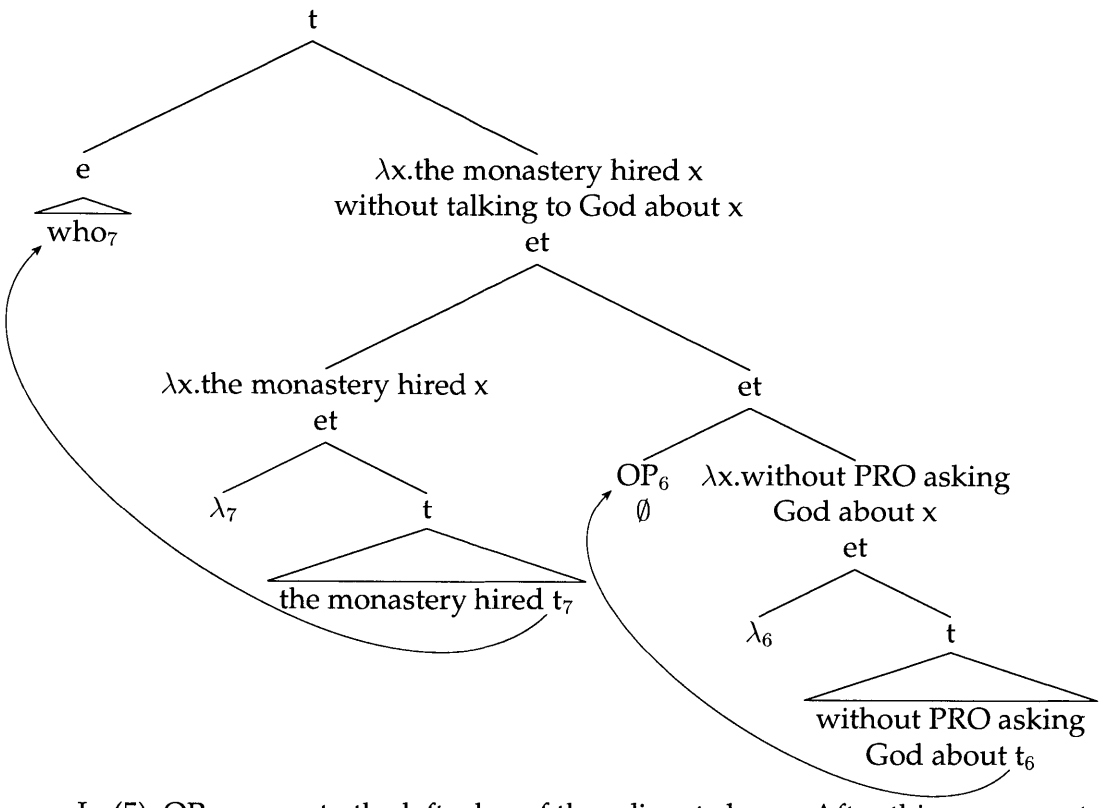
First, this movement chain is required to link the null operator to the matrix subject. Second, this movement chain is needed to create a DegP of a semantic type that can compose with the matrix clause without resulting in a type clash. Before null operator movement, the embedded clause is of type <d<st>><st>, while after null operator movement (to the edge of the embedded clause) creates a node of type <e<d<st>><st>>, and only a type <e<d<st>><st>> node can combine with an adjective of type <e<d<st>>> (for more details on this, see the structure in (7)). Without this null operator movement, it would be impossible to construct a well-behaved LF for the gapped degree phrase, and it would also be impossible to guarantee that the gap site and the matrix subject refer to the same entity.

Understanding why this is true involves taking a closer look at the semantics. Semantically, combing a DegP containing a null operator with the matrix clause via a semantic operation (like predicate modification) is the only way to ensure that two arbitrary arguments (the null operator and the matrix subject) refer to the same entity. Without movement of the null operator to the left-edge of the degree phrase, this kind of semantic combination would not be possible. Nissenbaum & Schwarz (2011) propose that the filler-gap relation between the matrix subject and embedded gap site in gapped degree phrases is analogous to filler-gap relations in parasitic gap constructions. In parasitic gaps, these relationships are established via two parallel instances of movement (and lambda abstraction) and predicate modification, as illustrated in (5). Because parasitic gap constructions are comparatively simpler than gapped degree phrases, this section will review the semantics of parasitic gap constructions before moving on to the semantics of gapped degree phrases.

---

more theory-neutral name for the construction.

(5) LF for a Nissenbaum-style parasitic gap



In (5), OP_6 moves to the left edge of the adjunct clause. After this movement operation has taken place, the adjunct clause—which was initially type <t>—is type <e,t>. The adjunct clause is able to combine with the matrix clause because the matrix clause *also* involves an instance of movement, or, more specifically, because the matrix clause involves an instance of lambda extraction that creates a type <e,t> node below the *wh*-word's final landing site (directly above λ_7). These two type <e,t> nodes—the matrix node above λ_7 and the adjunct node above OP_6–can combine with each other via predicate modification. Importantly, the adjunct clause must combine with the matrix clause below the *wh*-word (it must "tuck in"). A crucial consequence of combining the clauses via predicate modification is that *who* will be understood as both the antecedent of the matrix gap *and* the adjunct gap. For more details about this process, and arguments in favor of analyzing parasitic gaps structures with the semantic representation given in (5), see Nissenbaum (2000).

According to Nissenbaum (2000), predicate modification is sufficient to derive most, but not all, parasitic gaps constructions. For more complex structures, like those in (6), which do not involve an adjunct clause straightforwardly adjoining

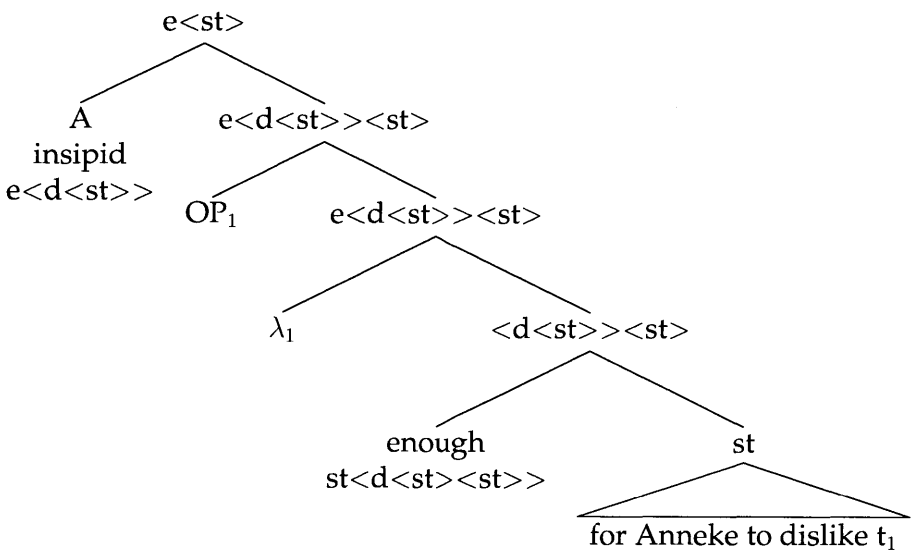
to a matrix clause, a slightly different semantic operation is required.

(6) a. this article, which$_i$ [we filed t_i] [OP$_j$ [without PRO reading t_j]]
b. his theory, which$_i$ [people who argue against t_i] [OP$_j$ [often end up adopting t_j]]
c. this theory, which$_i$ I convinced [OP$_j$ [several proponents of t_j]] to reject t_i

To capture more complex examples like (6), Nissenbaum (2000) proposes a "generalized conjunction" operation called COMPOSE[4]. COMPOSE, essentially, is an extension of predicate modification; unlike predicate modification, however, COMPOSE can apply to a more flexible range of semantic types. Where predicate modification allows the semantics to conjoin two <et> nodes, COMPOSE allows the semantics to combine any two nodes if they are the same semantic type (e.g., two nodes of type <e<d<st>>>) or if one node take the other in its scope, e.g., one node is <<**d**<**st**>><st>> and another is type <d<st>>. The latter is exactly the case we see in gapped degree phrases. Following the literature on comparative constructions and gradable adjectives (cf. Meier 2003 and the references therein), we assume that gradable adjectives are of type <e<d<st>>> and degree words, such as *too* and *enough*, are type <st<<d<st>><st>>>. The type structure for a gapped degree phrase, up to the adjective layer, is given in (7).

[4]Technically, Nissenbaum (2000) uses a generalized conjunction operation called CONJOIN. This operation is extended slightly and renamed COMPOSE in Nissenbaum & Schwarz (2011). In Nissenbaum & Schwarz (2011), the operation COMPOSE totally subsumes the operation CONJOIN: everything that CONJOIN can do, *compose* can do, too.

(7) Partial LF for gapped degree phrase



COMPOSE allows the type e<d<st>><st> and e<d<st>> nodes to combine with each other. However, after COMPOSE takes place there is still a type <e> variable that must be linked to both the embedded gap site and the gradable adjective, which also requires a type <e> argument. This composition would not be possible if the null operator moved to another position in the embedded clause, such as below the degree word or within the CP (e.g., the type <st> node in (7)). If the operator were to move to either of those position, COMPOSE would not be able to take place (because the types would not align), leading to a type clash and an inability to link the gap site and the matrix subject. For a fuller account of the specifics and technical details of COMPOSE, including the full semantic computation for a gapped degree phrase, see Nissenbaum & Schwarz (2011). Importantly, this dissertation will assume that semantically necessary movement steps are present in the syntax, as well as at LF.

This analysis potentially raises a question of phase boundaries: though the movement chain in (7) is unspecified, it potentially violates the phase impenetrability condition (PIC), at least on an analysis where the finite CP is obligatorily phasal. As discussed in Chapter 1, this dissertations assumes Wurmbrand's (2017) theory of contextually determined phasehood. Under this movement, movement to spec-DegP would not incur a PIC violation. Recall from the introduction to this dissertation that Wurmbrand's theory proposes that not all CPs are phases, and not all phases are CPs; instead, phases should be understood as "the highest head in the cyclic domain." Recall from Chapter 1 that

this assumptions allows a number of projections—FocPs, TopPs, FinPs, CPs, etc—to serve as phases, and avoids requiring a stipulation that vPs and CPs are phases because of some accident of birth. In accordance with Wurmbrand, this dissertation will argue that the DegP in gapped degree phrases is the phase boundary for the embedded clause—not the CP.

2.3 Background: Degrees, evaluators and thresholds

The embedded clauses of gapped degree phrases ((8), repeated from (7)) contain two obvious parts: an embedded CP (*for Anneke to read*) and the degree word (*too*) that licenses the CP[5]. The degree word itself is licensed by the adjective that predicates the subject.

(8) a. Jonathan Franzen is too self-absorbed for Anneke to read _.
b. Anneke is too intelligent _ to read Jonathan Franzen.

Evidence that the embedded CP is truly an argument of the degree word comes from pairs like (9), taken from Lasnik & Fiengo (1974:536) where the CP—called the *standard* in the gradable adjectives literature—cannot occur without the presence of the degree word (9a-c). In addition to the *standard*, degree words can also license an optional *evaluator* who evaluates the standard relative to their own belief worlds. An example of this is given in (10). In (10a) *Anneke* evaluates an uncomfortably thin mattress as suitable for Jonathan Franzen to sleep on, in accordance with a belief world where Jonathan Franzen deserves little material comfort. Note that, in (10b), without a degree word to license her, *Anneke* cannot serve as an evaluator.

(9) a. This mattress is {thin/thick}.
b. *This mattress is {thin/thick} to sleep on[6].
c. This mattress is {too thin/thick enough} to sleep on.

[5]Degree words can themselves appear without an overt CP as their complement., i.e., 'Jonathan Franzen is too self-absorbed.' A prediction of the analysis presented here is that such constructions contain a (perhaps contextually specified) CP argument that is semantically present but not overt, analogous to comparative deletion.

[6]While the ungrammaticality of (9b) is clear, a small class of adjectives are, in fact, grammatical in these types of constructions, such as *'These flowers are pretty to look at'*. This chapter will not discuss these kinds of constructions.

(10) a. For Anneke, this cardboard mattress is thick enough for Jonathan Franzen to sleep on.
b. *For Anneke, this cardboard mattress is thick.

The importance of evaluators, and the strength of their connections with the degree words that license them, is best understood when the semantic properties of degree words are further considered. Essentially: degree words are modal operators that position a CP *standard* in relation with a *threshold* across possible worlds. When a threshold is not determined by the limits of the natural world, *evaluators* establish whether or not a threshold is met, by contributing the belief worlds that are used to determine that outcome[7].

First, consider the modal properties of degree words like *too* and *enough*. Both *too* and *enough* can express deontic and epistemic modalities (11); they can also act as a kind of counterfactual (12).

(11) Jonathan Franzen too tall to ride the carousel.
a. Deontic: because the theme-park rules say that it's so.
b. Circumstantial: because he doesn't fit in the seats.

(12) a. Jonathan Franzen's latest novel is *too boring* to read, but Doug did, anyway.
b. Jonathan Franzen's latest novel isn't *interesting enough* to read, but Doug did, anyway

Meier (2003) claims that the facts in (11)-(12) follow from the fact that *too* and *enough* both take a covert modal as their complement. This is schematized in (13)[8], where the copula takes an AP as its complement, which itself takes a DegP as its complement. The head of the DegP, *too*, requires a covert modal as its complement, while the DegP itself takes a larger clause.

(13) a. Bertha is too young to drive.
b. Bertha is [$_{AP}$ young [$_{DegP}$ [$_{Deg}$ too CAN] to drive a car]]

[7]This description suggests that evaluators are not present in the derivation when a threshold is determined by the limits of the natural world. Another way to conceive of this fact is to propose that, in objective cases like (i), the evaluator could, in some sense, be understood as an omniscient evaluator who is aware of all the limits of the natural world.

[8]This analysis is similar to other, more recent analyses that have proposed the presence of covert modals in bi-clausal structures, such as Harves and Kayne (2012) for *have* and Kayne (2009) for *be* (i).

i. a. Bertha has to buy a car.
a'. Bertha$_i$ has [NEED t$_i$ to buy a car]
b. Bertha is to arrive at 7.
b'. Bertha$_i$ is [SUPPOSED t$_i$ to arrive by 7].

This analysis is sufficient to predict that counterfactual uses of degree words, as in (12): the covert modal selected by the degree word can be evaluated against a number of possible worlds. In the counterfactual examples, it is being evaluated against an idealized world, not the actual world.

The remainder of this dissertation, mainly for simplicity, will collapse the degree words and the modal operator they select: it will treat the degree word itself as a comparative modal operator—or, at least, a terminal head with modal properties (instead of a comparative element that embeds a modal operator). The key contribution of this section should be the understanding that degree words are able to establish whether or not a threshold is met, and this that evaluation is done through comparisons with possible worlds[9]. When an evaluator is introduced, those possible worlds are their belief worlds. When an evaluator is not overt, it is implicit and contextually determined.

Importantly, the possible worlds against which a threshold is evaluated are not identical across all degree constructions. Some thresholds, like (14a) are established by the limits of the natural world; others, like (14b) are dependent on the possible worlds against which they are being evaluated. Subjective thresholds like (14b) are *judgement calls*; when the degree word introduces an overt *evaluator* argument into the derivation, the evaluator's belief worlds are the possible worlds in which the boundaries of the threshold are established.

(14) a. Five hundred degrees fahrenheit is hot enough for a novel to catch fire. (objective threshold)
b. Ten cents is too expensive for a Jonathan Franzen novel. (subjective threshold)

When evaluators are present, we only have information about how the evaluators understands the threshold. This explains why, when an evaluator is combined with an objective threshold, the result is distinctly strange, (15).

(15) a. #For Anneke, five hundred degrees fahrenheit is hot enough for a novel to catch fire. (objective threshold)
b. For Anneke, ten cents is too expensive for a Jonathan Franzen novel. (subjective threshold)

It's worth noting that, while there is no phonologically overt evaluator in (14b), it

[9]Whether *too/enough* determine sufficiency/excess or deficiency/lack depends on whether the degree word occurs with positive or negative polar adjectives (e.g., *tall* and *old* are positive while *short* and *young* are negative). For more information, see Meier (2003).

is also clear that (14b) is not an objective fact[10]. The most obvious interpretation of (14b) is that some contextually relevant entity (most likely the author or speaker) set the threshold for the appropriate value of a Jonathan Franzen novel, relative to their belief worlds. Even though an overt evaluator is only present in (15b), an evaluator argument must also be present in (14b), or else it would be impossible to understand the belief worlds within which the threshold and standard will be compared. The meanings of (14b) and (15b) suggest that an evaluator is always present in a judgement call, even if the evaluator is not phonologically overt.

The remainder of this section looks at *where*, in a judgement call, the evaluator is introduced. This section will provide evidence that evaluators, when they are licensed by a comparative degree word like *too* or *enough*, are introduced internal to the DegP[11]. The data from this section, unsurprisingly, will draw on examples with phonologically overt evaluators. However, I will assume that silent evaluators are introduced in the same position as their overt counterparts.

One form of support for the claim that evaluators are introduced internal to the DegP comes from word-order facts in sentences like (16)-(18), where the evaluator linearly follows both the degree word and the adjective.

(16) a. CONTEXT: Chris needs to run an experiment that will only be successful if performed in warm weather. He cannot run it without his advising committee's approval. Chris and his committee do not always agree.
b. It's too cold [$_{PP}$ for his advising committee] [$_{CP}$ for Chris to run the experiment][12]

[10]If it were, it would be difficult to reconcile with the fact that, on November 2 2016, almost all Jonathan Franzen novels were priced above $0.10 on Amazon.com (though the majority of his novels were quite inexpensive).

[11]Degree words like *too* and *enough* are not the only constructions that can license evaluators, but they are the only constructions this chapter will examine. Evaluators that are not licensed by a degree word like *too* or *enough* may be introduced in other positions–for example, directly above a CP—particularly if a degree word is not present. See Section 3.1 for a discussion of which degree words can license gapped degree phrases.

[12]Some readers report an ambiguity in these constructions, leading (16) to have the following interpretations:

(i) a. In his advising committee's opinion, it's too cold for Chris to run the experiment (because they believe the low temperature will lead to his failure. (intended)
b. The members of Chris's advising committee are too cold for Chris to run the experiment (the experiment must be run in a freezer, and they have to be in that freezer during the run) (unintended).

Because the majority of readers (including myself) don't have these readings, they won't be dis-

(17) a. CONTEXT: Anneke harbors a deep-seated hatred of Jonathan Franzen. Doug believes that Jonathan Franzen is the literary voice of their generation. Unfortunately, Anneke and Doug have a group of mutual friends, who invite both of them to parties.
b. It's too awkward [$_{PP}$ for Anneke] [$_{CP}$ for Doug to be at parties]

(18) a. CONTEXT: The superintendent of a school district is a very conservative man, prone to avoiding controversy. He determines which books the entire district is (not) allowed to assign.
b. *Purity* is too controversial [$_{PP}$ for the superintendent] [$_{CP}$ for any high school to assign it next semester]

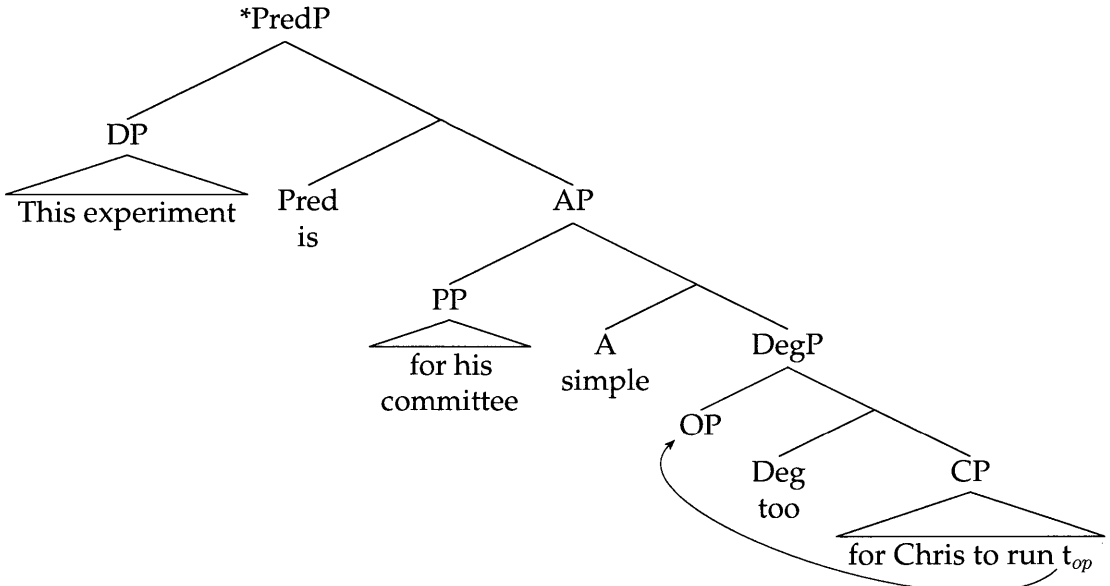
Additional evidence for introducing evaluators internal to the DegP comes from movement facts. While evaluators can occur sentence-internally in *gapless* degree phrases, they cannot occur sentence-internally in *gapped* degree phrases (19), cf. (16).

(19) *This experiment is too simple [$_{PP}$ for his committee] [$_{CP}$ for Chris to run __].

Recall from Section 2 that gapped degree phrases require an instance of null operator movement to spec-DegP, the highest specifier position of the embedded clause. Section 4 will provide additional syntactic evidence for this movement step, while Section 2 has shown that this movement step is semantically necessary. Without this movement step, there is no way to bind the embedded gap site with the matrix subject. We know that this binding is required, because there is an obligatory co-reference between the embedded clause gap site and the matrix subject. Given this, the contrast between (19) and (16) is mysterious if sentence-medial evaluators are introduced outside the DegP—say, as part of the AP. Abstracting away from word-order, a structure where the evaluator is introduced above DegP would look like (20).

cussed further in this chapter.

(20) GDP with evaluator introduced *above* DegP
This experiment is too simple [PP for his committee] [CP for Chris to run __].

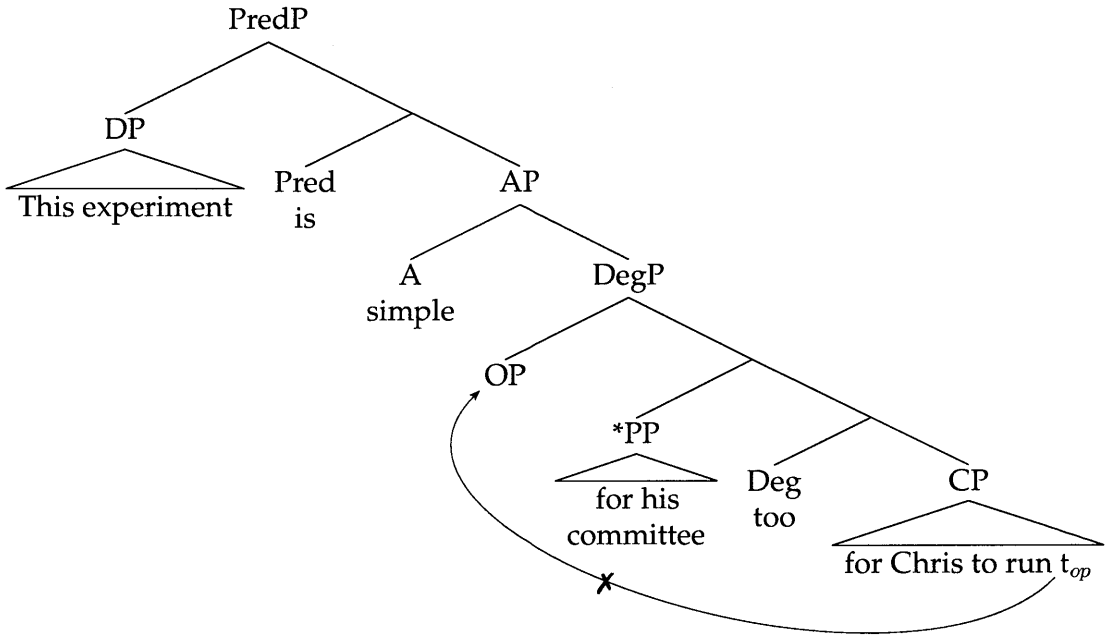


The contrast between (20) and (16) is mysterious if sentence-medial evaluators are introduced outside the DegP—say, as part of the AP. In (20), there is no reason why the movement chain would be prevented from taking place, and in fact no reason why the sentence would be ungrammatical. However, if the evaluator is introduced DegP internally, as in (21), we predict the ungrammaticality of (19). When the evaluator is introduced internal to the DegP, as in (21), it functions as an intervener to the necessary null operator movement[13][14]. The tree in (21) assumes that Deg has multiple specifiers, one in which an argument (the evaluator) can be introduced and another which can function as a landing site for movement.

[13]In (21), Deg is represented as having multiple specifiers–one hosting the evaluator and the other hosting the null operator.

[14]Careful readers who note that gapped degree phrase null operator movement is depicted as a single movement step in (20) may wonder how the presence of an evaluator argument could be an intervener for *A*-movement. As section 2.4 will discuss, object gapped degree movement should be understood as involving two steps

(21) GDP with evaluator introduced *internal to* DegP
This experiment is too simple [PP for his committee] [CP for Chris to run __].



Given these facts, the remainder of this chapter will assume that evaluators, when they co-occur with *too/enough* constructions, can be base generated internal to the DegP.

2.3.1 A note on specific degree words

This paper has so far discussed only two degree words—*too* and *enough*. For simplicity, this chapter will mainly give examples with *too*. Unless otherwise stated, the judgments given for *too* gapped degree phrases are parallel with the judgments for *enough* gapped degree phrases. With the exception of the word order and constituency differences noted in Section 2.2, gapped degree phrases formed with either degree word appear more or less identical.

While this chapter will limit itself to discussions of *too* and *enough* gapped degree phrases, it's worth noting that some other comparative degree words can license gapped degree phrases (22)—while other comparative degree words cannot (23). Developing a typology of which degree words can occur with gapped degree phrases and explaining that typology in full is an interesting problem left for future research.

(22) a. Jonathan Franzen is rather boring for Anneke to read __.
b. Jonathan Franzen is kind of boring for Anneke to read __.
c. Jonathan Franzen is awfully boring for Anneke to read __

(23) a. *Jonathan Franzen is very boring for Anneke to read __.
b. *Jonathan Franzen is so boring for Anneke to read __.
c. *Jonathan Franzen is entirely boring for Anneke to read __.

2.4 DegP-internal movement chain

This section describes the specific details of the DegP-internal movement chain previously proposed by Nissenbaum & Schwarz (2011) for gapped degree phrases, as schematized in (4). Much of the work in this section will be done by comparing gapped degree phrases to *tough*-constructions. When the structures and movement chains of gapped degree phrases and *tough*-constructions are compared, the following relationship emerges: gapped degree phrases are structurally "larger" than *tough*-constructions. While *tough*-constructions and gapped degree phrases both contain movement chains of the same *length*, the movement chain in *tough*-constructions spans the entire sentence—an element Merged in the lower clause moves to the subject position of the higher clause. Gapped degree phrases involve A-movement chain that does not reach the matrix clause—gapped degree movement, Nissenbaum & Schwarz's (2011) analysis has already shown, is entirely internal to the DegP. Thus the DegP internal to the gapped degree phrase can be thought of as containing an instance of *tough*-movement (with the degree word behaving in a way that is syntactically parallel to the role of the *tough*-predicate in *tough*-constructions). Taking this comparison seriously immediately makes predictions about the properties of Deg and its specifier. Deg must have a ϕ-probe, like T. This makes sense, given that the evaluator arguments in gapped degree phrases were previously shown to be interveners to gapped degree movement. Similarly, spec-DegP will have the properties of an A position, as least as far as ϕ-feature driven movement is concerned.

2.4.1 Ā-movement

Since Chomsky (1977), *tough*-constructions have been argued to contain an Ā step, beginning in the embedded object position and ending in the embedded

spec-CP. However, since Chomsky (1977), there has been debate about *which kind* of syntactic element undergoes that Ā-movement. In Chomsky (1977), it was a null operator. However, more recent improper movement analyses of *tough*-constructions propose that it is the object DP itself that undergoes this Ā-movement step[15]. This paper will also assume it is the embedded object DP—and not a null operator—that undergoes Ā-movement in TCs, as (partially) schematized in (24)[16].

(24) __ is tough [$_{CP}$ Jonathan Franzen C [$_{TP}$ for Anneke to read <Jonathan Franzen>.]]

Support for proposing an Ā-movement step in *tough*-constructions comes from the fact that *tough*-movement display many properties of Ā-movement constructions. Take, for example, evidence from island effects (Chomsky 1977). In TCs, the gap site can be separated from its filler by a clause boundary (25a). However, the gap site cannot be separated from its filler by an island boundary (25b).

(25) a. Jonathan Franzen is tough [$_{CP}$ for Anneke to say [$_{CP}$ that she has read __]] (clause)
b. *Jonathan Franzen is tough [$_{CP}$ for Anneke to talk about [$_{DP}$ the book written by __]] (island)

Additionally, *tough*-movement from the higher position of double object constructions (DOCs) is prohibited (26a). This restriction is found in other Ā-movement constructions, like *wh*-movement, as well (26b).

(26) a. *Anneke was tough to give *t* this book. (DOC *tough*-extraction)
b. *Who did you give *t* this book? (DOC *wh* movement)

Chomsky (1977:104) additionally states that it is particularly "natural" to claim

[15]Hybrid approaches—such as Hicks (2009), where a null operator containing a phonologically overt DP is moved from the embedded object position to the edge of the embedded clause, after which the overt DP is sub-extracted and moved into the matrix clause—have also been proposed to explain the unique properties of *tough*-constructions. See the introduction of this dissertation for a review of Hicks' argument, which is not adopted here.

[16]For clarity and consistency, this dissertation will assume an overt movement analysis of *tough*-constructions throughout. This approach is, I believe, the most thoroughly defended in the literature and requires the fewest number of outside assumptions. Additionally, this approach is best suited to explain the split-antecedent and partial control facts presented later in this chapter. That being said, the core proposal of this chapter—that gapped degree phrases are structurally larger than TCs, and contain an instance of *tough*-like-movement within their embedded clauses—is on no way dependent on any particularly analysis of *tough*-constructions.

that *tough* constructions involve Ā-movement, as *tough*-constructions have "analogous forms in which [a] *wh*-phrase may directly appear" (27).

(27) a. a tough chair *on which* to sit
b. an easy violin *on which* to play sonatas

Tough-constructions, like *wh*-constructions, also license parasitic gaps (28). This section has previously discussed Nissenbaum (2000), an analysis of parasitic gaps that does not depending on a previous instance of Ā-movement (though see Engdahl 1983 for a theory where parasitic gaps are dependent on a previous instance of Ā-movement). However, there still appears to be some connection between Ā-movement and parasitic gaps, as shown in (29), a classic example from Engdahl, where passive (A) movement does not license parasitic gaps, but *wh*-movement (Ā) does:

(28) a. That candidate was easy to hire *t* [without interviewing *pg*]. (*tough* PG)
b. Which candidate did you hire *t* [without interviewing *pg*]? (*wh* PG)

(29) a. *The candidate was hired (by Adam) _ without interviewing _. (passive movement)
b. *Which candidate did Adam hire _ without interviewing _? (*wh*-movement)

Given the data in (29), it stands to assert that there is still some connection between Ā-movement and licensing parasitic gaps. If the fact that *tough*-constructions license parasitic gaps cannot be taken as evidence that they are Ā-movement constructions, it is at least consistent with an Ā-movement analysis.

Final support for the claim that *tough*-constructions involve an Ā-movement step comes from the fact that *tough*-extraction can occur over the subject of the embedded CP without causing a relativized minimality violation. Because the embedded object must move over another argument—the embedded subject (either overt or PRO)—relativized minimality requires that this step *must* be Ā-movement. A-movement would result in a relativized minimality violation, with the embedded subject functioning as a defective intervener.

(30) a. __ is tough [$_{CP}$ Jonathan Franzen C [$_{TP}$ for Anneke to listen to <Jonathan Franzen>.]] (Ā)

b. *__ is tough [$_{CP}$ Jonathan Franzen C [$_{TP}$ for Anneke to listen to <Jonathan Franzen>.]] (A)

The structure given in (30) presents *Anneke* as the subject of the embedded clause, following Hartman (2012), but contra Longenbaugh (2016), who argues that *for Anneke* is a higher PP introduced above the embedded clause. This chapter will continue to use Hartman's structure and analysis for *tough*-constructions—see Section 4.2 for a more details discussion of Hartman's analysis.

Object gapped degree phrases behave identically to *tough*-constructions with respect to Ā-movement tests. gapped degree phrases license parasitic gaps (31) and disallow extraction of the indirect object in DOCs (32).

(31) Jonathan Franzen's novels are inane enough for Anneke to burn *t* after reading *pg*]

(32) *Anneke is too smart to give __ this book. (DOC extraction)

Additionally, gapped degree phrases do not permit their null operator to move across an island boundary (33a). However, the gapped degree phrase null operator can be separated from its gap site by a clause boundary, as in (33b).

(33) a. *Jonathan Franzen is too arrogant [$_{CP}$ for Anneke to talk about [$_{DP}$ the book written by __]]. (island)

b. Jonathan Franzen is too arrogant [$_{CP}$ for Anneke to say [$_{CP}$ that she has talked to __]]. (clause)

Gapped degree phrases also have a corresponding overt *wh*-construction (35)[17].

(34) a. She thinks 35 miles/week is *too small a base* ***on which*** to run a marathon.

b. I don't believe this is *a big enough hook* ***on which*** to hang a coat.

Finally, gapped degree phrases also allow movement of the extracted object past the subject of the embedded CP. As briefly discussed in Section 1, this fact suggests that the movement step from the embedded object position to embedded spec-CP is not ϕ-driven movement. Generally, Ā-movement is *not* thought to be

[17]I only nominally have these constructions in my grammar; these examples are results from the first page of Google searches for "too * a * on which to" and "a * enough * on which to," which yielded 274,000,000 and 255,000,000 search results, respectively. Other native English speakers who have this construction in their grammars agree with these judgments.

ϕ-driven. Because the ϕ-features of the embedded subject do not block movement of the embedded object, we expect that the object Ā-moves at least as high as the embedded CP (35). Recall that this is *null operator movement*, which occurs internal to the DegP.

(35) a. ... [$_{AP}$ shy [$_{DegP}$ too [$_{CP}$ OP C [$_{TP}$ for Anneke to talk to t_{op}.]]] (Ā)

b. *... [$_{AP}$ shy [$_{DegP}$ too [$_{CP}$ OP C [$_{TP}$ for Anneke to talk to <t_{op} >.]]] (A)

This chapter makes the following assumptions about movement: An extracted argument may only cross a phase boundary if it first moves to the specifier of that phase. Both CP and vP function as phases. Ā-movement is successive-cyclic, and Ā extracted arguments move through spec-vP on their way to spec-CP (Fox 1999, Rackowski and Richards 2005, van Urk and Richards 2013). Given these assumptions, the Ā step of TCs and gapped degree phrases can be represented as follows (with traces used in place of copies, for space):

(36) *tough*-constructions
__ is tough [$_{CP}$ Jonathan Franzen C [$_{TP}$ for Anneke to [$_{vP}$ t_I read t_I.]]]

(37) Gapped degree phrase
Jonathan Franzen is [$_{AP}$ insipid [$_{DegP}$ too [$_{CP}$ OP C [$_{TP}$ for Anneke to [$_{vP}$ t_{op} read t_{op}.]]]

Split-antecedent tests support distinguishing the Ā-movement chains of TCs and gapped degree phrases as in (36)-(37), at least in terms of *what* kind of element undergoes Ā-movement in each construction. The example in (38), taken from Brillman (2015), show that gapped degree phrases can occur with split antecedents (38a) while TCs cannot (38b).

(38) a. Katherine is friendly enough and Gretchen is compassionate enough (for me) to introduce __ to each other. (GDP)

b. *Katherine is easy and Gretchen is tough (for me) to introduce __ to each other. (TC)

This follows naturally from the analysis proposed here, where a null operator undergoes Ā-movement in gapped degree phrases and an overt DP undergoes Ā-movement in *tough*-constructions. In gapped degree phrases, the null operator moves within the phasal DegP, and then links up to its split antecedent, base generated in the matrix clause via a semantic operation called COMPOSE

(Nissenbaum & Schwarz 2011). In *tough*-constructions, the split antecedent would need to be moved from the embedded clause to the matrix clause. This kind of movement is predicted to cause a binding theory violation. This account explains why (38b) is ungrammatical, unlike (38a). Thus, the ungrammaticality of the *tough*-construction (38b) is expected on an analysis where *tough*-constructions contain a single movement chain, from the embedded object position to the matrix subject position.

2.4.2 A-movement

This section adopts and defends an improper movement analysis of *tough*-constructions and gapped degree phrases. To do this, this section shows that the derivations of both TCs and gapped degree phrases contain an A-movement step, which crucially must occur after the Ā-movement step discussed in Section 3. The improper *tough*-movement structure is schematized in (39), repeated from **??**, cf. (36). Evidence for this movement chain in *tough*-constructions comes from Hartman (2009, 2012), which shows that (a) English is a language sensitive to certain types of relativized minimality effects, where an intervening argument can block A-movement and (b) these relativized minimality effects are seen in *tough*-constructions, where the presence of an argument above the CP can indeed block *tough*-movement. Arguments for an improper movement chain in gapped degree phrases come from applying Hartman's tests to gapped degree phrases. These tests produce results that mirror the *tough*-construction data[18].

(39) Tough (improper) movement[19]:
Jonathan Franzen is tough [$_{CP}$ t_I for Anneke to talk to t_I]

A Key:
——— = A-bar
..... = A

The (improper) movement chain for object gapped degree phrases is illustrated in (40) cf. (37). Unlike *tough*-constructions, however, this movement chain does not extend into the matrix clause. In gapped degree phrases, the entire improper movement phrase is internal to the embedded clause. Recall, here, that the embedded clause for a GDP is the DegP that licenses the CP and contains the full operator movement chain.

[18]For analyses of *tough*-constructions that do not assume an A-movement step, see Chomsky (1977, 1981), Fleisher (2013) and the references therein.

[19]This example excludes the necessary successive-cyclic Ā-movement step from the embedded object position to spec-*v*P for maximal clarity.

(40) Gapped degree phrase movement:
Jonathan Franzen is [$_{AP}$ shy [$_{DegP}$ OP too [$_{CP}$ t_{op} C [$_{TP}$ for Anneke to [$_{vP}$ t_{op} talk to t_{op}.]]]

English is a defective intervention language

Hartman (2009, 2012) provides support for an improper movement analysis of *tough*-constructions by showing that *tough*-constructions are sensitive to defective intervention effects when an oblique experiencer is present. Oblique experiencers are oblique arguments optionally licensed by some predicates, including *tough* adjectives and ECM verbs, as illustrated in (41).

(41) Oblique experiencers
a. It was hard [$_{PP}$ on Mary] [$_{CP}$ for her boyfriend to give up sugar] (TC)
b. The prosecutor proved [$_{PP}$ to the jury] [$_{CP}$ that the defendant was guilty] (ECM)
c. It was proven [$_{PP}$ to the jury] [$_{CP}$ that the defendant was guilty]

The argument status of these experiencers is shown by *wh*-extraction tests. Because the DP can be grammatically *wh*-extracted from the experiencer PP, we predict these oblique PPs to be arguments, not adjuncts. This is a crucial fact, given Hartman's proposal that only arguments will block a lower argument from A-movement above them in defective intervention languages. If this constituents were not arguments, we would not expect them to allow their DPs to be *wh*-moved.

(42) a. Who was it hard [$_{PP}$ on t_{wh}] [$_{CP}$ for Mary to give up sugar]?
b. Who did the prosecutor prove [$_{PP}$ to t_{wh}] [$_{CP}$ that the defendant was guilty]?
c. Who was it proven [$_{PP}$ to t_{wh}] [$_{CP}$ that the defendant was guilty]?

In both active and passive ECM constructions, oblique experiencer block A-movement, (43)-(44). This contrast can be reduced to a relativized minimality violation. Since an argument cannot A-move over another argument[20], ECM constructions containing both an oblique experiencer an argument that must

[20]There are two exceptions to this rule, perhaps English specific. The English verbs *seems* and *appears* license A-movement (raising) across an intervening experiencer (i). These verbs are anomalous in this regard; to the best of my knowledge, there is no account about why they do not show the same defective intervention effects found in other constructions, like ECM verbs.

(i) a. Jonathan Franzen seems **to Anneke** [t to have stolen the coffee]

A-move across the experiencer are predicted to be ungrammatical. This prediction is borne out in (43a) & (44a), where A-movement across an experiencer is ungrammatical. When the oblique experiencer is not present, A-movement of a lower argument is predicted to be licit. This prediction is borne out in (43b) & (44b).

(43) Active ECM verb

a. *The prosecutor proved the defendant [$_{PP}$ **to the jury**] [$_{TP}$ t to be guilty]

b. The prosecutor proved the defendant [$_{TP}$ t to be guilty]

(44) Passive ECM verb

a. *The defendant was proven [$_{PP}$ **to the jury**] [$_{TP}$ t to be guilty]

b. The defendant was proven [$_{TP}$ t to be guilty]

Hartman shows that these defective intervention effects can be obviated by moving the experiencer to a non-intervention position elsewhere in the clause (45)[21]. In both (45a) and (45b), I argue, again following Hartman and previous analyses of ECM, that the defendant has moved to the higher clause for case reasons. The subject position of the tenseless infinitive is not a case position, which means that leaving *the defendant* in-situ would result in a structure where *the defendant* could not receive case. This structure is predicted to be ungrammatical, while a structure where *the defendant* moves to the higher clause to receive case is predicted to be grammatical.

b. Jonathan Franzen appears **to Anneke** [t to have written several terrible novels]

[21]This dissertation does not make a claim on why this word order fact is true, or how exactly it relates to relativized minimality.

(45) a. [$_{PP}$ **To the jury**], the prosecutor proved the defendant [$_{TP}$ t to be guilty].
b. [$_{PP}$ **To the jury**], the defendant was proven [$_{TP}$ t to be guilty]

Tough-movement and defective intervention

The remainder of this chapter depends on an analysis of *tough*-construction where two kinds of prepositional arguments are possible: as in (45a), it should be possible to adjoin an oblique experiencer PP to the structure and for the embedded clause to be able to house a *for*-infinitival subject. Within the literature, there is some debate surrounding whether this is an appropriate structural analysis of *tough*-predicates. For example, Longenbaugh (2016) claims that *tough*-constructions embed a complement too small to contain a *for*-infinitival subject. Under Longenbaugh's analysis, any preposition-DP pair present in a *tough*-construction would need to be an oblique argument outside of the embedded clause. Nevertheless, this chapter will adopt the proposals put forth in Hartman (2012) which provide a variety of arguments (both from the syntax and semantics) in support of the structure depicted in (45) (though see Keine & Poole 2017 for an alternative approach).

Crucially, Hartman (and parallel accounts) shows that *tough*-constructions are also sensitive to RM effects. As seen in Section 4.2.1, *tough*-predicates can optionally license an oblique PP experiencer. When the oblique experiencer is not present, *tough*-movement is grammatical (46b). As with ECM active and passive constructions, *tough*-moving an argument from the embedded CP over a PP experiencer argument results in ungrammaticality (46c).

(46) a. It is enjoyable ([PP to me]) [$_{CP}$ for John to eat strawberries.]
b. Strawberries are enjoyable [$_{CP}$ for John to eat t]
c. *Strawberries are enjoyable [$_{PP}$ to me] [$_{CP}$ for John to eat t]

Similarly, moving the PP experiencer to a non-intervention position in the clause is enough to avoid the relativized minimality violation and license *tough*-movement again (47), from Chomsky (1977).

(47) a. It is easy [$_{PP}$ for the rich] [$_{CP}$ for the poor to do the work]
b. *The work is easy [$_{PP}$ for the rich] [$_{CP}$ for the poor to do __].
c. The work is easy [$_{CP}$ for the poor to do __].
d. [$_{PP}$ For the rich], the work is easy [$_{CP}$ for the poor to do __].

These facts suggest that the final step in the *tough*-movement chain is an A-movement step. However, this analysis hinges on the assumption that this argument is the subject of the embedded CP, *not* an experiencer licensed by the *tough*-predicate. Hartman (2009, 2012) has several arguments for this. The strongest comes from *tough* predicates that license oblique experiencer with a preposition other than *for*, as in (48).

(48) a. It was hard [$_{PP}$ **on** Mary] [$_{CP}$ for her boyfriend to give up sugar]
b. It is enjoyable [$_{PP}$ **to** John] [$_{CP}$ for his granddaughter to eat strawberries]

When the examples in (48) involve *tough*-movement, only the *for* PPs (the embedded subjects) are grammatical (49)-(50). This suggests that it is the oblique experiencer, not the embedded subject, that is absent from these constructions.

(49) a. Sugar was hard for Mary to give up.
b. *Sugar was hard on Mary to give up.

(50) a. Strawberries are enjoyable for John to eat.
b. *Strawberries are enjoyable to John to eat.

Additional evidence comes from partial control. In (51a), the oblique experiencer partially controls the embedded PRO. However (51b) lacks a partial control reading (and is ungrammatical). This is easily explained if (51b) lacks an oblique experiencer and a PRO to partially control. Thus, partial control facts suggest that *tough*-constructions lack an oblique experiencer and contain an overt (non-PRO) embedded subject.

(51) a. It's tough [$_{PP}$ for Mary] [$_{CP}$ PRO to meet on the bridge]
b. *The bridge is tough [$_{CP}$ for Mary to meet on __]

Gapped degree phrases and defective intervention

Gapped degree phrases have a parallel distribution with regard to defective intervention tests. Recall that gapped degree phrases, if they are judgement calls, license an oblique evaluator argument above the degree word in spec-DegP. This argument can be either phonologically overt or silent (and contextually specified). In either case, this argument contributes the belief worlds which allow the CP standard to be evaluated against the threshold put forward by the degree word itself. The example in (52) shows that, like with *tough*- and ECM constructions, a

phonologically overt argument (in this case the evaluator) is only allowed when there is no gapped degree movement. Put another way, (52) shows that overt evaluators cannot co-occur with null operator movement internal to the DegP embedded clause.

(52) a. It is too cold [$_{PP}$ for Bob] [$_{CP}$ for Chris to run this experiment]
b. *This experiment is too complex [$_{PP}$ for Bob] [$_{CP}$ for Chris to run __].

This is a crucial contrast with gapless degree phrases, which do not contain an instance of null operator movement and allow for an overt evaluator.

(53) This experiment is too complex [$_{PP}$ for Chris] [$_{PP}$ for Mary to use it as one of her exam questions].

This contrast is not predicted if the null operator undergoes Ā-movement alone. If that were the case, the null operator would be able to cross over the evaluator, as it can cross over the subject of the embedded CP. However, this contrast is predicted if the null operator undergoes a final A-movement step, parallel to the final movement step in *tough*-constructions. If there is no final A-movement step, it is difficult to explain the grammaticality of (53b).

The gapped degree phrase data parallels the *tough*-construction data in two key ways. First, when the phonologically overt evaluator is not present, null operator movement can occur, as in (54). Second, when the phonologically overt evaluator is moved to a non-intervention position, null operator movement is licensed again (55).

(54) The experiment is too simple [$_{CP}$ for Chris to run __].

(55) [$_{PP}$ For his advisor], this experiment is too simple [$_{CP}$ for Chris to run __].

As with the *tough*-construction data, this analysis hinges on *for Chris* in (54) being an embedded subject, not an oblique evaluator. Again, evidence for this comes from partial control. A partial control reading is only possible when there is no DegP internal movement, such as the gapless degree phrase in (56a). In (56a), the evaluator *for Mary* partially controls the embedded subject PRO. However, (56b), which involves null operator movement, does not have a partial control reading (and is so ungrammatical). This is only explained if *for Mary* is the embedded subject, and there is no evaluator licensed.

(56) a. It's too cold [$_{PP}$ for Mary] [$_{CP}$ PRO to meet on the bridge]

b. *The bridge is too cold [$_{CP}$ for Mary to meet on __]

Parallel to the data from Hartman, we see that (56b) is grammatical if the evaluator is fronted to the left of the matrix clause, as in (57).[22]

(57) For Mary, this bridge is too cold to meet on _.

This supports an improper movement analysis of gapped degree phrases, where the final movement step, from spec-CP to spec-DegP, is an A-movement step. This is schematized in (58), repeated from (40).

(58) Gapped degree phrase movement:
Jonathan Franzen is [$_{AP}$ shy [$_{DegP}$ OP too [$_{CP}$ t_{op} C [$_{TP}$ for Anneke to [$_{vP}$ t_{op} talk to t_{op}.]]]

2.5 Subject gapped degree phrases

So far, this chapter has focused primarily on object gapped degree phrases. This has been primarily for clarity: *tough*-constructions do not license subject extraction, and the structural and behavioral parallels are more immediately obvious when *tough*-constructions and object gapped degree phrases are directly compared.

However, this fact is, itself, something of a mystery: if gapped degree phrases and *tough*-constructions are so similar, why do gapped degree phrases license subject extraction while *tough*-constructions do not? The examples in (59) show that *tough*-constructions with subject extraction are obviously ungrammatical (59a), while subject gapped degree phrases are perfectly acceptable (59b).

[22] An interesting question both here and in Hartman's account is why only phonologically overt arguments serve as interveners to movement. In gapped degree phrases, only overt evaluators can serve as interveners to movement–neither phonologically null interveners nor silent traces/copies of overt evaluators can block gapped degree movement. Similarly, Hartman's data shows that only overt oblique PP arguments–not traces or silent copies of these arguments–block *tough*-movement or the formation of ECM constructions. This connection between an arguments phonological overtness and its ability to serve as an intervener is left for further research.

(59) a. *Anneke is tough __ to read Jonathan Franzen[23] (*tough*-construction)
b. Anneke is too smart __ to read Jonathan Franzen (subject gapped degree phrase)

This is particularly mysterious because this chapter has proposed that there are only three significant differences between gapped degree phrases and *tough*-constructions: (i) gapped degree phrases have a structurally larger embedded clause, compared with *tough*-constructions; (ii) gapped degree phrases involve null operator movement, while *tough*-constructions involve the movement of a phonologically overt DP; and (iii) gapped degree phrase movement is entirely internal to the embedded clause, while *tough*-movement spans both the matrix and embedded clauses.

An ideal analysis of gapped degree phrases, particularly one that compares them to *tough*-constructions, would be able to predict why gapped degree phrases allow subject extraction while *tough*-constructions prohibits it. However, the current analysis is not sufficient to do this: currently, there is nothing in any of the three differences between gapped degree phrases and *tough*-constructions that would predict this asymmetry.

Chapter 3 will introduce a new puzzle piece, a constraint called *spec-to-spec anti-locality* that will allow us to bridge this gap. Essentially, the addition of spec-to-spec anti-locality will directly link the fact that gapped degree phrases license subject extraction to the fact that they have a larger embedded clause structure, compared to *tough*-constructions. Similarly, the relatively smaller size of the embedded clauses in *tough*-constructions will be enough to explain why *tough*-constructions cannot license subject extraction.

2.6 Conclusions

This chapter has shown that the seemingly mysterious properties of (object) gapped degree phrases can be easily understood given an analysis that claims that (object) gapped degree phrases have a *larger than usual* embedded clause structure. Particularly, in gapped degree phrases, the embedded clause is headed by a DegP layer, directly above CP. The presence of this DegP layer

[23]This example, (59a), has a marginally grammatical reading where the act of reading a Jonathan Franzen itself would make Anneke tough. This reading is clearly not a *tough*-construction interpretation, and, though this example's namesake would undoubtedly agree with the sentiment, this chapter will disregard these marginal interpretations.

straightforwardly explains the two most puzzling properties of gapped degree phrases: (i) their similarity to *tough*-constructions and (ii) their semantic composition, which appears parasitic gap-like in its requirement that the matrix argument be bound to the embedded gap site. Syntactically, this larger structure allows for the proposal that gapped degree phrases contain an instance of (something like) *tough*-movement entirely within their embedded clauses. Semantically, this larger structure sets the stage for the semantic operation COMPOSE, which cannot operate over CPs, to occur, binding the matrix argument and embedded gap site. This chapter positions gapped degree phrases as a first example of what can be gained from a theory of syntax that allows for a variety of embedded clause structures and sizes.

Chapter 3

Infinitives and anti-locality

3.1 Introduction

This chapter introduces the notion of *specifier-to-specifier* (spec-to-spec) anti-locality. Broadly, spec-to-spec anti-locality is a constraint that limits Ā-movement operations that are "too short" (see Section 3 for a formal definition of this term). This chapter will review two distinct but closely related spec-to-spec anti-locality proposals: one from Erlewine (2016), the other from Brillman and Hirsh (*to appear*). Both theories of spec-to-spec anti-locality restrict short subject Ā-movement from spec-TP to an adjacent spec-CP position — under both theories, this movement operation meets the definition of "too close". This chapter will explore the interaction of spec-to-spec anti-locality and subject movement more generally. A core goal of this chapter is to show how spec-to-spec anti-locality has the ability to unify a wide range of subject/non-subject asymmetries. Importantly, spec-to-spec anti-locality does not make an explicit reference to subjecthood. Instead, spec-to-spec anti-locality constraints A-movement configuration that subjects—the highest arguments in the derivation—are disproportionately likely to appear in: an Ā-movement chain with one link in spec-TP and another in an adjacent spec-CP. As a result, subjects are the arguments most likely to operation within the scope of spec-to-spec anti-locality. As such, subjects also function as the clearest demonstrations of how spec-to-spec anti-locality operates within the grammar.

This chapter will be organized as follows: Section 2 will discuss the types of subject/non-subject asymmetries that spec-to-spec anti-locality can and cannot capture. Section 3 discusses the exact definition of both Erlewine's and Brillman

and Hirsh's spec-to-spec anti-locality constraints. This section also discusses previous anti-locality constraints, and lays out the theory of syntax that both Erlewine and Brillman and Hirsh's spec-to-spec anti-locality constraints require. Section 4 discusses how spec-to-spec anti-locality compares to a competing theory that has been proposed to unify subject/non-subject asymmetries, Criterial Freezing (Rizzi & Shlonsky 2007). This section shows that both theories are sufficiently powerful to capture a range of cross-linguistic data. Section 4 will also continue to illustrate the types of configurations where anti-locality is predicted to be able to restrict subject movement. Section 5 moves to English, and discusses a number of English subject/non-subject asymmetries that are predicted by anti-locality.

Much of this chapter, particularly sections 3 and 5, represent joint work with Aron Hirsch, and that much of the analysis included in this section has been adopted from or inspired Brillman & Hirsch (2016).

3.2 On the (noted) weirdness of subject extraction

In many languages subject extraction appears to be special, compared to the extractions of objects and other arguments. Cross-linguistically, subject/non-subject extraction asymmetries can manifest in a variety of ways, including: complementizer realization (see Koopman and Sportiche 2007 for French); (anti-)agreement morphology, (see Diercks 2010 for Lubukusu, Ouhalla Ouhalla 1993 for Berber); and more restricted extraction possibilities for subjects, (see Cole Cole 1985 for Quechua, Borer Borer 1984 for Hebrew). For a more extensive cross-linguistic overview of noted subject/non-asymmetries, see Rizzi and Shlonsky (2007). Subject/non-subject asymmetries are also common in English. For example, subjects cannot undergo *tough*-movement in English, (1a), while object can (1b)—an asymmetry that was first discussed in Chapter 2 of this dissertation.

(1) *Tough*-constructions:
   a. *Anneke is tough _ to see Ian. (subject gap)
   b. Ian is tough for Anneke to see _. (object gap)

Of course, *tough*-constructions are only one example of English subject/non-subject restrictions. The *that*-trace effect is another. Here subjects cannot Ā-extract from an embedded clause introduced by the complementizer

that, while non-subjects can (2a-b). Even more puzzling, this effect appears connected to the presence or absence of an overt complementizer *that* in the derivation. When no overt complementizer *that* is present, subjects are free to extract from an embedded clause (2c).

(2) *That*-trace effects:

a.	*Who does Bill think that _ saw John?	(subject extracted with *that*)
b.	Who does Bill think that John saw _?	(object extracted with *that*)
c.	Who does Bill think _ saw John?	(subject extracted without *that*)
d.	Who does Bill think John saw _?	(object extracted without *that*)

Similarly, there are a number of asymmetries between subject and non-subject matrix *wh*-questions. One example is given in (4): subject questions do not allow *do*-support, while non-subject questions require it (Koopman 1983).

(3) Matrix *wh*-questions and *do*-support:

a.	Who talked/*did talk to Bill?	(subject question)
b.	Who did John talk to _?	(object question)

These English examples show that subjects can Ā-extract differently than non-subjects. However, beyond that basic pattern, it's not clear how to further unify this suite of behavior—it's not obvious how to relate the absence of *do*-support to a restriction on *tough*-moving subjects. This chapter proposes that spec-to-spec anti-locality offers a united explanation for the apparently special behavior of subjects across a wide variety of apparently unrelated Ā-constructions. This chapter proposes that the exceptional status of subjects in each of the English examples in (1)-(3) (as well as the remaining examples that this chapter will discuss) can be traced to the same source: a restriction that prohibits Ā-movement whose origin and landing site are "too close" together in the syntax. A formal definition of "too close" will be given in Section 3, and this definition will include movement chains with one link in spec-TP and another in a neighboring spec-CP. The types of subject/non-subject examples that this chapter will discuss are those with, on first investigation, appear to require their subjects to move directly from spec-TP to a neighboring spec-CP.

This restriction–limiting subject movement from spec-TP to an adjacent spec-CP–immediately creates a tension with the data: subject extraction is obviously possible in many circumstances (see (2)), the most obvious of which are probably matrix subject *wh* questions. The goal of this chapter is not to propose a

constraint that restricts all subject movement, but rather to discuss a constraint that restricts certain instances of subject extraction. An auxiliary goal of this section is to propose that subject movement, contrary to appearances, never involves Ā-movement from spec-TP to an adjacent spec-CP position. Spec-to-spec anti-locality bans subject movement from spec-TP to spec-CP, but it does not ban subject movement. One key fact of anti-locality is that it allows longer subject movement chains—from spec-TP to spec-CP *across an intervening projection*, or from spec-vP to spec-CP directly—to occur unrestricted. Section 3 spells out these conditions formally, and describes the full range of strategies that a language can access in order to allow subject Ā extraction while still respecting anti-locality.

Spec-to-spec anti-locality is not the first constraint that has been proposed to capture a wide range of subject/non-subject asymmetries. That most obvious competitor to spec-to-spec anti-locality is Criterial Freezing, first proposed in Rizzi & Shlonsky (2007). Given this, portions of this chapter will directly compare the predictions of Criterial Freezing and spec-to-spec anti-locality, with the hope of showing that anti-locality has the more robust predictive power.

3.2.1 Subjects aren't actually special

One of the most baffling facts about subject/non-subject asymmetries is that subjects only *sometimes* seem special. The *that*-trace effect (2) is perhaps the clearest example of this: the subject seems special when it occurs in an embedded clause that contains an overt *that* and it cannot extract, but the subject does not seem special when it occurs in an embedded clause which lacks and overt *that* and normal extraction is permitted. A theory of subject/non-subject asymmetries that refers to subjecthood directly will be hard-pressed to capture alternations like the *that*-trace effect, where external factors (in this case, the presence or absence of an overt complementizer) determine whether a subject behaves like a non-subject arguments or exceptionally.

Spec-to-spec anti-locality does not make any references to subjecthood. It does not, for instance, require a designated SubjectP or [+SUBJECT] feature (though it is also not inconsistent with such proposals). Rather, spec-to-spec anti-locality limits the Ā-movement paths available to all arguments. As previously mentioned, one movement path that spec-to-spec anti-locality restricts is movement from spec-TP to an adjacent spec-CP. This is a syntactic configuration that subjects are are disproportionately likely to encounter, because the EPP requires subject to move

to spec-TP and theories of successive cyclic Ā-movement and phasehood require all arguments to move through spec-CP. As a result, spec-to-spec anti-locality, while not designed to restrict the movement paths of subjects specifically, in practice applies most commonly to subject movement operations. In essence, when a subject is unable to participate in A-movement operation because of an anti-local syntactic configuration, they will behave distinctly from other (lower) objects. When there is no anti-locality configuration blocking subject movement, subjects will behave (more or less) like other arguments. In this way, anti-locality is able to capture the occasional specialness of subjects while simultaneously leaving the extraction possibilities of other arguments generally unaffected.

This also suggests which types of subject/non-subject asymmetries spec-to-spec anti-locality should be able to capture, and which languages should display anti-locality effects. Movement from spec-vP to spec-CP is not anti-local. As such, anti-locality only predicts subject/non-subject asymmetries in languages with a strong EPP that requires a full argument to pass through spec-TP. Similarly, since spec-to-spec anti-locality is constraint on Ā-movement chains, it is not predicted to explain subject/non-subject asymmetries that do not involve Ā-movement. For example, spec-to-spec anti-locality is not predicted to explain why object raising is banned while subject raising is allowed. Similarly, very local A-movement operations, such as short subject movement from spec-vP to spec-TP are not ruled out by anti-locality, suggests that anti-locality is only an Ā-movement constraint. Additionally, because of the interaction of the EPP and theories of phasehood, anti-locality is only predicted to account for subject/non-subject asymmetries that involve movement through both TP and CP at some step of the derivation. Finally, anti-locality is only predicted to be able to account for subject/non-subject asymmetries where *subject* movement is restricted; it is not necessarily predicted to be able to be able to explain asymmetries where subject are able to Ā extract more freely than objects. The most well-known subject/non-subject asymmetry that does not fit into this paradigm is the French *qui/que* effect, discussed in depth in Section 4.5. In the *que/qui* alternation, objects are restricted from extracting from embedded clauses introduced by *qui*. In order to extract from *qui* clauses, objects would not be required to undergo any Ā-movement step that is "too short," and so this effect at least partially falls outside the scope of anti-locality. This is not to claim that anti-locality has no place in any *que/qui* analysis (Section 4.5 proposes a potential solution to the *que/qui* effect that relies on the interaction of anti-locality and a

certain theory of phasehood) or any analysis of a similar phenomenon, but rather to explain that some subject/non-subject asymmetries can either not be accounted for by anti-locality, or can only be accounted for by the interactions of anti-locality and other facets of the grammar.

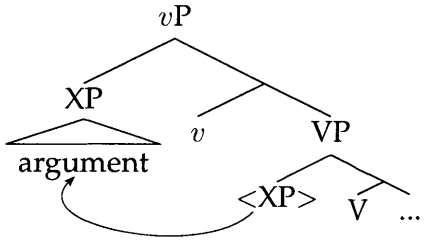
3.3 Anti-locality constraints

This section discusses previous theories of anti-locality, paying particular attention to Erlewine's (2016) and Brillman & Hirsh's (2016) theories of spec-to-spec anti-locality. In addition to discussing previous anti-locality proposals, this section will also outline the kinds of subject/non-subject alternations that spec-to-spec anti-locality predicts. This section will spend significant time comparing the proposals in Erlewine and Brillman & Hirsh. Though these proposals are very similar, they interact differently with distinct aspects of the grammar, and choosing one theory or another requires adopting slightly different assumptions about syntax more generally. In discussing Erlewine's and Brillman & Hirsh's spec-to-spec anti-locality proposals, this section will be able to formalize what have previous discussed as abstract notions of distance. Additionally, once anti-locality is fully defined and described, this section will also discuss its predictions, and the types of subject/non-subject asymmetries that we can expect anti-locality to explain.

3.3.1 Anti-locality: an abbreviated history

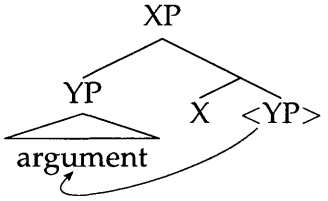
Anti-locality—the general concept that movement operations can be *too short* as well as *too long*—is not a concept novel to this dissertation. General theories of anti-locality can be traced back to (at least) Grohmann (2000), who proposed a version of anti-locality dependent on a tripartition of the clause into three "Prolific Domains": one for thematic relations, one for agreement, and one for discourse relations (roughly analogous to vP, TP and CP, respectively). Grohmann argued that, while movement *across* domains (e.g., from the thematic domain into the agreement domain) was possible, movement *within* a single domain was prohibited. Grohmann proposed that a PF constraint restricted anti-local movement. Thus, the movement chain outlined in (4), which depicts movement from spec-VP to spec-vP, is impossible under Grohmann's theory.

(4) Grohmann's anti-locality constraint (updated from 2000:41 ex 10)



Like Grohmann, Abels (2003) also proposed a constraint that limited certain short movement steps. Abels was less concerned with marking boundaries within the clause, and instead proposed that anti-locality was something integral to the definition of and motivation for movement itself. Abel's anti-locality ruled out very short movement from the complement position of an XP to spec-XP, as illustrated in (5). Abels proposed that movement to the specifier of a projection was a method of establishing a relationship with the head of that projection, and that complements were already in a similar relationship with their heads. Thus, Abel's proposal sought to eliminate a redundant movement operation from the grammar.

(5) Abels's anti-locality constraint (2003:12 ex4)



Spec-to-spec anti-locality is somewhat different from Grohmann's and Abels's theories of anti-locality. This chapter will first start with a discussion of Erlewine's (2016) definition of spec-to-spec anti-locality in Section 3.2, before moving to a discussion of Brillman & Hirsh's (2016) definition in Section 3.3.

3.3.2 Erlewine's spec-to-spec anti-locality

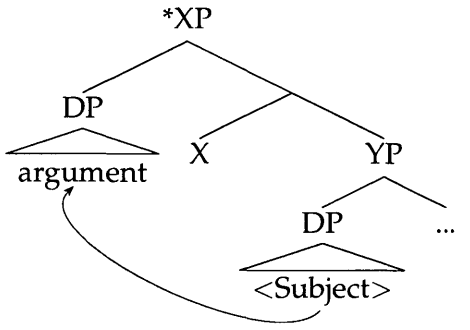
Erlewine's definition of spec-to-spec anti-locality is given in (6). The primary goal of Erlewine's constraint is to rule out Ā-movement chains that are "too short." In Erlewine's terms, "too short" can be understood as moving from the specifier of one XP to the specifier of a structurally adjacent XP, without crossing any intervening maximal projections. Though somewhat tangential to the current discussion, it's worth noting that Erlewine's constraint was originally proposed to operate in a theory of OT syntax. In that system, spec-to-spec anti-locality was a

highly ranked, though not inviolable constraint.

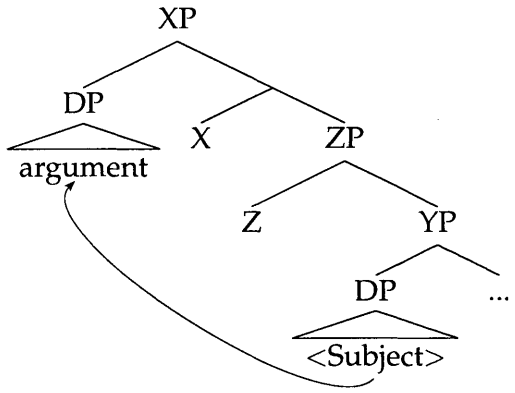
(6) **Spec-to-Spec Anti-Locality (Erlewine, 2016)**
Ā-movement of a phrase from the specifier of XP must *cross* a maximal projection other than XP. Movement from position α to β *crosses* γ if and only if γ dominates α but does not dominate β.

Erlewine's constraint is unlike Grohmann's and Abel's constraints in that is applies only to Ā-movement, not movement in general. This formation of anti-locality predicts that the construction in (7) is ungrammatical, as it involve short movement from spec-YP to spec-XP that does not cross any intervening projection. The example in (8), however, is grammatical under Erlewine's anti-locality: the DP crosses ZP when it moves from spec-YP to spec-XP, and thus it is long enough to be licensed.

(7) Violates Spec-to-Spec Anti-Locality



(8) Satisfies Spec-to-Spec Anti-Locality



An important feature of Erlewine's anti-locality is that it restricts movement only in certain configurations. A consequence of this is that an anti-local configuration can be obviated. While movement between spec-YP to spec-XP is restricted in (7), the same movement operation is possible in (8). It is not a specific property of XP or YP that prohibits or allows movement in (7)-(8), it's the relationship XP and YP. Specifically, it's the fact that no other projection intervenes between XP and YP.

Thus, Erlewine's anti-locality predict that the following alternations will occur: languages sensitive to anti-locality should display at least one configuration where movement between two projections is prohibited when those projections are adjacent, but possible when another head intervenes between the two

projections. For example, this predicts that subject movement from spec-TP to an adjacent spec-CP should be ruled out, but that subject movement from spec-TP to a *non-adjacent* spec-CP (i.e., movement to spec-CP across an intervening maximal projection) should be possible. Importantly, Erlewine's anti-locality has particular consequence for subject extraction, since subject Ā-movement is often assumed to involve movement from spec-TP to the specifier or an adjacent CP.

Kaqchikel Agent Focus

Erlewine (2016) claims that Kaqchikel Agent Focus alternations display exactly the type of pattern predicted by the anti-locality constraint in (6). In Kaqchikel, matrix questions (and other Ā-movement constructions) show an asymmetry in verbal agreement that depends on whether a subject or non-subject is extracted. The canonical agreement pattern for regular transitive clauses is shown in (9), where the subject controls agreement on the verb.[1] In the following examples, a dashed line indicates an agreement (not movement) relation.

(9) Iwïr x-***u***-tëj ri wäy ri a Juan
yesterday ASP-A_{3sg}-eat the tortilla Juan
'Yesterday, Juan ate the torilla.'

Non-subject questions show the same canonical pattern as (9), where the subject is the agreement controller:

(10) Achike x-***u***-tëj ri a Juan?
what ASP-A_{3sg}-eat Juan
'What did Juan eat?'

Subject questions, however, do not follow the canonical pattern. Unlike (10), the subject in (11) does not control agreement. Instead, the sentence displays a special form of the verb, which is marked with what is called an *Agent Focus* morpheme. In sentences that allow for the special Agent Focus morpheme, canonical agreement is ungrammatical:

(11) Achike *x-***u**-tëj / ✓x-∅-tj-**ö** ri wäy?
who ASP-A_{3sg}-eat / ASP-eat-AF the tortilla
'Who ate the tortilla?'

[1]For clarity, the examples in this section do not gloss Kaqchikel object agreement (which is null in all examples given).

Erlewine makes the following assumptions about Kaqchikel subject agreement and Agent Focus. First, he proposed that the Kaqchikel EPP is a violable syntactic constraint, ranked below anti-locality in an OT theory of syntax. This means that Kaqchikel sentences do not require a subject to move through spec-TP, though such movement is still possible in the language. Additionally, Erlewine proposes that the presence or absence of subject agreement on the Kaqchikel verb can be used to track whether a Kaqchikel subject has passed through spec-TP, as Erlewine's proposes that Kaqchikel subject agreement manifests as a result of an Agree relation between T and a subject DP in spec-TP. In (10)-(11), the presence of subject agreement shows that the subjects have moved to spec-TP during the course of the derivation. However, the lack of subject agreement in (12) shows that the subject has not passed through spec-TP. Instead, in (12), Erlewine proposes that the subject has moved directly from its thematic position in spec-vP to an information structure position in spec-CP.

This alternation, as analyzed by Erlewine, captures half of the predictions of spec-to-spec anti-locality: Kaqchikel appears to only allow subject Ā-movement when the movement chain extends directly from spec-vP to spec-CP, and it appears to ban short subject movement from spec-TP to spec-CP. However, Erlewine also observes another fact about Kaqchikel Agent Focus alternations that is directly predicted by anti-locality and ties the occurrence of Agent Focus to the structural proximity between CP and TP. Erlewine observes that subject extraction triggers Agent Focus only when TP and CP are adjacent in the structure; when there is a maximal projection intervening between TP and CP, Agent Focus does not occur. This is shown in (12), when an adverb intervenes between TP and CP. Note, also, that the inclusion of the adverb in (12) completely resolves the asymmetry: in (12), the inclusion of Agent Focus is, in fact, ungrammatical.

(12) Achike **kanqtzij** ✓ x-*u*-tëj / *x-tj-*ö* ri wäy?
who **actually** ASP-A_{3sg}-eat / ASP-eat-AF the tortilla
'Who actually ate the tortilla?'

Erlewine assumes that adverbs are housed in dedicated projections on the clausal spine (as in Cinque 1999), and that *kanqtzij* attaches as an AdvP between TP and CP[2]. The subject can now A-move to spec-TP, enter into an Agree relation with T,

[2]Erlewine does not distinguish whether or not he believes that the adverb is the head or the specifier of the AdvP (and such a distinction would have no bearing on his analysis). The modified

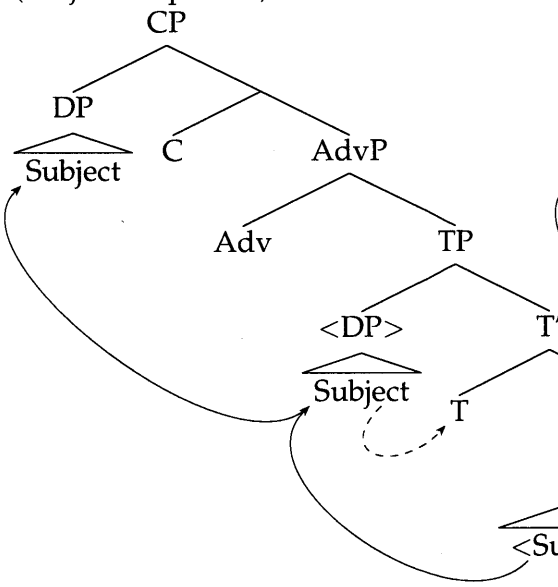
and then Ā-move to spec-CP, across the intervening AdvP. The inclusion of this projection, Erlewine argues, is what makes Agent Focus ungrammatical in (12), and allows for canonical agreement. Comparing (11) and (12) directly, this Agent Focus effect seems less like a subject movement phenomenon, and more like a *short* subject movement phenomenon.

To review: canonical subject movement (which triggers canonical subject agreement) is impossible when the subject would need to undergo movement directly from spec-TP to spec-CP without crossing any intervening projections. However, canonical subject agreement is once again possible when the movement chain is *lengthened* such that the subject crosses both TP and an adverbial projection on its way to spec-CP. In Agent Focus constructions then, the subject moves directly from spec-vP to spec-CP, violating the lower-ranked and violable Kaqchikel EPP. Because the subject never passes through spec-TP, Agent Focus—a type of default agreement—surfaces on the verb in lieu of canonical subject agreement. In (12), the presence of AdvP obviates the anti-local configuration by lengthening the path of movement from spec-TP to spec-CP. In (12), then, there is no need to violate the Kaqchikel EPP, as movement from spec-TP to spec-CP crosses *both* TP and AdvP, respecting Erlewine's anti-locality constraint.

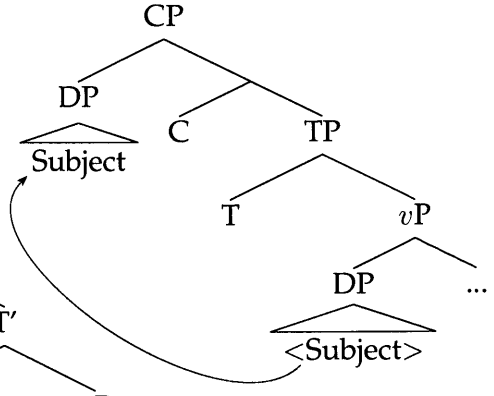
These movement chains are illustrated in the examples below. In (13), the movement chain proposed for the grammatical version of (12) (no Agent Focus), an intervening adverb allows subject movement to proceed from spec-TP to spec-CP. In (14), the movement chain proposed for the grammatical version (Agent Focus) of (11), the Agent Focus surfaces because the subject never moves through spec-TP.

anti-locality constraint introduced in this chapter requires that adverbs be the specifiers of AdvPs.

(13) Verbal agreement
(subject in spec-TP)



(14) Agent Focus
(nothing in spec-TP)



This discussion and the structures in (13)-(14) show that Erlewine's spec-to-spec anti-locality does not reduce to a ban on subject interrogation or extraction. Rather, Erlewine's spec-to-spec anti-locality constitutes a set of restrictions on how subject extraction can take place. The Kaqchikel facts provide two ways to get around spec-to-spec anti-locality's restrictions on short subject movement. A language can lengthen a subject Ā-movement chain by inserting an intervening XP, in this case an AdvP, between CP and TP, or a language can lengthen a subject Ā-movement chain by extracting the subject from its thematic position in vP directly to an information structure position in the CP. Later, this chapter will show that there is an additional (trivial) way to satisfy anti-locality: instead of extracting a subject, a language can allow the subject to remain *in situ*.

Coon & Henderson's (2017) response to Erlewine

Section 3.2.1 outlined the data that Erlewine originally claimed demonstrated the footprint of anti-locality. Under Erlewine's account of Kaqchikel syntax, Agent Focus alternations follow directly from spec-to-spec anti-locality. Importantly, this is not the only subject/non-subject asymmetry that appears to adhere to the predictions of by Erlewine's spec-to-spec anti-locality—sections 4 and 5 of this chapter will also detail a number of other examples that look to be motivated by a spec-to-spec anti-locality constraint.

Coon and Henderson (2017), in a reply to Erlewine (2016), claim that Erlewine's theory of agreement—particularly, the claim that transitive subject agreement is established when the subject moves through spec-TP—is not a viable analysis for Kaqchikel. Coon & Henderson (2017) note that this analysis differs from how agreement has been proposed to work in a number or related Mayan languages, such as Chol and Yucatec Maya. In those languages, transitive subjects are argued to remain *in-situ* and be licensed low in the derivation (c.f. Aissen 1992, Coon 2013). Coon & Henderson argue that Kaqchikel is consistent with other Mayan languages. They claim that only intransitive Kaqchikel subjects move through spec-TP, transitive Kaqchikel subjects remain in their thematic position unless extracted to an information structure position in the CP domain. If this is correct, Erlewine's specific analysis of Kaqchikel agreement syntax—which relies on transitive subjects moving through spec-TP to establish an agreement relation with the verb—cannot be upheld.

Additionally, Coon & Henderson provide examples of subjects that are apparently Ā-extracted, but do not require Agent Focus on their verb (15) (2016:9 ex18b) as well as cases where Agent Focus appears alongside extraction and the presence of an adverb (16) (2016:11 ex23). These examples, additionally, do not conform to Erlewine's analysis. In fact, Erlewine predicts that (15)—which shows a subject *wh*-questions with canonical agreement on the verb and without an adverb between TP and CP—should be ungrammatical, as it would require the subject to undergo anti-local movement from spec-TP (where subject agreement is established) to spec-CP. Similarly, (16) does not conform to the theory of OT syntax that underlies Erlewine's approach. The presence of the overt adverb allows the subject to move from spec-TP to spec-CP. Because this movement is possible, the derivation with the fewest number of violations should be the one where the subject satisfies the Kaqchikel EPP, and so Erlewine's analysis does not predict Agent Focus to be able to grammatically occur on examples like (16).

(15) [Achike]$_i$ x-**u**$_i$-k'äm pe r$_i$-ixayil pa nimaq'ij
who$_i$ x-ERG3S-bring DIR ERG3S-wife PREP party
'Who brought his (own) wife to the party?'

(16) ri **kanqitzij** y-e-ya'-**on** ru-q'ij k'ochi e ütz
REL **actually** y-ABS3P-give-**AF** ERG3S-day must ABS3P good
'Those who truly present his word must also be good.'

Additionally, Coon & Henderson (2017) propose that Agent Focus is not

connected to subject extraction. The full details of Coon & Henderson's counterargument to Erlewine are beyond the scope of this discussion, but their claims can be summarized as follows: They claim that Kaqchikel is a language with many silent elements, including a null copula and complementizer. The present of these silent elements makes it difficult to distinguish biclausal and monoclausal Kaqchikel sentences, at least on first glance. In fact, Coon & Henderson propose that structures like (11)—a subject *wh*-question that is grammatical with Agent Focus and ungrammatical with subject agreement—and (12)—a subject *wh*-question that includes an adverb and canonical subject agreement, and which cannot occur with Agent Focus—are actually quite distinct. Following the previous literature on Agent Focus in Mayan (e.g., Ordonez 1995), Coon & Henderson propose that Agent Focus actually tracks ergative subject extraction within a clause. They assign (11) the structure given in (17) and argue that Agent Focus is required to appear on the verb because it indicates that the subject has extracted to the CP domain.

(17) $[_{CP}$ achike$_i$ $[_{vP}$ t$_i$ x-tj-ö ri wäy]]
$[_{CP}$ who$_i$ $[_{vP}$ t$_i$ ASP-eat-AF the tortilla]]
'Who ate the tortilla?'

Coon & Henderson do not specifically discuss whether the movement chain in (17) involves subject movement through spec-TP, though if, following their analysis, only intransitive Kaqchikel subjects move through spec-TP, there is no external reason to expect the subject in (17) to move through spec-TP on its way to spec-CP. Thus, while the structure proposed in (17) does not provide evidence in favor of anti-locality, it does also not provide evidence *against* an Erlewine's anti-locality constraint (assuming his constraint is found to be active in other languages). While Coon & Henderson's response to Erlewine can be understood as an argument against spec-to-spec anti-locality as a explanation for Kaqchikel Agent Focus alternations, it cannot be takes as an argument against Erlewine's spec-to-spec anti-locality directly.

The reason that Agent Focus cannot occur on (12), Coon & Henderson argue, is because (12) is, in fact, a biclausal construction with a silent copula and complementizer—a structure very distinct from that in (11)/(17). Coon & Henderson propose that (12) is actually an example of something called Adverbial Predication, and that the adverb in (12) actually predicates a copular construction which is capable of embedding another clause with a null subject

pro. Because the *pro* subject of the embedded clause does not actually undergo extraction, there is no Agent Focus on the embedded verb. Though the copular subject extracts, the nature of the null copula means that overt Agent Focus does not surface in this particular extraction context. The Coon & Henderson structure for (12) is given in (18). Here, Coon & Henderson do not give evidence that the matrix subject is required to undergo anti-local Ā-movement. In fact, they do not specify the exact nature of the matrix subject movement chain, and so the structure in (18) abstracts away from this movement operation.

(18) achike$_i$ COP kanqtzij [$_{CP}$ (COMP) pro$_i$ x-u-tëj ri wäy]
who$_i$ COP actually [$_{CP}$ (COMP) pro$_i$ ASP-A3SG$_i$-eat the tortilla]
'Who (what is that) actually ate the tortilla?'

The nature of examples like (15) and (16) become easier to understand given this system. The sentence in (16), for instance, is how the example in (18) should look if the embedded subject does undergo extraction. Similarly, (17) can be understood as an structure where 'achike', *who*, does not undergo extraction at all, and so no Agent Marking occurs on the verb.

Given Coon & Henderson's response, it is not immediately clear that anti-locality is the best framework for understanding the Kaqchikel Agent Focus alternations. However, as previously mentioned, Coon & Henderson's analysis of Kaqchikel Agent Focus alternations does not constitute an argument against a general theory of spec-to-spec anti-locality. None of their structures, for example, require a subject to undergo short subject movement from spec-TP to spec-CP across no intervening material. To that end, this chapter will not attempt to adjudicate which model—Coon & Henderson or Erlewine—best captures the subtleties of Kaqchikel agreement facts. While Kaqchikel Agent Focus alternations were the original motivation for Erlewine's spec-to-spec anti-locality constraint, it is possible that this analysis will not to hold up to further scrutiny from linguists more familiar with the working of Mayan languages (ust as it is possible that Erlewine's analysis is ultimately correct, and that Kaqchikel handles its agreement relations distinctly from other Mayan languages). Indeed, a core goal of this chapter is to argue that the strength of anti-locality lies in the fact that its effects are quite far-reaching. To serve that goal, this chapter presents (what I hope are) a number of alternations that can be explained by anti-locality.

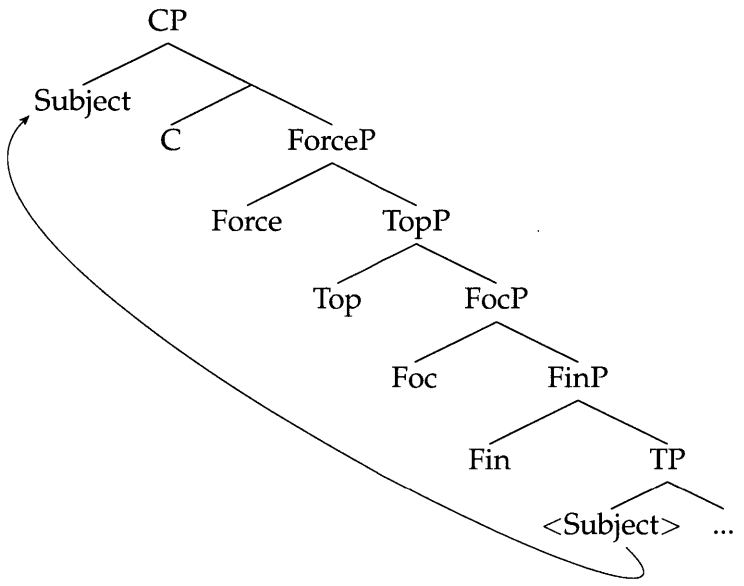
3.3.3 Brillman & Hirsh's specifier-motivated anti-locality

Brillman & Hirsh's (2016) definition of spec-to-spec anti-locality is given in (19); the differences between this definition and Erlewine's formal definition, given in (6) are underlined.

(19) **Spec-to-Spec Anti-Locality**
Ā-movement of a phrase from the specifier of XP must **cross** a specifier projected by a maximal projection other than XP. Movement from position α to β **crosses** γ if and only if γ dominates α but does not dominate β.

Brillman & Hirsh's definition of spec-to-spec anti-locality was proposed to account for a consequence of Erlewine's theory of anti-locality. Namely, Erlewine's spec-to-spec anti-locality depends on there being a very small syntactic structure at the left-periphery of the clause. Erlewine's spec-to-spec anti-locality is defined such that the existence of any intervening head between TP and CP would be sufficient to license subject movement between the two projections and so obviate an anti-local configuration. As such, Erlewine's definition is fundamentally opposed to syntactic models like Cartography, which propose the existence of numerous additional projections along the clausal spine between CP and TP. Under Erlewine's definition of spec-to-spec anti-locality, the movement chain in (20) is not anti-local: the movement chain is permitted even if all the heads between TP and CP are phonologically silent and lack an overt specifier.

(20) Anti-locality does not ban movement across overt or silent projections without specifiers[3]



The structure in (20) is permitted by Erlewine's anti-locality because the subject crosses four maximal projections other than TP (FinP, FocP, TopP, ForceP) before reaching spec-CP. One obvious challenge of Erlewine's spec-to-spec anti-locality is reconciling it with theories of syntax, e.g., cartography, that propose a large number of possible nodes that are often silent in the syntax[4]. This lends a certain fragility to Erlewine's definition of spec-to-spec anti-locality: the addition of a single node which is said to always be present between TP and CP—such as a FinP that would always be present in the derivation, indicating whether a clause is finite or non-finite—would, effectively, dismantle any expectation of observing anti-locality effects.

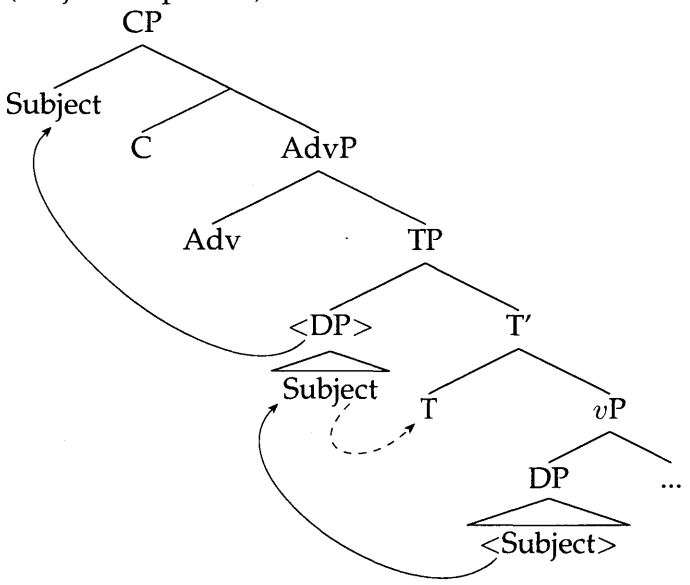
For example, a classical Cinquean approach to adverbs—where a series of AdvP projections (FrequencyP, MannerP, etc) are always present along the clausal spine, and where adverbs, when they are present in the derivations, occur as specifiers of those AdvP—poses a problem to Erlewine's original analysis of Kaqchikel. If AdvP is always present in the derivation, then there is no anti-local

[3]This dissertation will assume a theory of syntax in line with bare phrase structure (? 1994), where a head can only be understood to project a specifier if that specifier is either (i) phonologically overt or (ii) semantically active.

[4]Erlewine's anti-locality proposal could also be used as an argument against such theories of syntax—with the basic premise being that, since we see anti-locality effects in vanilla clauses, we would predict that projections like TopP and ForceP are not present in those constructions. To the best of my knowledge, neither Erlewine nor any of his adopters have made these arguments in the literature.

configuration in vanilla Kaqchikel *wh*-questions, and nothing to prevent the subject from agreeing with the T (and thus triggering canonical subject agreement on the verb) before then crossing TP and AdvP on its way to spec-CP, as illustrated in (21).

(21) Verbal agreement
(subject in spec-TP)



Brillman & Hirsh's (2016) definition of spec-to-spec anti-locality then, is an attempt to marry the spirit of Erlewine's original constraint with a theory of syntax that allows for a proliferation of silent heads along the clausal spine. However, this approach is also not without its complications. While Erlewine's definition of spec-to-spec anti-locality requires the assumption that there fewer heads than previously proposed within the syntax, Brillman & Hirsh's theory requires, at points, both the possibility of multiple specifiers and a theory of syntax with perhaps more specifiers than other arguments have previously proposed.

3.3.4 Predictions of spec-to-spec anti-locality

The majority of the remainder of this chapter will be dedicated to explaining how spec-to-spec anti-locality. When spec-to-spec anti-locality is generally discussed in the remainder of this chapter, it can be informally understood as a constraint that limits Ā-movement from the specifier position of one XP to the specifier position of its neighboring XP, across no intervening material. This informal

definition will only be used in the (many) cases where Erlewine's and Brillman & Hirsh's definitions of spec-to-spec anti-locality make the same predictions. In the cases where Erlewine's and Brillman & Hirsh's formal definitions of spec-to-spec anti-locality make divergent predictions—or require divergent analyses—this chapter will discuss each theory individually.

A more general concept of spec-to-spec anti-locality—which rules out short Ā-movement that does not cross "intervening material," a term left intentionally undefined—makes a number of predictions about the typology of Ā-movement operations that should potentially be able to occur in the grammar. Most obviously, spec-to-spec anti-locality rules out short-distance subject movement from spec-TP to spec-CP. However, spec-to-spec anti-locality does not constitute a ban on subject movement between spec-TP and spec-CP. Importantly, anti-local configurations can be *obviated*: that is, an anti-local configuration that bans movement in one instance (e.g., short distance subject extraction from spec-TP to an adjacent spec-CP) can be *neutralized* (by, e.g., the addition of an AdverbP projection with a phonologically overt adverb in its specifier between CP and TP along the clausal spine). When the anti-local configuration is neutralized, movement can then proceed uninhibited (e.g., a subject can undergo longer-distance movement from spec-TP to spec-CP across an overt specifier of an AdvP).

This strategy—the insertion of a maximal projection with an active specifier between two adjacent maximal projections—is not the only way to obviate an anti-local configuration. There are two additional logical possibilities. First, an anti-local configuration can be obviated if movement can occur along a longer path, without the insertion of additional phonological material. This strategy for subject movement was briefly discussed in the summary of Coon & Henderson's (2017) response to Erlewine (2016), where the subject of a transitive Agent Focus construction can be understood to move directly to spec-CP from its thematic position in spec-vP. Later, this chapter will also show that subject movement out of embedded clause provides another example of this: spec-to-spec anti-locality is predicted to rule out movement of a subject from the embedded subject position (spec-TP) into a matrix clause if the subject would first need to move directly from the embedded spec-TP to the embedded spec-CP, a standard assumption of phase theory. However, spec-to-spec anti-locality is not predicted to restrict embedded subject movement if, for instance, an expletive subject is allowed to sit in spec-CP, allowing the subject to move to spec-CP directly from its thematic position.

Additionally, anti-locality is trivially satisfied if an argument doesn't move at all; cases where a subject remains in-situ trivially satisfy anti-locality. Discussions of *wh*-in-situ will be particularly important for a the discussion of the relationship between spec-to-spec anti-locality and matrix *wh*-questions. One consequence of spec-to-spec anti-locality is that English matrix subject *wh*-questions and relative clauses must be understood as *wh*-in-situ constructions, since anti-locality predicts that the movement chain required by most analyses of subject *wh*-questions and relative clauses is ungrammatical (Section 5 of this chapter will provide evidence for understanding these examples as *in*-situ constructions).

The remainder of this chapter aims to show how the predictions of spec-to-spec anti-locality can explain a wide range of subject/non-subject asymmetries, as well as demonstrating that spec-to-spec anti-locality can capture the same breadth of subject/non-subject alternations captured by competing linguist theories. Section 4 presents a number of examples taken from Shlonsky & Rizzi's (2007) discussion of *Criterial Freezing*, the most comprehensive alternative theory of subject/non-subject asymmetries currently available outside of an anti-locality approach. Section 5 discusses how spec-to-spec anti-locality operates across at least three English subject/non-subject asymmetries. For each asymmetry, this chapter will show that ungrammaticality arises when spec-to-spec anti-locality is violated, and disappears when the anti-locality configuration is resolved. Importantly, the English data give evidence not only for the obviation strategies that Erlewine proposed were active in Kaqchikel, but also for the two additional obviation strategies discussed in this section.

3.3.5 Spec-to-spec anti-locality and a theory of syntax

Of course, the predictions that spec-to-spec anti-locality makes are also dependent on theory of syntax within which the constraint operates. Unless otherwise stated (e.g., in a discussion of a specific proposal from the literature, or a particular phenomenon from a language other than English), this dissertation (and this chapters) assumes the following theory of syntax.

On subjects: We assume that the EPP is active in all languages, though its strength may differ across languages. English is assumed to have a strong EPP which requires a full DP to sit in spec-TP in order to be satisfied. However, other languages may have a weak EPP, where verbal agreement, a clitic, or something other than a full DP can satisfy the EPP. Contra Erlewine's analysis of Kaqchikel

Agent Focus alternations, this dissertation will not assume that the EPP is a violable constraint; the remaining analysis operates under the assumption that if the EPP is not satisfied in a given derivation, that derivation is ungrammatical.

On phases: This dissertation does not assume the standard analysis of phasehood. The standard analysis proposes, roughly, that all CPs, DPs and vPs are phases. Under this approach, only certain syntactic heads can be phase heads. These heads are either born as true phase heads—in which case, they require arguments to move through their specifiers in order to vacate the phase—or they can be "defective" (e.g., unaccusative vP)—in which case, arguments can vacate the phase without successive-cyclic movement through the phase edge. As discussed in the introduction to this dissertation, my analysis assumes the theory of phasehood put forward in Wurmbrand (2017), where phasehood is contextually specified. Under this analysis, phasehood is a property of a complete cyclic domain, not a specific head. A cyclic domain can be understood as a domain that contains a predicate that takes at least one argument and any elements that this predicate takes as an argument. A complete cyclic domain can be understood as any cyclic domain where all the theta roles available to a certain predicate have been assigned to arguments within its scope. This definition allows for a much more flexible definition of phasehood. Particularly, this analysis allows for cyclic domains that are *smaller* than previously proposed. For example, a VP might embed a bare TP–and that TP can be considered a complete cyclic domain if all the requirements of T are met. If that bare TP meets the definition of a complete cyclic domain, that TP itself will be the phase edge. For arguments supporting this proposal, see Chapter 1 of this dissertation. This understanding of phasehood, however, is still subject to the PIC: an argument cannot move out of a phase without first moving (successive-cyclically) to the phase edge.

On A and Ā-movement and locality: This dissertation assumes that there is some fundamental difference between A and Ā-movement. At the very least, A and Ā-movement operations are subject to different types of restrictions. Particularly, only Ā-movement is subject to spec-to-spec anti-locality; A-movement operations are proposed (and assumed) to be unaffected by spec-to-spec anti-locality constraints. This dissertation also assumes that there are two kinds of locality constraints present in the grammar. The first is a condition of Ā extraction out of a phase edge, the PIC, which limits long distance Ā extraction across phase boundaries. This locality constraint is assumed to interact significantly with anti-locality. The second locality constraint is relativized

minimality, which limits the ability of arguments to A-move across another argument. Relativized minimality is assumed to be something that effects A-movement operations alone, and so is not predicted to interact with anti-locality.

This section lays out the set of assumptions needed to understand the arguments presented in this chapter. However, it also does not discuss any connections between the constraints proposed in this chapter—for example, it states that spec-to-spec anti-locality exclusively affects Ā-movement operations, and that certain locality constraints exclusively effect A-movement operations, but does not assume any connection between these statements. For a less conservative discussion of the relationship between A and Ā-movement and spec-to-spec anti-locality, see Section 6.

3.3.6 Interim summary

This section has discussed both Erlewine's and Brillman & Hirsh's spec-to-spec anti-locality constraints. The definition of Erlewine's constraint is given in (22), repeated from (6), and the definition of Brillman & Hirsh's constraint is given in (23), repeated from (19).

(22) **Erlewine's spec-to-spec anti-locality**
Ā-movement of a phrase from the specifier of XP must *cross* a maximal projection other than XP. Movement from position α to β *crosses* γ if and only if γ dominates α but does not dominate β.

(23) **Brillman & Hirsh's spec-to-spec anti-locality**
Ā-movement of a phrase from the specifier of XP must **cross** a specifier projected by a maximal projection other than XP. Movement from position α to β **crosses** γ if and only if γ dominates α but does not dominate β.

At a high level, both definitions of spec-to-spec anti-locality make consistent predictions. When discussing this cases, this chapter will refer to a *general theory* of spec-to-spec anti-locality. However, there are also points where the distinct definitions of Erlewine's and Brillman & Hirsh's constraints lead to distinct predictions. In those cases, this chapter will discuss the predictions and trade offs of each constraint individually.

This section has discussed why anti-locality is disproportionately likely to affect subjects. Subjects are the arguments most likely to undergo short Ā-movement because of the interaction of the EPP and conditions on phasehood.

CP layers are often phase layers (even under a contextualized specified theory of phasehood), and a CP is often adjacent to TP. Because the EPP requires that subjects A-move to spec-TP, any subject Ā-movement operation is disproportionately likely to be affected by spec-to-spec anti-locality.

This section has also discussed three strategies that languages are predicted to use to avoid requiring movement that violates a general spec-to-spec anti-locality constraint. First, a language can lengthen a subject chain by including an additional projection (under Brillman & Hirsh's formulation, an additional projection with a semantically active specifier) between TP and CP on the clausal spine. Second, a language can lengthen A-movement chain by allowing the subject to move directly from its thematic position in spec-vP to spec-CP, without EPP-driven movement through spec-CP. Finally, a language can trivially satisfy anti-locality by allowing the subject to remain in-situ.

3.4 Anti-locality cross-linguistically

This section will discuss four "classic" cross-linguistic subject/non-subject asymmetries: question formation in Imbabura Quechua and the distribution of resumptive pronouns in Hebrew complex relative clauses, and anti-agreement effects in Tarafit Berber and the *que/qui* effect in French. The first two asymmetries are also discussed in Rizzi and Shlonsky's (2007) paper on Criterial Freezing, a competing theory on the unification of subject/non-subject asymmetries. This section will discuss how three of the four disparate effects can be captured by spec-to-spec anti-locality. This section will also discuss how anti-locality is also only predicted to explain one *half* of the *que/qui* alternation—the fact that subjects are banned from extracting from embedded clauses introduced by *que*–but not necessarily the other half—the fact that objects are banned from extracting from embedded clauses introduced with *qui*.

This section has two main goals. First, to show that there is evidence for the existence of a spec-to-spec anti-locality constraint from a number of unrelated languages and to outline the type of cross-linguistic alternations that anti-locality is predicted to be able to account for. Second, this section aims to show that spec-to-spec anti-locality has at least the same predictive power as Criterial Freezing. Importantly, this section will primarily discuss how anti-local configurations restrict possible subject movement operations in these languages. Section 5, which turns to English, will discuss a variety of scenarios where

spec-to-spec anti-locality ban certain subject movement operations, but where obviating the anti-local configuration in question allows for the possibility of subject movement again.

3.4.1 Criterial Freezing: a primer

Rizzi & Shlonsky's (2007) theory of Criterial Freezing, building on Rizzi (2003), essentially claims that subjects become "frozen"—i.e., unavailable for further movement operations—when they reach a "criterial" position in the syntax. Here, a criterial position is defined as "a position dedicated to some scope-discourse interpretive property." In this way, Rizzi & Shlonsky's theory makes a somewhat more explicit mention of the nature of subjecthood, compared to spec-to-spec anti-locality. Rizzi & Shlonsky (2007) take the EPP ("the requirement that clauses have subjects") to indicate that subjecthood is itself a scope-discourse interpretive property; objects do not show Criterial Freezing effects because there is no analogous object-EPP property, and so the object position is not predicted to be a meaningful scope-discourse position. Their theory follows from observations that other arguments that move to scope-discourse positions (e.g., *wh*-words, topics, foci) cannot undergo further Ā-movement. This makes Criterial Freezing, and the Subject Criterion (Rizzi & Shlonsky's interpretation of the EPP) formally analogous to proposals like the Topic Criterion and Focus Criterion, which limit the movement of topics and foci.

Like spec-to-spec anti-locality, Criterial Freezing can be obviated. Rizzi & Shlonsky state that Criterial Freezing should not limit subject movement when "the thematic subject is allowed to skip the criterial Subject position" or when a language employs "strategies for forming Ā chains on embedded subjects without moving them" (2007:2). It's worth noting that this second strategy applies to more than simply leaving a subject *in-situ*—many structures involving resumptive pronouns are included in this obviation strategy in Rizzi & Shlonsky' (2007).

Importantly, this claim makes the prediction that subjects which do not undergo Criterial Freezing will have scope-discourse properties that are systematically different from the scope-discourse properties of subjects that do undergo Criterial Freezing[5]. Rizzi & Shlonsky also adopt a theory where a designated SubjectP is the Criterial position in which subjects are frozen; thus

[5]To the best of my knowledge, this Rizzi & Shlonsky do not provide any evidence to support this claim.

their theory also claims that all languages should involve a designated SubjectP along the clausal spine. Thus, the specifier of SubjectP constitutes Rizzi & Shlonsky's EPP position (though this dissertation assumes that spec-TP is the canonical EPP position, and so will treat spec-TP as the designated subject position).

3.4.2 Pied-piping in Imbabura Quechua

As first noted by Cole (1985), Imbabura Quechua (Ecuador) has two strategies for question formation: *wh*-movement (24a) and clausal pied-piping (24b). In object *wh*-questions, both options are available, as illustrated in (24). In (24a), the *wh*-word *ima-ta*, "what," moves from the embedded clause to the left-periphery of the matrix clause, illustrating the availability of *wh*-movement in Imbabura object questions. In (24b), however, it is the entire finite embedded clause, *ima -ta wawamiku-chun*, "what the child eats," that moves to the left-periphery of the matrix clause. This shows that clausal pied-piping is an available interrogation strategy for object *wh*-questions in Imbabura Quechua.

(24) Object *wh*-questions in Imbabura Quechua

a. ima -ta -taj Maria -ka Juzi miku-shka *t* -ta kri-n?
what ACC Q Maria TOP Juzi eat-NMLZ *t* ACC belive-AGR
'What does Maria believe that Juzi ate?' (R&S:2007, ex. 15)
object wh-movement ✓

b. ima -ta wawamiku-chun -taj Maria kri-n *t*?
what ACC child-eat.FIN Q Maria believe-AGR *t*
'What does Maria believe that the child eats?'
Lit: '[what the child eats] does Maria believe?' (R&S:2007, ex. 16)
object clausal pied-piping ✓

Imbabura subject *wh*-questions behave differently. Unlike object *wh*-questions, straightforward subject *wh*-movement is not possible in Imbabura embedded subject *wh*-questions (25a). However, the clausal pied-piping strategy is still available, as illustrated in (25b).

(25) Subject *wh*-questions in Imbabura Quechua

a. *pi -taj Maria -ka *t* chayamu-shka -ta kri-n?
who Q Maria TOP *t* arrive-NMLZ ACC believe-AGR?
INTENDED: 'Who does Maria believe arrived?' (R&S:2007, ex. 17)
subject wh-movement *

b. [pi chayamu-shka -ta -taj] Maria *t* kri-n
[who arrive-NMLZ ACC Q] Maria *t* believe-AGR
'Who does Maria believe has arrived?'
Lit: [Who has arrived] does Maria believe?
subject clausal pied-piping ✓

Criterial freezing offers a straightforward analysis of these facts: Rizzi & Shlonsky assume that subject extraction from an embedded clause is ruled out because the subject would have had to undergo movement to the embedded SubjectP, a Criterial Freezing position. Following that movement the embedded subject would be frozen in place within the embedded clause and so unable to move higher. There is no parallel restriction for object *wh*-movement because the object does not need to move to any Criterial position within the embedded clause. Thus the object can freely extract into the matrix clause. Pied-piping the entire embedded clause allows the subject to stay in the Criterial Position while still permitting an interrogated element to move to a question position in the CP. This analysis should be possible to extend to other languages that display this footprint with respect to clausal pied piping and subject Ā-movement.

This alternation also follows easily from anti-locality: embedded subject *wh*-movement is prohibited because it would require anti-local Ā-movement. This dissertation will assume that pied-piping represents movement of a complex QP to the left periphery, following Cable (2010). In (25a), an embedded subject QP would be required to undergo Ā-movement from the embedded spec-TP to the embedded spec-CP. Because this movement crosses no intervening material, it would be prohibited. However, Imbabura Quechua still employs a subject interrogation strategy. The language can be understood to allow both arguments and finite embedded clauses to be QPs. These QPs can move to the left-periphery of the matrix clause without violating spec-to-spec anti-locality. This step involves Ā-movement across the entire matrix clause—a domain containing at least one additional maximal projection and semantically active specifier—which satisfies spec-to-spec anti-locality under both Erlewine's and Brillman & Hirsh's

definition[6].

Subject *wh*-questions are not the only Imbabura Quechua constructions that appear to be constrained by spec-to-spec anti-locality. Subject relative clauses also show restrictions that might appear arbitrary in isolation, but are in fact predicted by anti-locality. Cole (1985) notes that Imbabura Quechua displays two types of relative clauses, which he labels "left branching" and "internally headed." We can understand this proposal as contrasting a head-final, null operator relative clause structure with a relative clause structure where the an overt DP moves within the RC itself, respectively.

Both relative clause structures disallow the relativization of embedded subjects. This is shown in (26) for Cole's left branching relative clauses (null operator structures) and in (27) for Cole's internally headed relative clauses (where an overt DP moves within the embedded clause). Though Imbabura Quechua displays multiple types of relative clauses, the language does not allow relative clauses to form by pied piping, and so it appears that Imbabura Quechua disallows relativized embedded subjects[7].

(26) Left branching relative clauses (Cole 1985:56)

a. *[Marya Juzi-man ni-shka [_ Juan-ta riku-shka]] -ta warmi
[Marya Juzi-to say-NMLZ [_ Juan-ACC see-NMLZ]] -ACC woman
llugshi-rka
leave-PST.3
*'[The woman who Marya told Juze (that) _ saw Juan] left.
(embedded subject relativization ✗)

b. [chay [Marya Juzi _ riku-shka] -ta kri-j wawa]
[that [Marya Juzi _ see-NMLZ] -ACC believe-NMLZ.FUT child]
ña-mi ri-rka
already-VAL go-PST.3
'The child whom Marya believes Juze saw _ already left.
(embedded object relativization ✓)

[6]Of course, to fully illustrate that anti-locality underlies this subject/non-subject alternation, it would be ideal if we could find a minimal pair to (25a) where (i) there is intervening material between the embedded TP and CP and (ii) subject *wh*-movement is allowed. These situation are not necessarily common, but they can appear in certain languages. For example, as will be further discussed in Section 5.1, (at least some) TP-level adverbs are possible in English certain embedded clauses. However, it is not clear that Imbabura Quechua allows for either adverbs or adverbial phrases to intervene between T and C in embedded clauses, so it is unclear whether or not future fieldwork will be able to test these predictions. I leave this question open to future research.

[7]Additionally, Cole (1985:96) states that a similar restriction exists that bans embedded subject from moving to emphatic position in the left periphery, though he does not give any examples of this phenomenon.

(27) Internally headed relative clauses (Cole 1985:57)

a. *[Marya [warmi Juan-ta riku-shka] -ta] ni-shka llugshi-rka
[Marya [woman Juan-ACC see-NMLZ] -ACC] say-NMLZ leave-PST.3
*[The woman that Marya said _ saw Juan] left.
(embedded subject relativization ✗)

b. [Marya [Juan wawa-ta riku-shka] -ta] ni-shka llugshi-rka
[Marya [Juan child-ACC see-NMLZ] -ACC] say-NMLZ leave-PST.3
'[The child that Marya said that Juan saw _] left.'
(embedded object relativization ✓)

The facts in (26)-(27) are predicted under anti-locality. This is because they would both require a structure where an embedded subject undergoes anti-local movement from the specifier of the embedded TP to the adjacent specifier of the embedded CP, parallel to the ungrammatical subject *wh*-movement (QP movement) examples.

Imbabura Quechua matrix subject *wh*-questions also appear to respect anti-locality. Because the subject appears in the same position in declarative and interrogative clauses (and there is no difference in word order between the two), baseline subject *wh*-questions can be analyzed as *wh*-in-situ constructions. Under this analysis, the subject does not extract from the canonical position, trivially satisfying anti-locality (28) (from Cole 1985:19).

(28) a. [$_{subj}$ pi-taj] kan-paj mama-man ali wagra-ta kara-rka
[$_{subj}$ who-Q] you-POSS mama-to good cow-ACC give-PST.3
'Who gave your mother a good cow?'

b. [$_{subj}$ ñuka wawki-ka] ñuka mama-man ali wagra-ta kara-rka[8]
[$_{subj}$ my brother-TOPIC] my mother-to good cow-ACC give-PST.3
'My brother gave my mother a good cow.'

However, multiple matrix *wh*-questions that involve subject interrogation cannot be analyzed as *wh*-in-situ. In these situations, however, the subject can move *across* another overtly-moved *wh*-element, as in (29a). Cole (1985) treats (29b) as a *wh*-extraction construction, and in (29b) it is not possible for the subject to undergo short *wh*-movement to a position below the object *wh*-word. The subject *wh*-movement in (29a) respects anti-locality because the subject's movement path is lengthened by the intervening object. The subject *wh*-movement path in (29b) would have to be anti-local, as no intervening

[8]Though Cole glosses the example in (28) as containing a topic marker on the subject, he also claims that this is canonical subject marker in declarative contexts and does not indicate any movement of the subject within the left periphery.

material can lengthen its movement path[9].

(29) Multiple *wh*-questions

a. pi-taj, ima-ta-taj, riku-rka Utavalu-pi-ka
who-Q, what-ACC-Q, see-PST.3 Otavolo-in-TOP
'Who saw what in Otavolo?'

(SUBJ WH > OBJ WH ✓)

b. *ma-ta-taj, pi-taj, riku-rka Utavalu-pi-ka
what-ACC-Q, who-Q, see-PST.3 Otavolo-in-TOP
*What did who see in Otavolo?'

(OBJ WH > SUBJ WH ✗)

The contrast in (29) is, of course, not unique to Quechua. It is also not specific to anti-locality; (29) is an example of the kinds of superiority effects that have long been noted for languages that allow for *wh*-questions with multiple instances of overt *wh*-movement (e.g., Richards 1997). However, the contrasts in (29) are also consistent with the theory of spec-to-spec anti-locality outlined in this chapter.

3.4.3 Resumptive pronouns in Hebrew

Rizzi & Shlonsky note an interesting fact—originally discussed in Borer (1984:249-250) as evidence for the successive-cyclic nature of long-distance Ā-movement—about complex Hebrew relative clauses. As shown in (30a), complex Hebrew object relative clauses can include a resumptive pronoun in the original object position. However, it is also possible to front the object resumptive pronoun to an information structure position in the left periphery. This can either be a left periphery position of the most local CP, as in (30b), or of a more distant CP layer, as in (30c). In (30c), for example, the resumptive pronoun moves to the left edge of CP_2.

[9]However, these examples are not unambiguously due to spec-to-spec anti-locality, given the multiple *wh*-movement theories laid out in Richards (1997), which independently predict that the superiority violating example in (29b) will be ungrammatical.

(30) Complex Hebrew object relative clauses

a. kaniti et ha-šulxan [$_{CP_1}$ še xana amra [$_{CP_2}$ še dalya
I.bought ACC the-table [that Hannah said [that Dalya
ma'amina [$_{CP_3}$ še Kobi raca **oto**]]]
believes [that Kobi wanted him]]]
'I bought the table that Hannah said that Dalya believes that Kobi wanted (it).'
resumptive object pronoun in-situ

b. kaniti et ha-šulxan [$_{CP_1}$ še xana amra [$_{CP_2}$ še dalya
I.bought ACC the-table [that Hannah said [that Dalya
ma'amina [$_{CP_3}$ še **oto** Kobi raca _]]]
believes [that him Kobi wanted _]]
'I bought the table that Hannah said that Dalya believes that it (was the one which) Kobi wanted.' *resumptive pronoun fronted*

c. kaniti et ha-šulxan [$_{CP_1}$ še xana amra [$_{CP_2}$ še **oto** dalya
I.bought ACC the-table [that Hannah said [that him Dalya
ma'amina [$_{CP_3}$ _ še Kobi raca _]]]
believes [_ that Kobi wanted _]]
'I bought the table that Hannah said that it (was the one which) Dalya believes that Kobi wanted.' *resumptive pronoun fronted*

Complex Hebrew *subject* relative clauses, however, behave slightly differently. Subject relativization is possible in Hebrew, including complex relative clauses parallel to those in (30). In these subject relative clauses, a resumptive subject pronoun can be left *in-situ*, as in (31a). However, this subject resumptive pronoun cannot be pronounced in any higher information structure position, as shown in (31b) and (31c)[10].

[10]These facts are not the only resumptive pronoun asymmetries found in the grammar. For example, McCloskey (1990) notes the existence of what he calls the Highest Subject Restriction, or the fact that resumptive pronouns are not available in non-embedded subject positions in many languages, including Irish and Hebrew (Rasin 2017). Under an analysis where relative clauses are formed viA-movement of an NP/DP to the specifier of a CP, this restriction is easily explained by spec-to-spec anti-locality, since that movement chain would require anti-locality movement (see Koopman and Sportiche 1983, Engdahl 1983, Pesetsky 1998, Rasin 2017, among others for evidence for such a syntax). However, there are also a variety of arguments against such an analysis; Chomsky 1977, Doron 1982, Sells 1984, Shlonsky 1992 and McCloskey 2002 and Rasin 2017 all argue that at least some relative clauses are not formed viA-movement, and that, when resumptive pronouns occur in those relative clauses, they can be understood as garden variety pronouns that are bound to a relative head/operator via some binding mechanism other than movement. Under this analysis of relative clause structures, it is not clear that spec-to-spec anti-locality predicts the Highest Subject Restriction. For this reason, this chapter leaves the question of whether or not spec-to-spec anti-locality can be used to explain the Highest Subject Restriction to future research.

(31) Complex Hebrew subject relative clauses

a. kaniti et ha-šulxan [$_{CP_1}$ še xana amra [$_{CP_2}$ še dalya
I.bought ACC the-table [that Hannah said [that Dalya
ta'anna [$_{CP_3}$ še **hu** ya'ale harbe kesef]]]
claimed [that he will.cost much money
'I bought the table that Hannah said that Dalya claimed that (it) will cost a lot of money.' *resumptive subject pronoun in-situ*

b. *kaniti et ha-šulxan [$_{CP_1}$ še xana amra [$_{CP_2}$ še dalya
I.bought ACC the-table [that Hannah said [that Dalya
ta'anna [$_{CP_3}$ **hu** še _ ya'ale harbe kesef]]]
claimed [**he** that _ will.cost much money
resumptive pronoun cannot be fronted to spec-CP$_3$

c. *kaniti et ha-šulxan [$_{CP_1}$ še xana amra [$_{CP_2}$ še **hu** dalya
I.bought ACC the-table [that Hannah said [that **he** Dalya
ta'anna [$_{CP_3}$ _ še _ ya'ale harbe kesef]]]
claimed [_ that _ will.cost much money
resumptive pronoun cannot be fronted to spec-CP$_2$

Rizzi & Shlonsky claim that the contrast between the object resumptive pronoun facts in (30) and the subject resumptive pronoun facts in (31) can be reduced to an effect of Criterial Freezing. This section will discuss their proposal. However, this section will also argue that the facts in (30)-(31) also follow from spec-to-spec anti-locality.

Background: Hebrew resumptive pronouns are not links in A-movement chain

Before discussing complex Hebrew subject relative clauses, it's worth discussing the mechanics at play in (30)-(31)[11]. Following Shlonsky (2004), this analysis assumes that the relative clauses in (30)-(31) all contain some relative operator at their left periphery. Consistent with the semantics assumed elsewhere in this dissertation, we assume that this operator moved to the left periphery from lower in the clause, though the particulars of this movement are prediction-neutral for the present discussion.

Importantly, the mechanics of operator movement within the relative clause CP are prediction neutral for this analysis because Shlonsky proposes that the relationship between the relative head/operator and the resumptive pronoun is one of *binding*, not movement. That is to say, the operator and the resumptive

[11]However, this analysis is also prediction-neutral with a head-external relative clause analysis of (30)-(31), which would not involve operator movement within the relative clause CP.

pronoun do not represent two links in the same movement chain, instead they are simply bound to the same antecedent. Evidence for this analysis comes from examples like (32), which show that Hebrew relative clauses are impervious to island effects.

(32) hine ha-sendvic še lo yada'ti mi 'axal oto
here the-sandwich that NEG knew.1SG who ate it
'Here is the sandwich that I didn't know who ate.'

If the resumptive pronoun in (32) represented a link in the movement chain for the DP *ha-sendvic*, 'the sandwich,' we would predict (32) to be ungrammatical, as it would require A-movement chain that included an island violation. The grammaticality of (32) suggests that Hebrew resumptive pronouns are not links in A-movement chain, but rather pronouns that are able to bind to a relative clause head/operator through a binding mechanism not tied to movement. I leave the specifics of this binding operation open (though see Chomsky 1977, Doron 1982, Sells 1984, McCloskey 2002, Shlonsky 1992, among other, for similar analysis of relative clauses and resumptive pronouns).

Importantly, the data in (30) shows that Hebrew resumptive pronouns can undergo movement on their own and participate in an independent movement chain. Specifically, in (30b) the resumptive pronoun moves from its base generated position low in CP_3 to an information structure position in spec-CP_3. A similar movement chain is visible in (30c), where *oto* moves from the object position in CP_3 to an information structure position in CP_2.

Two kinds of resumptive pronouns in Hebrew

The data in (30)-(32) are a simplification of the full range of Hebrew resumptive pronoun facts. For instance, the resumptive pronoun in non-island CP relative clause constructions in (33) is optional, while the resumptive pronoun following the preposition in (34) is obligatory[12].

[12]While the examples in (33) are glossed identically, Rasin (2017) notes that the examples in (33a) and (33b) have slightly different interpretations. Specifically, (33a) can have either a *de re* interpretation (meaning that Dani is looking for some woman who actually exists) or a *de dicto* interpretation (meaning that Dani is looking for some woman who may or not exists), while (33b) can only have the *de re* interpretation. See Rasin (2017) for an analysis of these facts, which are beyond the scope of the present discussion.

(33) Optional resumptive pronouns in Hebrew

a. dani yima et ha-iša še-hu mexapes
Dani will-find ACC the-woman that-he seeks
'Dani will find the woman he is looking for.' (*Rasin 2017*)

b. dani yima et ha-iša še-hu mexapes **ota**
Dani will-find ACC the-woman that-he seeks **her**
'Dani will find the woman he is looking for.'
(*Rasin 2017, citing Doron 1982*)

(34) dani yima et ha-iša še-hu xolem al-eya/*∅
Dani will-find ACC the-woman that-he dreams about-her/*∅
'Dani will find the woman he desires.' (*Rasin 2017*)

Rasin (2017), building on work by Sichel (2014), treats the resumptive pronouns in (33) and (34) as fundamentally different elements. Rasin (2017) analyses the resumptive pronouns in (33b) to be a true pronoun that is bound to is relative operator/head—in Rasin's terminology, a *syntactic resumptive pronoun*—while the resumptive pronoun in (34) is analyzed as the ϕ-features of A-moved NP being realized on the D head embedded by the preposition–in Rasin's terminology, (34) involves a *morphological resumptive pronoun*. Importantly, only syntactic resumptive pronouns are independent pronouns that can participate in their own movement chains; morphologically pronouns represent links in relative-clause internal movement chain and so cannot undergo additional movement on their own.

It's worth asking whether or not the optionality of the resumptive pronoun in (33) is predicted to bear on the analysis of the subject/object resumptive pronoun asymmetry in (30)-(31). Rasin (2017), following McCloskey (2002), proposes that the difference between (33a) and (33b) reduces to complementizer choice. Specifically, McCloskey argues that (at least in Irish), the presence of a resumptive pronoun correlate with whether or not the complementizer *aN* or *aL* is selected in a relative clause. The complementizer *aN* allows binding without movement, the operation required for a resumptive pronoun to occur. The complementizer *aL*, meanwhile, signals movement, and so cannot co-occur with optional resumptive pronouns. Like Irish, Rasin (2017) proposes that Hebrew has both an *aN* and an *aL* complementizer form, though, in Hebrew, the two forms are homophonous and both realized as *še*.

Assuming this analysis of the Hebrew complementizer system, all the complementizers in (30)-(31) must be *aN*, which allows for binding without

movement and does not license movement within the relative clause.

Explaining Hebrew resumptive pronoun alternations

The facts in (30) are trivial under both Criterial Freezing and spec-to-spec anti-locality. These facts are trivial under a Criterial Freezing analysis because only subjects move through the designated subject position where freezing occurs. Because objects are not proposed to ever move through this designated position, Rizzi & Shlonsky predict that object movement will not be limited in examples like (30). Similarly, no path in the object movement chain in (30) is predicted to be anti-local (as no path involves short Ā-movement), and so spec-to-spec anti-locality does not restrict the possible resumptive pronoun movement paths depicted in (30).

The facts in (31), however, are within the domain of both Criterial Freezing and spec-to-spec anti-locality. Under spec-to-spec anti-locality, (31b) is prohibited because the subject would first have undergo anti-local movement from spec-TP_3 to spec-CP_3. The example in (31c) is ungrammatical for similar reasons: in order for the subject resumptive pronoun to get to spec-CP_2 it would first need to undergo anti-local movement from from spec-TP_3 to spec-CP_3. Without that movement step, the subject pronoun could not evacuate CP_3 without causing a PIC violation (as the pronoun would need to exit the phase without first moving through the phase edge). We know that CP_3 is a full CP phase because of the presence of the overt complementizer *še*.

Under Criterial Freezing, Rizzi & Shlonsky argue, the subject resumptive pronoun would, by nature of being a subject, need to move to a Criterial subject position in the lowest embedded clause, CP_3. Once the subject resumptive pronoun has moved to that position it would be frozen, ruling out (34b).

3.4.4 Anti-agreement as anti-locality: the case of Tarafit Berber

The previous two sections discussed subject/non-subject asymmetries that were also observed in Rizzi & Shlonsky's (2007) paper on Criterial Freezing. This portion of the chapter moves away from Rizzi & Shlonsky to discuss a different kind of subject/non-subject asymmetries: anti-agreement effects.

Because anti-agreement effects are widespread and occur in a large number of (unrelated) languages, this portion of the chapter will concentrate on one famous anti-agreement effect which has already been discussed extensively in the

literature: Tarafit Berber. Particularly, this section will review a subject/non-subject anti-agreement asymmetry and propose that this asymmetry can be cleanly explained by the interactions between spec-to-spec anti-locality and Tarafit Berber's true agreement system[13]. This subject/non-subject asymmetry is somewhat different from the Hebrew and Quechua examples, however. Tarafit Berber does not display an overt ban on certain subject Ā-movement operations. Instead, as with Erlewine's analysis of the Kaqchikel Agent Focus alternation discussed in Section 3.2.1, the Tarafit Berber verb shows a distinct agreement paradigm when it co-occurs with a subject that undergoes Ā extraction, compared to when it co-occurs with a non-subject that undergoes Ā extraction. However, despite the fact that they appear superficially similar, this section will argue that Tarafit Berber anti-agreement effects and Kaqchikel Agent Focus are distinct asymmetries.

In declarative sentences, Tarafit Berber verbs agree with their subjects, as shown in (35a). However, as first noted by Ouhalla (1993), this is not the case for subject questions. In (35b), a grammatical subject *wh*-question, the verb no longer shows canonical agreement with the interrogated subject; instead a different form of the verb is used. This form is traditionally called the "participial" form of the verb. However, I will descriptively gloss this morpheme as AAE (anti-agreement). Tarafit Berber anti-agreement morphology is fixed, regardless of the person, number or gender of the subject. Like in Kaqchikel, the canonical agreement form is ungrammatical in contexts where anti-agreement is grammatical, as illustrated in (35c). This is the classic footprint of anti-agreement.

(35) Berber anti-agreement facts

a. T-zra tfruxt Mohand
FEM.SG-see girl.FEM.SG Mohand
'The girl saw Mohand.' canonical agreement

b. Man tamghart ay y-zri-n Mohand?
which woman.FEM.SG C AAE-saw-AAE Mohand
'Which woman saw Mohand?' non-canonical agreement

c. *Man tamghart ay t-zra Mohand?
which woman.FEM.SG C FEM.SG-saw Mohand
canonical agreement * in subject *wh*-question

[13]As always, this discussion should not be misunderstood as a claim that all anti-agreement alternations can be captured using anti-locality. Instead, this is a proposal that anti-locality be understood as one of the tools available to help make sense of a rich and complex series of alternations and puzzles.

Though this section will focus exclusively on Berber subject *wh*-extraction, it's worth noting that Berber anti-agreement appears to be a (short) subject Ā-movement effect, and is not limited to *wh*-movement contexts. The same anti-agreement pattern is also found in subject relative clause and focus constructions, as in (36) (from Ouhalla 1993:479).

(36) a. tamghart [$_{RC}$ nni y-zri-n _ Mohand]
woman [$_{RC}$ C AAE-saw-AAE _ Mohand]
'the woman who saw Mohand'
b. tamghart-a ay y-zri-n _ Mohand
woman-DEM C AAE-saw-AAE _ Mohand
'It's this woman who saw Mohand.'

Additional evidence for Berber anti-agreement as a subject Ā-movement effect is given in (37), where the direct object is *wh*-extracted. In (37a) normal agreement occurs, while in (37b), the anti-agreement participial form is ungrammatical.

(37) Berber Anti-agreement is a subject movement phenomena
a. min y-wsha Jamal i Mena?
what MASC.SG-give Jamal to Mena
'What did Jamal give to Mena? (Elouazizi 2005:122, ex 7)
b. *min y-shwi-n Jamal i Mena
what AAE-give-AAE Jamal to Mena
INTENDED: 'What did Jamal give to Mena?' (Elouazizi 2005:122, ex 7)

Before diving into a spec-to-spec anti-locality analysis of the alternations in (36)-(37), it's worth spelling out some assumptions about Tarafit Berber itself and anti-agreement as a general phenomenon.

Background assumption: anti-agreement is still agreement

First, we can examine the Tarafit Berber anti-agreement morpheme, *y-in*, in more detail. Since Ouhalla (1993), the literature has glossed *y-in* as an anti-agreement circumfix. However, Baier (2016) proposes that *y-in* is actually two separate morphemes: a default agreement morpheme (y) and a *wh*-agreement morpheme (n). Though this chapter will continue to use a theory-neutral gloss (AAE) for both segments of the morpheme, it will adopt much of Baier's (2016) analysis of Berber anti-agreement.

Specifically, this chapter will assume, following Baier (2016, 2017), that *y-in* is a

true agreement morpheme[14]. This means that *y-in* represents the result of an Agree relation having taken place between a probe and a goal. In other words, *y-in* is not simply the morpheme that is realized when an Agree relationship fails to be established, but rather it is a true agreement morpheme that indicates a specific Agree relationship.

In his discussion of Tarafit Berber anti-agreement, Baier (2016) argues for collapsing the distinction between *wh*-agreement—a phenomenon in which a designated [+WH] agreement morpheme is realized on the verb—and anti-agreement—a phenomenon where a verb displays a lack of overt agreement in the presence of Ā-movement. Baier (2016) argues that the two morphological effects are instances of the same abstract principle: a verb entering into an Agree relationship with an Ā trace[15]. To Baier, *wh*-agreement is the result of this phenomenon occurring in a language which showcases a dedicated *wh*-agreement morpheme, while anti-agreement is the result of this phenomenon occurring in a language without a designated *wh*-agreement morpheme. Essentially, Baier (2016) argues that anti-agreement is a sub-type of *wh*-agreement (for more discussion of *wh*-agreement in general, see Chung 1994, 1998, Georgopoulous 1991, and Watanabe 1996, among others).

Importantly, for *wh*-agreement to occur, an Agree relation only has to be established between the verb and a [+WH] element, not necessarily an Ā trace. As shown in (38) for Abaza (Baier 2016, citing O'Herin 2002), a Northwest Caucasian language spoken in Russia and Turkey, *wh*-agreement can take place between a verb and a full DP containing a *wh*-word (38a) or between a verb and the null Ā operator in a relative clause (38b). Both (38a) and (38b) showcase the absolutive *wh*-agreement morpheme 'yə.'[16] In the following examples, following Baier's glossing conventions, agreement morphemes are underlined and agreement controllers are wavy underlined.

[14]To Baier (2016), these are actually a set of two similar but distinct true agreement morphemes, though this distinction is prediction neutral for the present discussion.

[15]This is Baier's terminology. However, while Baier uses the term "trace" this theory is also compatible with a theory where the verb enters into a relationship with an Ā trace or any other (c)overt element with a [+WH] feature.

[16]For a full overview of the Abaza agreement system, which is beyond the scope of this discussion, see Baier 2016 and O'Herin 2002.

(38) Abaza *wh*-agreement

a. a-č^{w}al dzač'^{w}əya yə-ta-wa
DEF-sack **what** ABS.WH-in-PRS
'What's in the sack?'

(*wh*-agreement with *wh*-DP)

b. [$_{CP}$ **OP**$_i$ *pro* yə-awə-y-ʃ^{j}tə-z] a-haq^{w}-dəw$_i$
[$_{CP}$ **OP**$_i$ 3SG.M ABS.WH-3SG.M-throw-PST] DEF-stone-big
'the big rock that he threw'

(*wh*-agreement with a null operator)

Given that Tarafit Berber show anti-agreement effects in both subject *wh*-questions and subject relative clauses, it makes sense to assume that Tarafit Berber allows for both *wh*-DP and null operators to enter into an Agree relation with a higher head Probe, and the result of both of these agreement operations is anti-agreement.

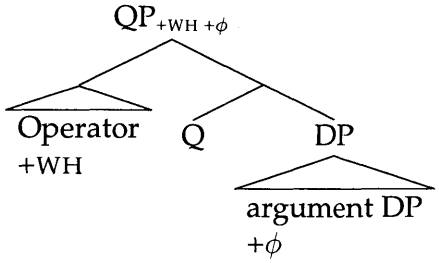
Background assumption: the syntax of Tarafit Berber

Before continuing the discussion of how anti-agreement effects show that spec-to-spec anti-agreement is active in Tarafit Berber, we need to establish some assumptions about the structure and syntax of Tarafit Berber more generally.

First, this chapter will assume, following Cable (2006), that the arguments that undergo *wh*-movement in Tarafit Berber are syntactically complex QPs with the structure given in (39) (similar assumptions were made about the Imbabura Quechua *wh*-system in Section 4.3). In (39), the DP argument is the complement of the Q head, and the null question operator is the specifier of the larger QP. The DP is [+ϕ] and the operator argument in spec-QP is [+WH]. Following theories of feature inheritance (e.g., Richards 2007), the entire QP argument itself is both [+WH, +ϕ]. Following Baier's assumptions about *wh*-agreement, the QP is a possible goal for both [+WH] and [+ϕ] probes[17].

[17]For further details about how agreement interacts with a complex QP and a theory that collapses *wh*-agreement and anti-agreement, see Baier (2016).

(39) Structure of a complex QP



Second, similar to Erlewine's analysis of Kaqchikel Agent Focus and following Kinjo's (2017) analysis Tamazight Berber, this chapter will assume that Tarafit Berber subject agreement relationships are established between the verb and the element that sits in the EPP position, spec-TP[18]. I will assume that, in Tarafit Berber, subject Agreement triggers subject movement. Given that Tarafit Berber anti-agreement (a sub-type of *wh*-agreement) can be triggered in both subject *wh* questions and subject relative clauses, it appears that Agree relations can be established between either an overt DP with a [+WH] feature or an Ā-movement operator with a [+WH] feature, so long as the [+WH]-marked element is able to move to spec-TP. This carries with it the additional assumption that the Tarafit Berber EPP can be satisfied by either an overt DP or a null operator.

Third, this discussion will make the following assumptions about Tarafit Berber clause structure. Following Ouali (2011), this discussion will assume that Tarafit Berber VSO word order is derived by moving the verb to some higher position in the information structure or aspect domain. While this discussion leaves open the exact target position of Tarafit Berber verb movement, this position must crucially be above the subject and below spec-CP (as *wh*-extracted Tarafit Berber subjects still preceded their verbs). However, examples in the remaining sections of this discussion will abstract away from verb movement operations[19].

[18]See Ouali 2011 for an alternative approach, claiming the Berber subjects remain in their thematic position unless extracted to an information structure position in the left periphery, which Baier 2016 adopts

[19]Assuming the verb moves to any position other than C, the verb would have to be analyzed as moving to a position that intervenes between T and C. This analysis would predict that Erlewine's version of spec-to-spec anti-locality should not be active in Tarafit Berber—or, at least, it should not constraint subject *wh*-movement and its relationship to anti-agreement. This is because subject movement from spec-TP to spec-CP would not be anti-local under Erlewine's definition (see Baier 2017) for a discussion of why Tarafit Berber is incompatible with Erlewine's definition of anti-locality). However, as there is no evidence that the Berber position that triggers head-to-head verb movement requires a designated, semantically active specifier, this analysis is still compatible with Brillman & Hirsh's definition of spec-to-spec anti-locality.

Analysis: anti-agreement and anti-locality

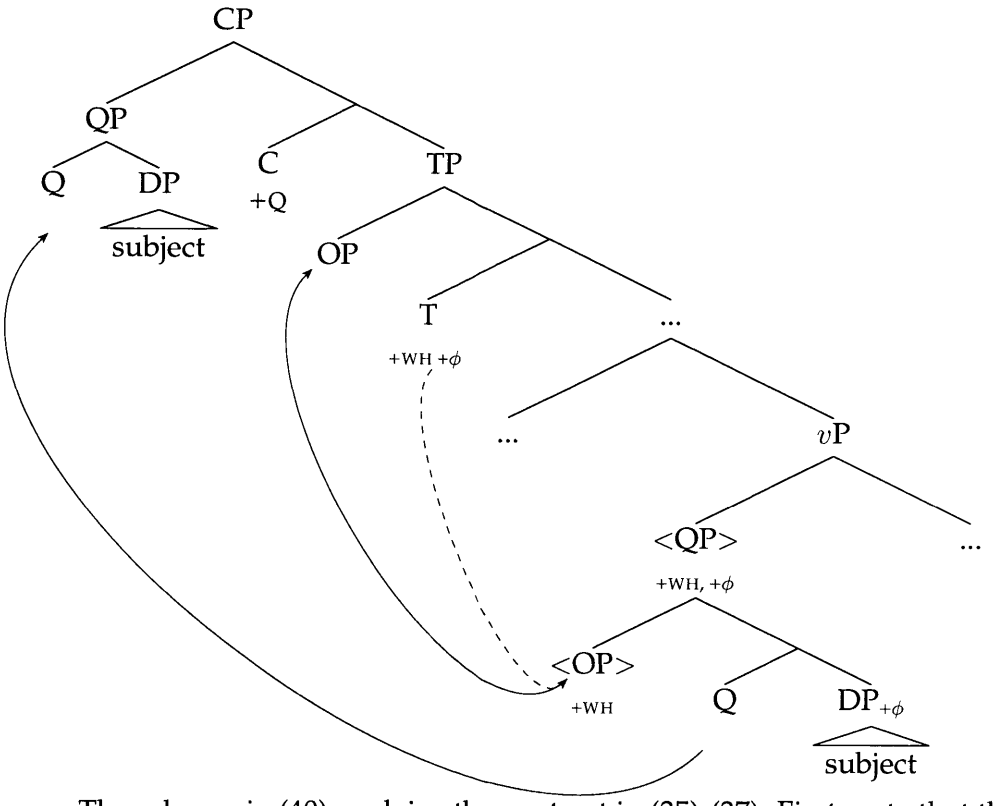
This chapter will analyze Tarafit Berber anti-agreement effects as resulting from the structure in (40). This analysis is supported by the complex QP structure outlined in (39) and the data from relative clauses that suggest Tarafit Berber anti-agreement relationships can be established between an Ā operator in spec-TP and the Tarafit Berber verb. In (40), the head T probes for an element with which it can Agree. Following Baier (2016) and Deal (2015, 2016), I assume that anti-agreement probes can be satisfied by either a head with [+WH] or [ϕ] features. The closest element to the probing T head is the complex QP. Because the T head can be satisfied be either a [+WH] or [ϕ] element, T enters into an Agree relationship with the null operator in spec-QP, the closest available argument. This agreement operation triggers movement of the null operator from spec-QP to spec-TP. This analysis predicts that anti-agreement is actually the expected morphological pattern when a null operator occupies the Tarafit Berber EPP position.

However, in (40), T is not the only syntactic head that probes for agreement; C also probes for an argument with a [+WH] feature. C cannot enter into an Agree relation with (and subsequently trigger the movement of) the [+WH] null operator, because Ā-moving the null operator from spec-TP to spec-CP would violate spec-to-spec anti-locality[20]. However, C is still able to enter into an Agree relationship with the entire [+WH] QP. In (40), then, C enters into an Agreement relationship with the QP, and subsequently triggers the movement of the QP into spec-CP. Because this movement chain crosses the TP projection, it satisfies both Erlewine's and Brillman & Hirsh's definitions of spec-to-spec anti-locality.

The structure in (40) abstracts away from material that intervenes between TP and vP in the tense/aspect domain (c.f., Ouali 2011 for a discussion of the structural projections that might occur in these positions), as well as Tarafit Berber verb movement. Because the verb never enters into a relationship with the [+ϕ] DP argument, ϕ agreement is never realized on the verb.

[20]If the entire QP had moved to spec-TP, movement of the entire QP would similarly be banned.

(40) Tarafit Berber anti-agreement



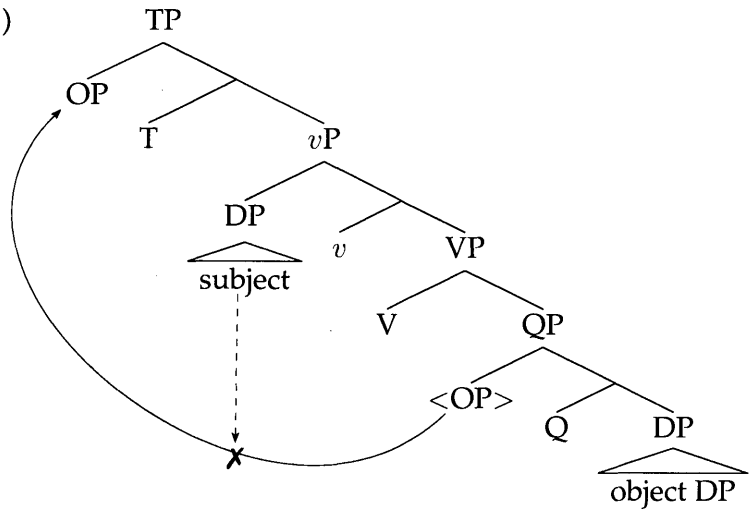
The schema in (40) explains the contrast in (35)-(37). First, note that the movement chains in (40) all respect spec-to-spec anti-locality. The null operator can A-movement from the specifier of the QP to spec-TP because spec-to-spec anti-locality does not restrict A-movement (this movement is licit regardless of whether or not any additional clausal projections sit between TP and vP in Tarafit Berber). The QP moves past TP, with the null operator in spec-TP, on its ways to spec-CP, satisfying both Erlewine's and Brillman & Hirsh's definition of spec-to-spec anti-locality. Thus, the grammaticality of (35b) is predicted. The ungrammaticality of (35c) is also predicted by anti-locality. This is because (35c) lacks operator agreement, and so would require a configuration where the entire QP that is introduced in spec-vP first A-moves to spec-TP (to satisfy the Tarafit Berber EPP) before Ā-moving directly to spec-CP. This movement path would cross no additional maximal projections or specifiers, and so would be banned by both Erlewine's and Brillman & Hirsh's definition of anti-locality.

A possible alternative analysis might involve having the DP subject embedded within the QP sub-extract to spec-TP, after which the QP could move directly to spec-CP from its thematic position. However, this movement should

be banned by relativized minimality, a locality constraint. In order for the DP to A-move to spec-TP, it would need to cross the null operator in spec-QP. Because Tarafit Berber is a language that allows null operators to establish Agree relationships with T and trigger agreement on the verb, we should also assume that null operator arguments are sufficiently feature-complete to cause a relativized minimality violation if the QP argument attempted to A-move past it.

This same logic, I propose, can also explain why (37b)—an object question that co-occurs with anti-agreement morphology—is ungrammatical. If the null operator is phi-complete enough such that another argument cannot move across it without incurring a relativized minimality violation, we also predict that the null operator will not be able to move across any other arguments without incurring the same violation. Assuming that Tarafit Berber object *wh*-questions also involve a complex object QP, and because objects are introduced below subjects in the derivation, there would be no way for the null operator to A-move from spec-QP to spec-TP without first crossing the subject argument, as illustrated in (41). This is the same general process that prevents an object argument from moving to spec-TP to satisfy the EPP.

(41)



In (41), the probe on T will attract the closest argument. Because the subject DP in spec-vP is structurally closer to T than the object QP, there is no way for the null operator to sub-extract to spec-TP from an object QP. Thus, the analysis provided here does not predict that anti-agreement would ever be able to grammatically co-occur with non-subject Ā extraction.

Though Tarafit Berber showcases a pattern that at least superficially appears very similar to Kaqchikel agent focus alternations, this section has not argued that Tarafit Berber anti-agreement and Kaqchikel agent focus alternations reduce to the same effect. Nor has this section proposed that Berber anti-agreement is a direct effect of anti-locality. Instead, this section has shown another way in which spec-to-spec anti-locality interacts with other facets of the grammar. In this case, it has described the interactions of spec-to-spec anti-locality and a complex *wh*-agreement system. In this discussion, spec-to-spec anti-locality has emerged as a constraint on short *wh*-movement in Berber; Baier's *wh*-agreement system, among other things, allows for Ā-movement to proceed in a way that is respectful of spec-to-spec anti-locality.

3.4.5 What anti-locality doesn't predict (a digression on *que/qui*)

This chapter has discussed a wide array of subject/non-subject asymmetries that anti-locality can potentially capture, ranging from resumptive pronoun alternations in Hebrew to anti-agreement effects in Berber. Additionally, this chapter has also stated that it will be able to show that anti-locality can capture a similar range of subject/non-subject effects in English, such as a lack of

do-support in matrix subject *wh*-questions. Before moving on to the English data, however, it's worth discussing the (type of) alternations that anti-locality is not predicted to (fully) capture or explain.

As previously stated, anti-locality is not proposed to be able to explain any subject/non-subject asymmetry; it's proposed to be able to explain, at least partially, any subject/non-subject asymmetry *that arises as a result of a subject needing to undergo short Ā-movement from spec-TP to an adjacent spec-CP*. Perhaps the most well-known subject/non-subject asymmetry that anti-locality is not predicted to fully explain is the *que/qui* alternation in French (see Section 2.1).

The first half of the (relatively famous) *que/qui* alternation can be summarized as follows: in standard embedded clause constructions, like (42a), French utilizes the complementizer *que*. When a French embedded object undergoes Ā-movement to the matrix clause, use of the complementizer *que* is still licit (42b). However, when the embedded subject is Ā extracted, *que* is ungrammatical (42c). From the example in (42), the French *que/qui* effect looks like a direct correlate of the English *that*-trace effect (Section 5.1) in a language where the overt complementizer cannot be omitted from the embedded clause.

(42) French subjects cannot extract across ***que***

a. J'ai dit **que** j'aim les pommes
I-AUX said *que* I.like the apples
'I said that I like apples.'
(***que** used in baseline case*)

b. Qui$_i$ tu crois [t$_i$ **que** Jean a vu t$_i$]
Who$_i$ you believe [t$_i$ **que** Jean has seen t$_i$]
'Who do you believe that Jean has seen?'
(***que** used with object extraction*)

c. *Qui$_i$ tu crois [t$_i$ **que** t$_i$ a vu Jean]
Who$_i$ you believe [t$_i$ **que** t$_i$ a vu Jean]
INTENDED: 'Who do you believe has seen Jean?'
(***que** impossible with subject extraction*)

This half of the *que/qui* alternation can be easily explained by spec-to-spec anti-locality. The subject cannot extract across *que* (42c) because such movement would be anti-local—the movement chain would require that the subject move directly from spec-TP to spec-CP across no intervening material, resulting in an spec-to-spec anti-locality violation (under both Erlewine's and Brillman & Hirsh's definition). A parallel explanation has has been used to explain why embedded subjects cannot undergo *wh*-movement in Quechua, why a resumptive pronoun

asymmetry exists in Hebrew, and why Berber subject *wh* questions cannot co-occur with canonical subject agreement. Similarly, this same analysis will also be proposed in Section 5.1 as a way of understanding the English *that*-trace effect.

However, the *que/qui* alternation goes beyond the examples in (42). French allows for the interrogation of embedded subjects, but French also requires that a special morpheme, *qui*, (sometimes argued to be a special form of the complementizer) be used in embedded subject-extraction contexts (43a). Currently, the problem still looks like it might be entirely explained by spec-to-spec anti-locality: the introduction of the special complementizer *qui* could correspond with a change in syntactic structure, one that obviates the anti-local structure that restricts subject movement in (42c). Potentially, *qui* could be the result of the complementizer *que* fusing after head movement with some additional, semantically active element. For Erlewine, this could be any element, so long as it is originally sat in a maximal projection between TP and CP. For Brillman & Hirsh, this element would have to be either the head of a maximal projection between TP and CP that projects a (possibly silent but) semantically active specifier, or some element from that specifier position itself. This would result in a larger clausal spine and A-movement chain long enough to respect anti-locality. In their Criterial Freezing (2007) paper, Rizzi & Shlonsky propose that this style of analysis—that *qui* is the realization of head movement to *que* from some additional layer in the derivation—accounts for the subset of the *que/qui* facts shown in (42). Alternatively, *qui* could correspond with a deeply shrunken embedded clause structure. For example, *qui* could be a realization of a bare TP, indicating an embedded clause structure small enough that the subject could move directly out of the embedded spec-TP and into the matrix clause, without having to move through some intermediate position like spec-CP. Neither of these avenues for analysis are unexpected or unpredicted, given spec-to-spec anti-locality.

The final piece of the *que/qui* alternation, seen in (43b), however, is not obviously connected to spec-to-spec anti-locality or anticipated by its predictions. As (43b) shows, object extraction is impossible across the *qui*, the morpheme that occurs when subject extraction is possible.

(43) (only) French subjects can extract across ***qui***

a. Qui$_i$ tu crois [**qui** t$_i$ a vu Jean]
Who$_i$ you believe [**qui** t$_i$ has seen Jean]
'Who do you believe has seen Jean?'

(compare with (42c))

b. *Qui$_i$ tu crois [**qui** Jean a vu t$_i$]
Who$_i$ you believe [**qui** Jean has seen t$_i$]
INTENDED: 'Who do you believe Jean has seen?'

*(object extraction not possible across **qui**)*

It's not obvious that spec-to-spec anti-locality should be able to explain why an object cannot undergo what looks like long distance movement. Anti-locality deals with explaining the relationship between nodes in a tree that are structurally near to each other. The relationship between the embedded object and the embedded complementizer is too distant to have any obvious connection with any provided definition of spec-to-spec anti-locality.

It's off course possible—perhaps, likely—that there might be a connection between an anti-locality analysis of the restriction of subject movement across *que* and the fact that object extraction is not permitted across *qui*. For instance, if a Rizzi & Shlonsky head-fusion approach is adopted, it is possible that the head that merges with *que* to form *qui* has some additional property that might limit object extraction, but that property would not directly follow from spec-to-spec anti-locality. Alternatively, a restriction on object extraction could be connected to an analysis of *qui* as a signifier of a base TP phase. For example, one possible analysis would be that *qui* signifies a bare TP phase structure that does not support multiple specifiers. If, in *qui* clauses, the subject sits in the specifier of a TP headed by *qui*, then the object would not be able to extract from within the embedded clause because of restrictions on successive-cyclic Ā-movement. There is no available phase edge position for the embedded object to move through (as the only available phase edge position is occupied by the embedded subject), and so object extraction is ruled out entirely. However, this rough analysis is based entirely on how outside principles—in this case, the particular properties of a proposed realization of T—happen to interact with anti-locality, not on the immediate predictions of anti-locality itself.

The purpose of this chapter is to show that a spec-to-spec anti-locality constraint is active in the grammar. The purpose of this discussion is not to claim that the *que/qui* alternation is completely unconnected to anti-locality, but rather to show that spec-to-spec anti-locality alone cannot capture and explain the full

range of the *que/qui* alternation facts. Expanding on the suggestions in the preceding paragraphs—taking this discussion from from a potential analysis and route for exploration to a well formed hypothesis with testable predictions—would involve careful investigation into the French TP and CP domains, and this chapter leaves this question open for future research. This section is meant as a concrete illustration of one of the types of subject/non-subject asymmetries that anti-locality is not totally suited to capture in isolation, as well as an illustration of how spec-to-spec anti-locality can be proposed to interact with other mechanisms in the syntax.

3.4.6 Interim summary

Finally, this section has explored (some of) the subject/non-subject asymmetries that spec-to-spec anti-locality is predicted to capture. Spec-to-spec anti-locality is particularly likely to explain subject asymmetries where embedded subject Ā extraction is banned in certain contexts but not others, such as the English *that*-trace effect. Anti-locality is not necessarily predicted to be able to tell the entire story when object Ā-movement is banned in certain contexts but subject Ā-movement is not. In these cases, such as the French *que/qui* alternation, spec-to-spec anti-locality might still be able to capture some of the story, but it is not predicted to be able to explain the entirety of the alternation.

3.5 Anti-locality Violations and Neutralizations: the English perspective

Previous sections of this chapter have shown contexts where anti-locality prohibits a certain movement operation, e.g., subject *wh*-movement in Imbabura Quechua. While this chapter has previously discussed strategies that a language can take to avoid incurring an anti-locality violation (e.g., pied-piping an embedded clause), it has only given scattered examples of situations where anti-locality can obviated (e.g., moving an operator to an agreement position, facilitating a longer movement chain). This section will focus on examples where not only does spec-to-spec anti-locality restrict subject Ā-movement, but where there is clear evidence that eliminating the anti-local configuration allows subject Ā-movement to proceed unrestricted. This section, which draws very heavily on Brillman & Hirsh (2016) will take three English subject/non-subject

asymmetries—the *that*-trace effect, the relationship between *tough* constructions and gapped degree phrases, and subject *wh*-constructions—and show the following: (i) for each asymmetry, anti-locality is sufficient to predict that the asymmetry exists and (ii) eliminating the anti-local movement configuration that underlies the asymmetry erases the asymmetry itself. This section not only gives evidence for the existence of a general spec-to-spec anti-locality constraint in the grammar, but shows how eliminating anti-local configurations from a construction also results in eliminating restrictions on subject Ā-movement.

3.5.1 *That*-trace effects

Our case study of English subject/non-subject asymmetries begins with the *that*-trace effect, repeated in (44) from (2):

(44) *That*-trace effects:

a.	*Who does Bill think that _ saw John?	(subject extracted, ✗)
b.	Who does Bill think that John saw _?	(object extracted, ✓)
c.	Who does Bill think _ saw John?	(subject extracted, ✓)
d.	Who does Bill think John saw _?	(object extracted, ✓

Analyzing the *that*-trace effect involves answering two questions. First, why are subjects prohibited from Ā extracting from an embedded clause across an overt complementizer *that* (44a-b). Second, why is subject Ā extraction not similarly restricted when the sentence lacks an overt complementizer (44c-d). This section will discuss these questions in this order, first discussing the role of anti-locality in explaining the ungrammaticality of (44a) before moving to the role of anti-locality in accounting for the grammaticality of (44c).

This section will largely present the *that*-trace effect as a *wh*-effect, as in (44). However, its worth noting that the *that*-trace effect is found across the full range of English Ā constructions, as shown in (45)-(46), modified from Pesetsky (2017), citing Bresnan (1977).

(45) No subject Ā-extraction across *that*

a. *This is the person who I thought that _ criticized Jonathan Franzen. (relativization)
b. *Anneke we think that _ criticized Jonathan Franzen. (topicalization)
c. *It was Anneke that we think that _ criticized Jonathan Franzen. (cleft)
d. *More people hate Jonathan Franzen than we think that _ love Patricia Lockwood. (comparatives)

(46) Subject Ā-extraction possible without overt *that*

a. This is the person who I thought _ criticized Jonathan Franzen. (relativization)
b. Anneke we think _ criticized Jonathan Franzen. (topicalization)
c. It was Anneke that we think _ criticized Jonathan Franzen. (cleft)
d. More people hate Jonathan Franzen than we think _ love Patricia Lockwood. (comparatives)

While the analysis in this section will be given in terms of the *wh*-movement context shown in (44), it should be able to transfer to each of the parallel Ā extraction contexts in (45)-(46).

An anti-locality account

First, focusing on the impossibility of subject extraction in (44a): The contrast between (44a) and (44b)—where subject Ā extraction is banned but object Ā extraction is permitted—is predicted if spec-to-spec anti-locality is operative in English. Assuming that the movement in all sentences in (46) is successive cyclic, the derivation for (44a) should look like (47):

(47) *Who does Bill think [$_{CP}$ <who> that [$_{TP}$ <who> saw John]]

In (47), the subject moves from spec-TP in the embedded clause to spec-CP in the embedded clause, before proceeding to spec-CP in the matrix clause. The first movement step from spec-TP to the local spec-CP violates both Erlewine's and Brillman & Hirsh's definitions of anti-locality (because the movement chain crosses neither an XP or any additional specifiers). Thus (44a) is rendered ungrammatical. The derivation for (44b), on the other hand, is given in (48):

(48) Who does Bill think [$_{CP}$ <who> that [$_{TP}$ John [$_{vP}$ saw <who>]]]

Here, the object moves to the local spec-CP from lower in the clause in (48). This movement path crosses more than one maximal projections (vP and TP) as well as at least one specifier (spec-vP, spec-TP), thus both Erlewine's and Brillman & Hirsh's definitions of anti-locality are respected, and so (48)/(44b) is predicted to be grammatical.

Neutralization

The contrast between subject extraction and non-subject extraction in (44) is consistent with anti-locality. However, to argue that the ungrammaticality of (44a) is in fact an anti-locality effect, we need to show that subjects *can* grammatically extract when their movement path does not violate anti-locality. We discuss three cases which confirm this prediction.

Case 1: An intervening projection (anti-*that*-trace effects)

The first method of obviating anti-locality is inserting an XP (with a specifier) into the structure intervening between TP and CP. There is evidence that obviating anti-locality in this way enables subject extraction when otherwise a *that*-trace effect would be observed. This comes from the so-called anti-*that*-trace effect. As observed in Bresnan Bresnan (1977) and Culicover (1993), (15b) with *for all intents and purposes* is improved over the baseline example in (15a).

(49) a. *Who does John think [$_{CP}$ _ that [$_{TP}$ _ served as president?]]
b. Who does John think [$_{CP}$ _ that [$_{AdvP}$ **for all intents and purposes** [$_{TP}$ _ served as president?]]]

Note that the inclusion of an intervening adverb also improves the parallel non-*wh* *that*-trace examples given in (45), as shown in (50).

(50) Adverb obviation with non-*wh* *that*-trace effects

a. This is the person who I thought that for all intents and purposes _ criticized Jonathan Franzen. (relativization)

b. Anneke we think that for all intents and purposes _ criticized Jonathan Franzen. (topicalization)

c. It was Anneke that we think that for all intents and purposes _ criticized Jonathan Franzen. (cleft)

d. ?More people hate Jonathan Franzen than we think that for all intents and purposes _ love Patricia Lockwood. (comparatives)

We assume that *for all intents and purposes* is the specifier of an AdvP intervening between the embedded TP and the local CP. The first step of subject movement from spec-TP to spec-CP in (50b) thus crosses TP and AdvP. Because this movement path crosses the AdvP, it is not anti-local by Erlewine's definition. Additionally, following a cartographic model where the phonological adverb itself is the specifier of a silent AdvP, this movement path is also permitted by Brillman & Hirsh's definition of anti-locality. The grammaticality of (50b) parallels Erlewine's analysis of the neutralization of the Kaqchikel agreement asymmetry in Section 3.2, as well as the Berber anti-agreement facts discussed in Section 4.4[21].

Crucially, our analysis predicts that not every adverbial linearly intervening between *that* and the subsequent verb should obviate *that*-trace effects – only an adverbial like *for all intents and purposes* that structurally intervenes between TP and CP will disrupt an anti-local configuration. The contrast between (51b) and (51c) is consistent with this prediction:

(51) a. *Who did John say [$_{CP}$ that [$_{TP}$ _ ran to the store]]?

b. Who did John say [$_{CP}$ _ that [$_{AdvP}$ **fortunately** [$_{TP}$ _ ran to the store]]?

c. *Who did John say [$_{CP}$ _ that [$_{TP}$ _ [$_{AdvP}$ **quickly** [$_{vP}$ ran to the store]]]]?

This contrast is also seen in non-*wh* examples of the *that*-trace effect, as shown in (52) for the relativization and cleft examples.

[21]Examples like (50) have been taken in Kandybowicz (2006) as evidence for a prosodic account of *that*-trace effects: a PF constraint prohibits a sequence of C and adjacent gap at the left edge of a prosodic phrase. The structural analysis presented here is also viable (and, we believe, can capture the rest of Kandybowicz's data, as well).

(52) a. This is the person who I thought [$_{CP}$ that [$_{AdvP}$ **fortunately** [$_{TP}$ _ criticized Jonathan Franzen]]]
b. *This is the person who I thought [$_{CP}$ that [$_{TP}$ _ [$_{AdvP}$ **quickly** criticized Jonathan Franzen]]]
c. Anneke we think [$_{CP}$ that [$_{AdvP}$ **fortunately** [$_{TP}$ _ criticized Jonathan Franzen]]]
d. *Anneke we think [$_{CP}$ that [$_{TP}$ _ [$_{AdvP}$ **quickly** criticized Jonathan Franzen]]]

Fortunately and *quickly* occur linearly in the same position in the pronounced string, and are of comparable prosodic weight. *Fortunately*, as a speaker-oriented adverb, can attach high in the structure, plausibly between TP and CP. *Quickly*, as a manner adverb, occurs lower in the structure, just above *v*P. The *that*-trace effect in (52a) obviates in (52b) with *fortunately*, but not in (52c) with *quickly*.

Case 2: Movement of the subject from lower in the clause

The second way of obviating anti-locality is to move the subject to spec-CP from a position in the clause below spec-TP. This occurs in (53b), which is grammatical, contrasting with the baseline example in (53a):

(53) a. *How many horses does John think [$_{CP}$ _ that [$_{TP}$ _ are in the barn?]]
b. How many horses does John think [$_{CP}$ _ that [$_{TP}$ there [$_{PredP}$ are _ in the barn?]]]

Because the expletive *there* is in the embedded spec-TP in (53b), the EPP is satisfied without the subject needing to move through this position. The subject thus moves to the local spec-CP directly from a lower position in the clause, such as spec-*v*P. This longer movement chain respects anti-locality. The way anti-locality is obviated in (53b) parallels Erlewine's analysis of the Kaqchikel Agent Focus effects discussed in Section 3.2 and the Tarafit Berber anti-agreement analysis discussed in Section 4.4[22].

Case 3: No (movement through embedded) spec-CP

As noted in the introduction to this section, one mystery of the *that*-trace effect is that subject extraction in examples like (44c), repeated in (54) and which lack a phonologically overt complementizer *that*, are completely grammatical:

[22]This fact was also part of Rizzi's (1990) account for why *that*-trace effects are not observed in Italian. Essentially, this proposal states that Italian subjects can move to spec-CP directly from a position within the verbal domain, though, Rizzi's proposal was based on the Empty Category Principle, rather than anti-locality.

(54) Who does Bill think _ saw John?

To account for the contrast between (44a) and (54), we suggest that movement of the subject in (54) can follow a different path than movement of the subject in (44a) – a path which respects anti-locality. Rather than *who* moving successive cyclically through the local spec-CP as shown in (48) for (44a), *who* in (54) can move to the matrix clause directly from spec-TP in the embedded clause. The movement path for (54) is (55):

(55) [$_{CP}$ Who does Bill think [$_{TP}$ <who> saw John]]

We propose that the tensed clause selected as a complement by a clausal embedding verb need not be a full CP. This is reflected in the structure proposed in (55), where *think* embeds a bare TP. Because the embedded clause in (55) does not have a CP layer, there is no short movement from spec-TP to spec-CP within the embedded clause. Anti-locality is thus respected, and the grammaticality of (55) is predicted.

There are two independently proposed theories that can explain how the presence of *that* in (44a) versus the absence of *that* (44c)/(55) can interact with the path of movement. First, recall the contextually specified theory of phasehood discussed in the introduction to this dissertation and Section 3.3 of this chapter. Under this theory, Wurmbrand states that embedded clauses without an overt *that* can be, but do not necessarily *need* to be, CPs. Under this theory, these embedded clauses can be as small as bare TPs. Additionally, this theory proposes that the highest maximal projection in the cyclic domain is, definitionally, the phase edge. By this account, TP is the phase edge in (55). In (55), when the embedded subject A-moves to spec-TP to satisfy the EPP requirement on the embedded clause, it has also undergone A-movement to the phase edge. Because the subject is already at the phase edge, the subject can move directly into the matrix clause without causing a PIC violation or having to undergo any additional successive-cyclic movement.

Additionally, Erlewine (2017), building on Fox and Pesetsky's (2006) system of cyclic linearization, proposes an alternative way of getting the grammaticality of (54). Erlewine proposes that the example in (55) really has the structure in (56), which is distinct from the structure presented in (55).

(56) $[_{CP}$ Who does Bill think $[_{CP}$ $[_{TP}$ <who> saw John]]]

According to Erlewine, as illustrated in (56), movement through spec-CP is not required when C lacks overt exponence. In (44a), *who* must move to the local spec-CP in order to be linearized to the left of *that* in C when the embedded clause is spelled-out. Because linearization requirements are imposed at PF, C is irrelevant for linearization when it lacks overt exponence, and successive cyclic movement is not similarly required in (54).

Conclusion

This subsection has shown that the *that*-trace effect aligns with the predictions of spec-to-spec anti-locality. Specifically, the presence of *that* in the embedded clause signifies that the embedded clause has a structure such that the subject would be required to undergo anti-local movement from spec-TP directly to spec-CP in order to extract from the embedded clause. Such movement is prohibited by anti-locality, and so Ā of an embedded subject across a phonologically overt *that* is banned. However, the absence of *that* indicates that such movement is not necessary. When *that* is not present, the embedded clause is compatible with a structure that allows the embedded subject to Ā directly from the embedded spec-TP position into the matrix clause. Additionally, the *that*-trace effect asymmetry is eliminated when the subject can Ā-move into the embedded spec-CP from lower in the derivation, or across an intervening adverb.

3.5.2 *Tough*-constructions and gapped degree phrases

There is a well-known restriction on Ā-movement from the subject position of infinitives, demonstrated in (57a), which contrasts with fully grammatical object extraction in (57b).

(57) a. *Who is it possible _ to see Mary? (subject extraction)
b. Who is it possible for Mary to see _? (object extracton)

Since Chomsky (1981), it is common to root the ungrammaticality of (57a) in case theory: because the subject of an infinitive is not a case position and *who* does not move to any other case position, in (57a) *who* cannot be valued for case and (57a) is ungrammatical.

This chapter argues that a case-based analysis cannot account for at least some cases of ungrammaticality involving the extraction of infinitival subjects, and that independent effects of anti-locality can be isolated. One such case is (58), a minimal pair to (57a), with a case-assigning complementizer *for*:

(58) *Who is it possible for _ to see Mary?

Who can be valued for case in the embedded spec-TP in (58), so (58) cannot be ruled out for case reasons. Rather, we argue that the ungrammaticality of (58) has the same source as the *that*-trace effect: its derivation involves a first step of movement from the embedded spec-TP to the local spec-CP, and that this movement step violates anti-locality ruling out the example[23].

(59) *Who is it possible [$_{CP}$ _ for [$_{TP}$ _ to see Mary]]

Evidence for a spec-to-spec anti-locality analysis of the ban on subject *tough* movement comes from comparing *tough* constructions, which do not permit subject extraction (60), to gapped degree phrases, which do allow for subject extraction (60).

(60) a. *Anneke is tough _ to talk to Ian . (subject gap TC)
b. Ian is tough for Anneke to talk to _. (object gap TC)

(61) a. Anneke is too smart _ to talk to Ian (subject GDP)
b. Ian is too smart for Anneke to talk to _. (object GDP)

Background on *tough* constructions and gapped degree phrases

First, a note on *tough*-movement. Chapter 2, which provided a syntax for gapped degree phrases, adopted an *improper movement* approach to *tough*-constructions, where an overt argument from the embedded clause itself undergoes movement to the matrix subject position. As noted in Chapter 2, this is not the only available view of the syntax of *tough*-constructions. Another approach proposes that a null operator is Ā-moved from spec-TP to spec-CP in the clause embedded by *tough* (62), and in the other, the surface matrix subject itself undergoes improper movement from the embedded subject position. The examples in (62) and (63)

[23]Unlike with *that*-trace effects, it is not possible to show obviation with insertion of a maximal projection intervening between TP and CP; *for*, as a case assigner to spec-TP, cannot be separated from TP by any intervening projection.

schematize what a subject *tough*-movement chain would need to look like, in the improper movement (62) and null operator (63) analyses, respectively.

(62) *He is tough [$_{CP}$ t [$_{TP}$ t to see Ian]]

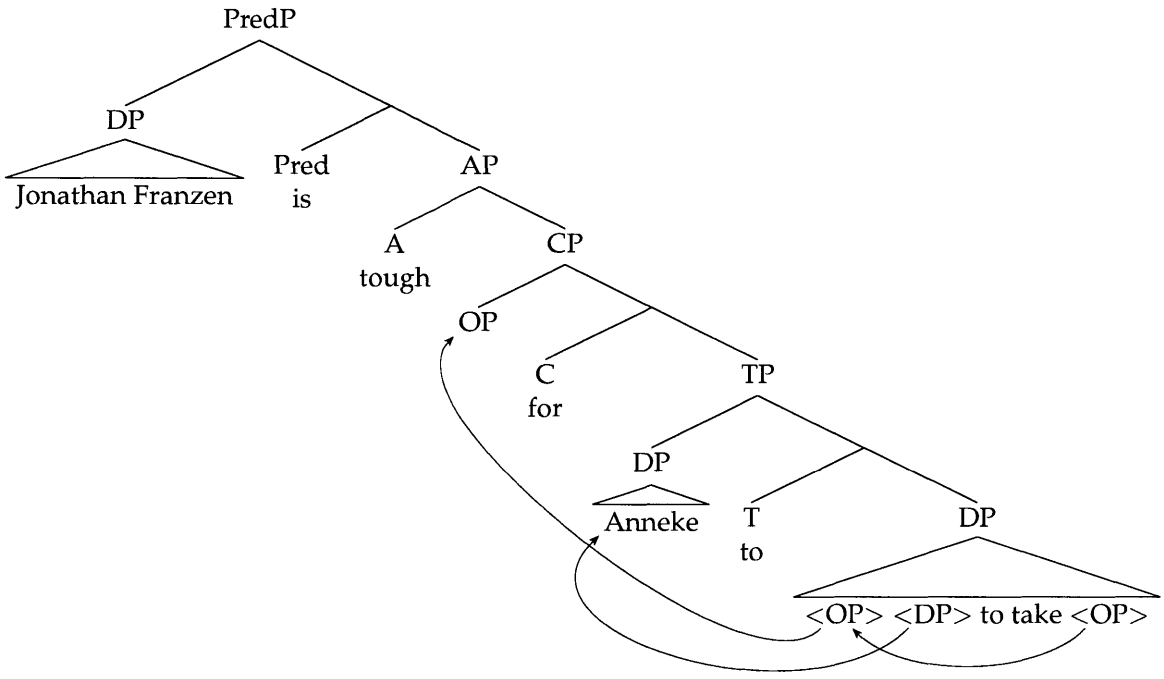
(63) *He is tough [$_{CP}$ OP [$_{TP}$ t to see Ian]]

This chapter will remain agnostic towards a particular analysis of *tough*-constructions (though see Chapter 4 for an acquisition argument that is most consistent with an improper movement approach to *tough*-constructions); figures throughout this chapter, however, will often depict a structure that is most compatible with a null operator analysis of *tough*-movement. This is done to make comparisons between *tough*-constructions and gapped degree phrases (a null operator constructions) as clear as possible. It's worth noting, though, as seen in (63)-(62), that a null operator and improper movement analysis of *tough*-movement are prediction neutral for the anti-locality analysis presented here: both possible movement chains would require movement from the embedded spec-TP to the embedded spec-CP across no intervening material, an anti-local movement step.

Moving, now, to a more direct comparison between *tough* constructions and gapped degree phrases. Gapped degree phrases look structurally "larger" than *tough*-constructions—at the very least, gapped degree phrases contain a degree word beneath their adjectival predicate that *tough*-constructions lack. These two constructions, and the relationship between them, was the subject of the second chapter of this dissertation. An important conclusion of that chapter was the proposal that *tough*-constructions, which are represented with a tree structure in (64), involve Ā-moving an element[24] to the edge of their embedded clause, the CP, while gapped degree phrases have the structure given in (65).

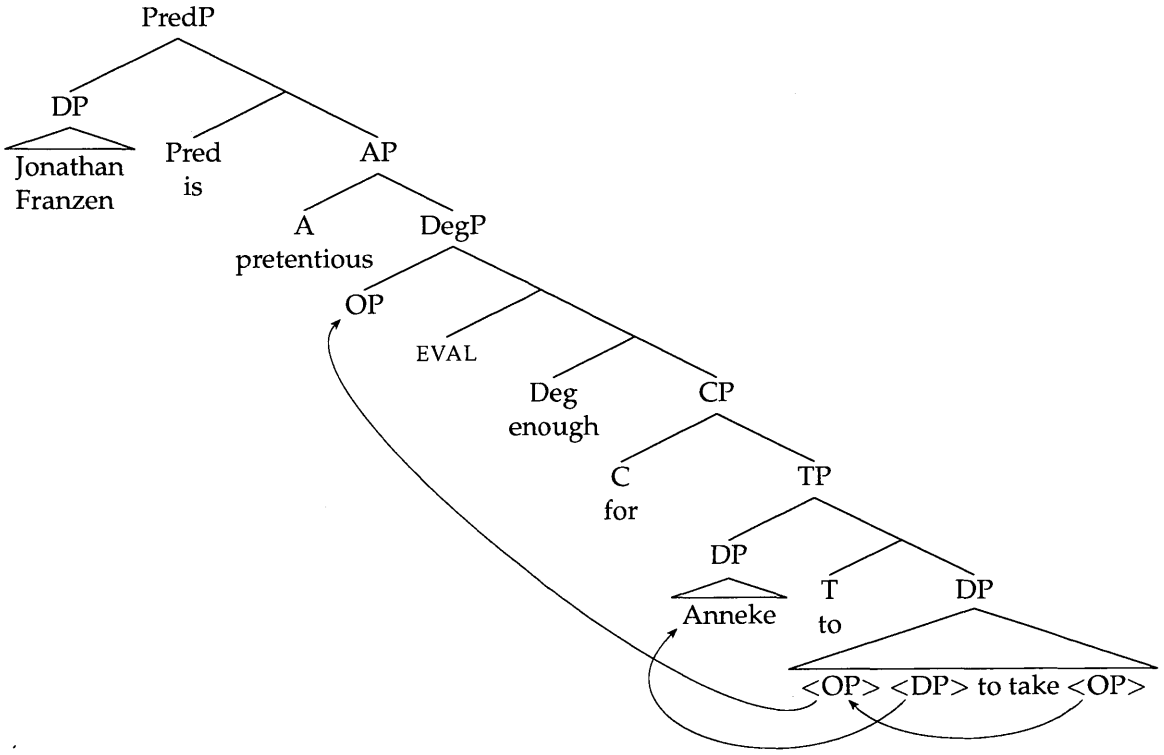
[24]In (i), this element is represented as a null operator. The predictions of the arguments in this chapter would not change if the element in (i) was represented as a phonologically overt DP.

(64) ***Tough*-movement schema**
"Jonathan Franzen is tough for Anneke to read."



In (64), the embedded object, a null operator, Ā-moves successive-cyclically to the edge of the embedded CP. Under both an improper movement and a null operator analysis of *tough* movement, this movement is syntactically and semantically necessary. Under an improper movement analysis of *tough* constructions, an overt DP (*"Jonathan Franzen"*, the parallel example to (64)) would A-move directly from the embedded clause spec-CP to spec-PredP, the subject position of the embedded clause. This movement would not be possible if the DP did not first move to the phase edge. The presence of an overt *for* infinitival complementizer indicates that the embedded clause is a CP structure and so the phase edge of the embedded clause must be spec-CP, even under a contextually-specified theory of phasehood. Under a null operator theory of *tough*-movement, the embedded object operator must move to the edge of the CP phase in order to eligible for semantic binding with the matrix subject, the overt DP *Jonathan Franzen*. This ensures the semantic identity between the interpretation of the null operator and the matrix subject.

(65) **Gapped degree movement structure**
"Jonathan Franzen is pretentious enough for Anneke to dislike."



The structure in (65) can be understood as follows: the embedded object is a null operator, which must move successive-cyclically to the edge of the embedded clause. In (65), however, the embedded clause does not end at the CP layer. Rather, the embedded clause extends to the DegP that it embedded directly under the adjectival predicate (see Chapter 2 for additional evidence supporting this claim). Because the embedded clause in (65) extends to the DegP—above the CP—and itself meets the definition of a cyclic domain, we can understand Deg as the phase head for the embedded clause. The highest spec-DP, then, would constitute the phase edge, not spec-CP. Within the embedded clause, the null operator can move directly from spec-vP to the highest spec-DegP. Like with the null operator analysis of the *tough*-construction schema in (64), the semantics require the null operator to move to the edge of the embedded clause. Null operators are, by themselves, semantically vacuous and can only be interpreted after been bound to an overt argument that c-commands them in the derivation. Within the literature, there are multiple proposals for how null operators can bind to their antecedents. This chapter will assume, following Chapter 2, that—at least in the context of gapped degree phrases—this semantic binding happens via a

semantic operation called COMPOSE (see Nissenbaum & Schwarz 2011 for the details of the original proposal, and Chapter 2, Section 2 for a brief summary). The precise mechanics of COMPOSE are complex and beyond the scope of this discussion. However, its requirements are both simple and important: COMPOSE can only happen if the null operator has Ā-moved to the edge of the embedded clause (within the movement, there will be a type mismatch between the embedded clause and the adjectival predicate that embeds it, and the derivation will fail at LF). Once the null operator has moved to this position, COMPOSE is able to bind it to the matrix subject. After COMPOSE, the interpretation of the null operator and all its copies are co-referential with the matrix subject.

The structure in (65) also shows a multiple specifier analysis of the DegP—the highest specifier is occupied by the null operator, the lowest specifier is occupied by a phonologically vacuous but semantically active judge/evaluator argument. Under Erlewine's definition of spec-to-spec anti-locality, the presence or absence of this DegP-internal evaluator argument is prediction neutral for an analysis of subject movement within gapped degree phrases. Under Brillman & Hirsh's definition of spec-to-spec anti-locality the presence of this specifier is crucial. Chapter 2 of this dissertation presented a semantic and syntactic argument for the presence of an evaluator argument internal to the DegP in cases like (65). Essentially, all degree constructions involve evaluating a CP threshold relative to some standard (Lasnik & Fiengo 1974). This evaluation is inherently modal (Meier 2003), meaning that it made relative to a set of possible worlds. In some cases, like (66a), the standard is objective, and so this evaluation is done relative to the rules of the natural world(s) in which the sentence occur(s). In cases like (66b), this evaluation is not as straightforward. This is because (66b) is not actually a statement about Anneke's literary preferences; it's a statement about someone's evaluation of Anneke's literary preferences. The existence of this evaluator is necessary for understanding the truth conditions and interpretation of (66b).

(66) a. That Jonathan Franzen novels is not thick enough to use _ as a doorstop.
b. That Jonathan Franzen novel is not interesting enough for Anneke to read _.

In (66b), the threshold is consistent across possible worlds. In addition to shifting with Anneke's preferences, it also shifts depending on the perspective of the evaluator, and how their belief worlds surrounding Anneke and her literary taste.

In fact, the evaluator's belief worlds can be formalized as the set of possible worlds within which the CP standard and degree threshold can be compared. Chapter 2 of this dissertation provides a more elaborated discussion of the inclusion of evaluator arguments in subjective degree phrases.

Comparing *tough*-constructions and gapped degree phrases

The literature has hypothesized that *tough*-constructions and gapped degree phrases are related since Lasnik & Fiengo (1974), and chapter 2 of this dissertation proposed a new way of understanding the relationship between the two constructions. This section will zoom in and focus on one similarity between the two constructions, the Ā-movement chain that both structures contain within their embedded clauses. Crucially, this section proposes that subject cannot *tough*-move because they are unable to Ā-move from spec-TP to spec-CP without violating spec-to-spec anti-locality (cf. (3a)-(64)). Gapped degree phrases, however, have an embedded clause that it one projection "larger" than *tough*-constructions: it consists of a DegP that selects a CP as its complement, not a bare CP. Because, both subjects and objects can move to the embedded clause edge in gapped degree phrases. This is illustrated in (67)-(68). The example in (67), like (65), shows object movement across TP, *v*P, and the specifiers of those projections. The example in (68) shows subject movement across DegP and its lowest specifier, a silent evaluator argument. The presence of this evaluator is crucial for an argument that relies on Brillman & Hirsh's formulation of spec-to-spec anti-locality, and prediction neutral for an analysis that relies on Erlewine's formulation.

(67) Anneke is [$_{AP}$ savvy [$_{DegP}$ OP EVAL enough [$_{CP}$ for [$_{TP}$ Ian to admire <OP>]]]]

(68) Ian is [$_{AP}$ savvy [$_{DegP}$ OP EVAL enough [$_{CP}$ [$_{TP}$ <OP> to admire Anneke]]]]

Chapter 2 provided extensive evidence to support the claim that *object* gapped degree phrases are movement (not control) constructions—one such argument was the fact that object gapped degree phrases licensed parasitic gaps. Similarly, a source of support for the claim that subject gapped degree phrases are also movement (rather than control) constructions is their ability to license parasitic gaps, as in (69a). Though somewhat degraded, (69a) is clearly improved over the example with a control construction in (69b), which is as deviant as the baseline

example of an unlicensed parasitic gap in (69c).

(69) a. ?This student is too young _ to take the bar exam [without us talking to *pg*]
b. *This student is eager PRO to take the bar exam [without us talking to *pg*]
c. *This student took the bar exam [without us talking to *pg*]

These facts help us understand the ungrammaticality of subject *tough*-constructions. The traditional approach to account for the ungrammaticality of subject *tough*-constructions like (60a) has again been case-based. In particular, it has been based on a restriction on Ā-movement which states that Ā-movement can only occur from a case position[25]. This cannot, however, account for the contrast between (60a) and (61a). The structures in (63) or (3a) for (60a) and (68) for (61a) both involve Ā-movement from a non-case position, so both should be ruled out if Ā-movement must originate in a case position.

We argue that spec-to-spec anti-locality provides a more straightforward analysis of the subject movement asymmetry between *tough*-constructions and gapped degree phrases. Both (63) and (3a) involve short movement from spec-TP to spec-CP, in (63) of the null operator and (3a) of *Anneke*, so are ruled out by anti-locality. Because movement in (68) goes from spec-TP directly to spec-DegP skipping spec-CP, movement in (68) crosses TP and CP, so respects anti-locality. Because gapped degree phrases have embedded clauses that include an additional DegP layer, compared to the embedded clauses in *tough*-constructions, anti-locality predicts the possibility of subject gapped degree phrases while prohibiting subject *tough*-constructions.

Combining degree words and *tough* predicates

A particularly clear examples of how spec-to-spec anti-locality interacts with the relationship *tough*-constructions and gapped degree phrases is the interpretation of sentences which combine a *tough*-predicate and a degree word, such as (70). The sentence in (70) is ambiguous between a *tough*-construction (where *Anneke* does not receive a Θ-role from *tough*, (70a)) and gapped degree phrase (where *Anneke* does receive a Θ-role from *tough*, (70b)) interpretation.

[25]Under an Erlewine (2017) analysis where successive cyclic movement is motivated by linearization considerations, it is not clear what would motivate movement to the intermediate spec-CP in (3a), which lacks a phonological complementizer. This perhaps suggests that a Doherty (1996) style analysis, where the size of the embedded clause where may better explain examples like (44).

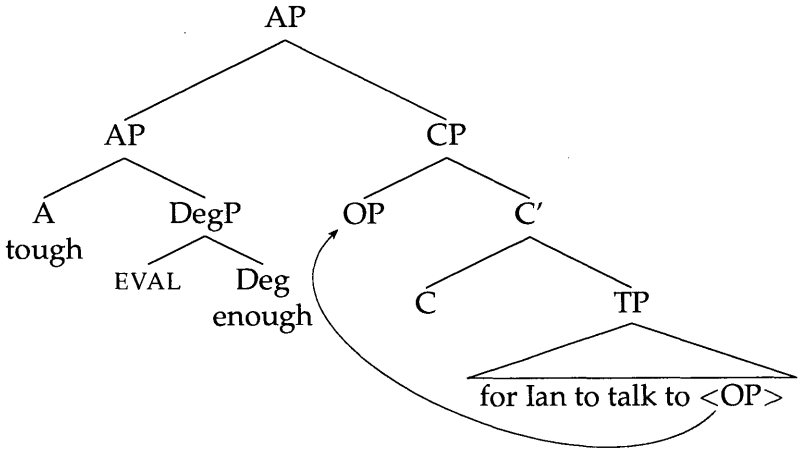
(70) Anneke is tough enough for Ian to talk to _.

a. It is difficult enough for Ian to talk to Anneke (any more serious interaction would be unthinkable) (TC reading)

b. Anneke herself is tough enough for Ian to talk to (he prefers intimidating women) (GDP reading)

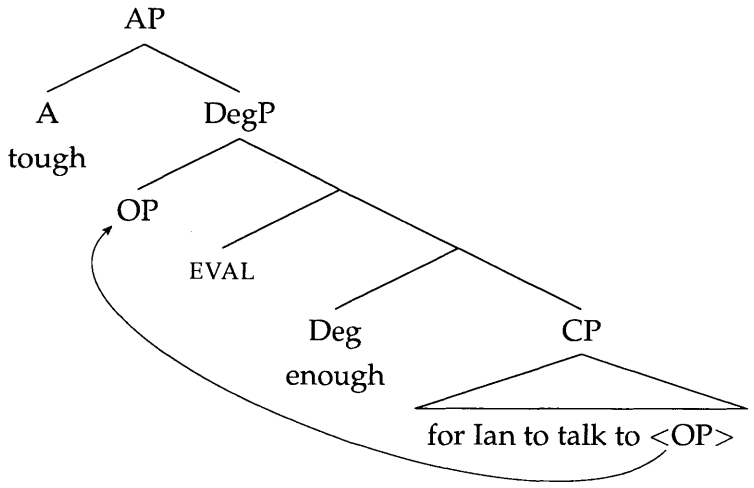
The ambiguity in (70) is structural. In (70a), the DegP adjoins to the AP, as schematized in (71). Despite containing a DegP, the (70a) interpretation of (70) corresponds to a *tough-* construction structure, where the DegP adjoins to and modifies the *tough*-predicate itself, (71).

(71) Anneke is [$_{AP}$ [$_{AP}$ tough [$_{DegP}$ enough]] [$_{CP}$ for Ian to talk to __]]



In (70b), the DegP instead adjoins to the CP layer, as schematized in (72). This is a gapped degree phrase structure. Importantly, in (72), the null operator, which originates as the complement to V in the embedded clause, moves directly to spec-DegP, not spec-CP.

(72) Anneke is [$_{AP}$ tough [$_{DegP}$ enough [$_{CP}$ for Ian to talk to __]]]

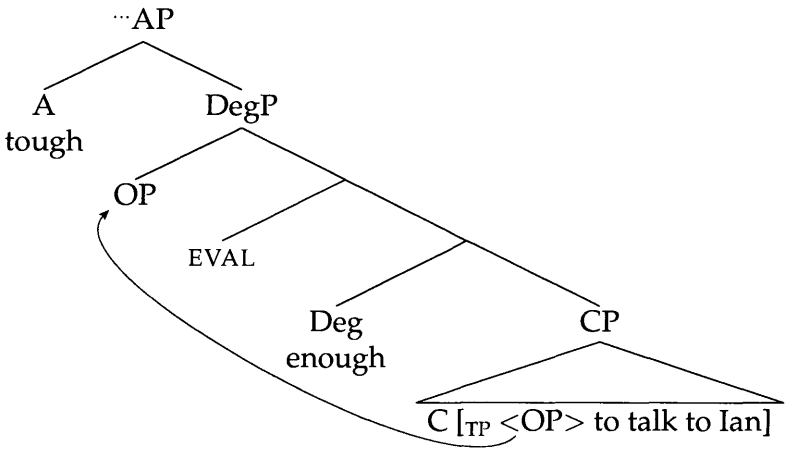


Now compare (73a) to (70), repeated in (73b):

(73) a. Ian is tough enough _ to talk to Anneke.
b. Anneke is tough enough for Ian to talk to _.

Unlike (73b) (an object gap construction), (73a) (a subject gap construction) is not ambiguous between a *tough*-construction and gapped degree phrase interpretation. Only the gapped degree phrase reading of (73a)—where *Ian* receives a Θ-role from *tough*—is available; the *tough*-construction reading is not. The structure for the single possible interpretation of (73b) is given in (74).

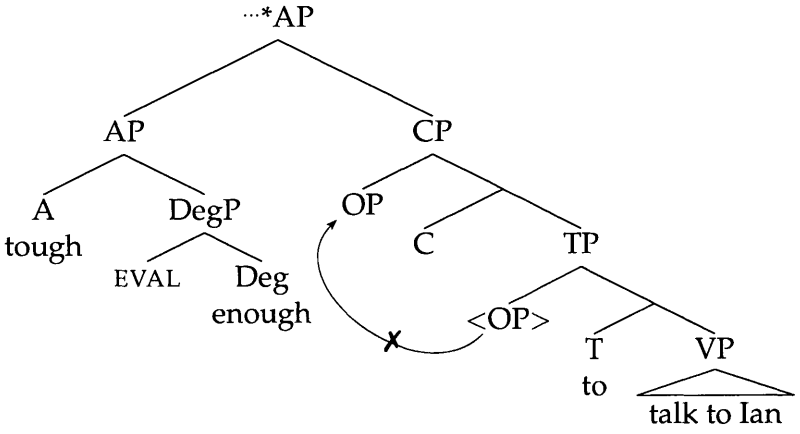
(74) Anneke is [$_{AP}$ tough [$_{DegP}$ enough [$_{CP}$ [$_{TP}$ __ to talk to Ian]]]]



The tree in (75) shows the (unavailable) *tough*-construction reading structure of (73a). In (75), the DegP layer adjoins directly to AP, modifying the

tough-predicate; it does not adjoin directly above the CP. *Tough*-constructions necessarily involve Ā-movement through spec-CP (Chomsky 1977). However, in (75), the null operator would need to move from spec-TP to spec-CP, crossing no intervening material. This movement violates anti-locality.

(75) *Anneke is [$_{AP}$ [$_{AP}$ tough [$_{DegP}$ enough]] [$_{CP}$ __ to talk to Ian]]



Subject gapped degree phrases and case theory

The discussion of gapped degree phrases began with a pair of examples tied to case theory. The first example illustrated the long standing observation that Ā-movement must extract from a case position. This restriction has previously (e.g., Chomsky 1981) been used to explain the ungrammaticality of examples like (57a), repeated in (76a). Specifically, in (76a) *who* cannot receive case in the embedded clause and so cannot be Ā-moved to a caseless *wh* position in the matrix clause. The second example illustrates the so-called "*for*-trace" effect—which is often considered to fall under the same umbrella as *that*-trace effects—repeated from (57b) in (76b). The ungrammaticality of (76b), which involves an embedded subject case marked by the preposition *for* and which can still not undergo Ā-movement into a *wh*-position in the matrix clause, shows that the ungrammaticality of (76a) is not be due to case theory alone. Following this initial case theory discussion, this chapter has gone on to propose that the examples in (76) are ungrammatical at least in part because they involve subject movement chains that would violate spec-to-spec anti-locality. However, this discussion has also left open the possibility that (76a) is "doubly ungrammatical," in that it involves both a violation of spec-to-spec anti-locality and case theory[26].

[26]However, it is also possible that (ia) involves a bare TP phase, and so can allow the subject to extract directly to the matrix clause without violating spec-to-spec anti-locality. According to this analysis, (ia) would be ungrammatical solely due a subject case restriction

(76) a. *Who is it possible _ to see Mary?
b. *Who is it possible for _ to see Mary?

The subject gapped degree phrase examples discussed in this section, however, pose an additional question for case theory. The remainder of this section will argue that subject gapped degree phrases involve Ā-movement of their subject from one caseless position to the other—that is, from the subject of the infinitive to the specifier of DegP. While this analysis longer raises questions about the interaction between case theory and spec-to-spec anti-locality, the grammaticality of subject gapped degree phrases is curious under a theory where all DPs need case. Eventually, this section will argue that the subjects of subject gapped degree phrases are like PRO in that they have special case properties. First, however, this section will review possible alternative analyses for why subjects can grammatical undergo Ā-movement from the caseless infinitival subject position to spec-DegP in gapped degree phrases.

There are a few potential paths for understanding the grammaticality of subject gapped degree phrases. One path discards previous assumptions about case. This would mean rejecting the twin assumptions that all nominals need to receive case, and that caseless arguments cannot undergo Ā-movement. Taking this path would mean arguing, essentially, that (76a) is ungrammatical only because it involves an anti-locality violation. This would also require assuming that case theory does not play a role in determining subject extraction restrictions from embedded infinitival clauses more generally. This path makes predictions that do not appear to be borne out. For example, this analysis would predict that (77), a minimal pair with (76a), is grammatical. This prediction is due to the fact that (77) involves a high attaching adverb adjoining between TP and CP in the embedded clause, meaning that movement from spec-TP to spec-CP would no longer violate spec-to-spec anti-locality. If spec-to-spec anti-locality where the only force in determining the ungrammaticality of (76a), (77) should be a well-formed English sentence. The example in (77), however, is strongly ungrammatical, so this kind of analysis does not appear fruitful.

(77) a. It is possible for all intents and purposes for Sue to run this city.
b. *Who is it possible for all intents and purposes (for) _ to run this city?

Another path forward is to propose that spec-DegP is a case position and that the subjects of subject gapped degree phrases are case-marked after they move to

spec-DegP. However, there is no empirical reason to believe that spec-DegP is a case position, or that degree words can case-mark DPs. Evidence against the proposal that degree words can assign case come from examples like (78), which show that gapped degree phrases behave like all other non-finite clauses in requiring that overt subjects (of gapless degree phrases and object gapped degree phrases) must be case marked by the preposition *for*.

(78) a. John is poor [$_{DegP}$ enough [$_{CP}$ for the monastery to hire him]]
b. *John is poor [$_{DegP}$ enough [$_{CP}$ the monastery to hire him]]
c. John is poor [$_{DegP}$ OP enough [$_{CP}$ for the monastery to hire _]]
d. *John is poor [$_{DegP}$ OP enough [$_{CP}$ the monastery to hire _]]

In (78a-b), a gapless degree phrase (which involves no null operator movement) is only grammatical if the preposition *for* can case-mark the subject of the embedded clause. In (78c-d), the same holds for object gapped degree phrases. This suggests that degree words cannot assign case to arguments in their clause. Similarly, there is no evidence to suggest that spec-DegP is a case position.

Subject gapped degree phrases are not the only constructions to have covert subjects that behave irregularly with respect to case. Abstracting away from issues of movement, the subjects of subject gapped degree phrases raise the same questions that PRO, the subject of control constructions, raises. A third way to understand the behavior or subject gapped degree phrases is to claim that subject gapped degree phrases, like control constructions, involve a subject PRO. However, crucially, this subject PRO needs to undergo Ā-movement in subject gapped degree phrases, unlike the PRO that occurs in baseline control constructions. While there is a still a considerable debate regarding whether or not PRO receives case[27], PRO is still able to appear in the subject position of a non-finite clause[28]. It's worth pointing out that shifting to an analysis where subject gapped degree phrases contain a subject PRO will not provide additional evidence about which theory of PRO best represents the full spectrum of control data, however this analysis may still be able to help us better understand subject gapped degree phrases. The next section will present evidence in favor of the argument that subject gapped degree phrases contain a subject PRO.

[27]Among proponents of the theory that PRO does receive case, there is an additional debate surrounding the kind(s) of case that PRO can receive.

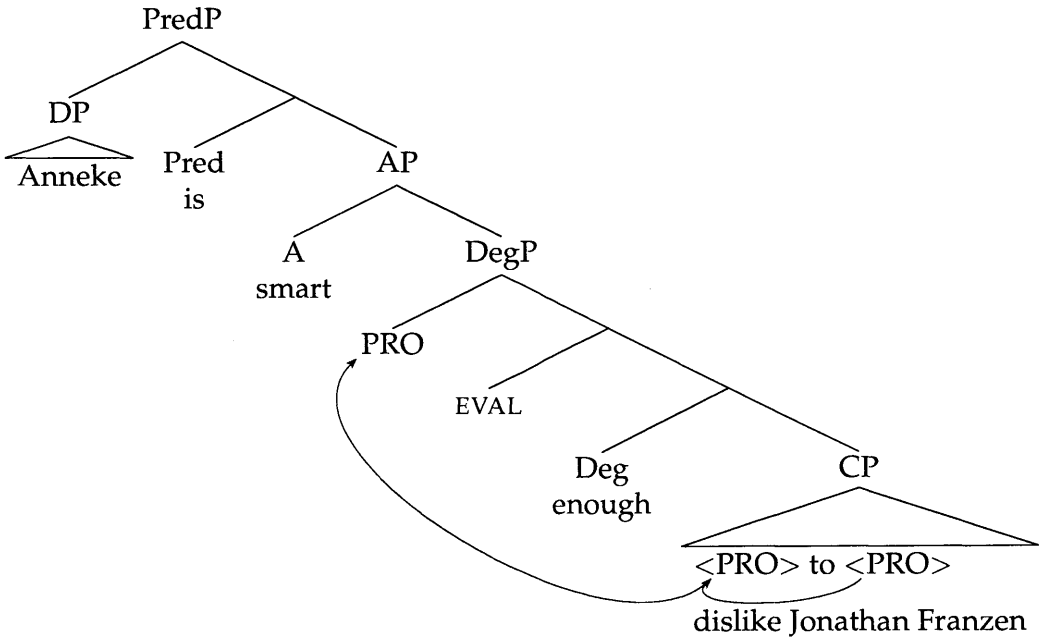
[28]Because any theory of PRO will be prediction neutral for the analysis in this section—anti-locality will be active in the grammar regardless of whether PRO is an exception to the Case Filter, receives default accusative case, or any analysis—this chapter will remain neutral between all competing theories of case marking PRO.

Subject gapped degree phrases are control and movement constructions

Thus far, this dissertation has represented both subject and object gapped degree phrases as null operator constructions, following Nissenbaum & Schwarz. This has been an entirely theory-internal choice: Nissenbaum & Schwarz (as well as previous discussions of gapped degree phrases, e.g., Lasnik & Fiengo) limited themselves to discussions of object gapped degree phrases. Chapter 2 discussed gapped degree phrases in detail, and showed that subject and object gapped degree phrases share a number of features and behaviors. Particularly given evidence that shows subject gapped degree phrases behave like movement constructions (e.g., Section 5.2.2), it has thus far made sense to represent subject and object gapped degree phrases as similarly as possible.

However, because English has two covert DPs, null operator and PRO, it is possible—at least naively—to represent subject gapped degree phrases as having either a null operator or PRO as their subject. Previously, this chapter has provided null operator representations of subject gapped degree phrases; a PRO representation of subject gapped degree phrases is given in (79).

(79) Anneke is smart enough to dislike Jonathan Franzen.



This section explores whether the subjects of subject gapped degree phrases should be represented as null operators or PRO. Additionally, this section explores the question of whether these two representations are distinct, and whether or not the distinction between PRO and null operators is meaningful. This discussion will conclude that subject gapped degree phrases should be understood as having a subject PRO, and that the distinction between PRO and null operators is meaningful. This is because, in addition to the evidence that suggests that subject gapped degree phrases are movement constructions, there is also evidence that shows that subject gapped degree phrases share some behaviors with PRO constructions. For example, subject gapped degree phrases can occur in partial control constructions, like (80). Note that subject and object gapped degree phrases differ their ability to license partial control. Object gapped degree phrases, unlike subject gapped degree phrases, are not compatible with a partial control interpretation, as in (81).

(80) Subject gapped degree phrases license partial control
The chair is too busy to meet at noon.

(81) Object gapped degree phrases do not license partial control
a. *This spice is too expensive for us to combine _.
b. *Mary is too stubborn for us to assemble _ in the town square.

Additionally, subject gapped degree phrases resemble control constructions in that they do not allow for an idiomatic interpretation in examples like (82).

(82) The cat is too important to _ be let out of the bag.

The evidence in (80)-(82) suggest that subject gapped degree phrases can be understood as having a subject PRO, as in (79). In the case of (79), the subject PRO moves to spec-DegP just as the object null operator of an object gapped degree phrase does. On this analysis, subject gapped degree phrases do not actually constitute a mystery for case theory. Or rather, subject gapped degree phrases do not constitute a *new* mystery, since PRO has long been considered mysterious within case theory.

Clearly, understanding subject gapped degree phrases as having a subject PRO does not necessarily clarify any questions about how PRO interacts with case theory. However, understanding how gapped degree phrases interact with case theory does become less pressing for the current analysis. Instead, subject gapped degree phrases emerge as another construction to be considered when developing a comprehensive theory of PRO.

While there is some debate within the literature about whether or not PRO can receive (or is required to receive) case, there have also been proposals for how to understand the distribution of PRO with respect to case theory. Earlier analyses of PRO assumed either that PRO did not receive case (Chomsky 1981), or that PRO received some special case specific to the infinitival subject position where PRO to most commonly proposed to occur (Chomsky 1995). However, data from languages with lexical case marking, like Russian and Icelandic, suggest that PRO can receive lexical case when it is the subject of lexical-case assigning verbs (c.f., Sigurdhsson 2008 for the Icelandic data) while recent evidence from Landau (2001, 2006) also suggests that PRO behaves relatively standardly with respect to case in Hebrew. While the data is complicated, it does suggest that PRO can receive case and also that PRO is not required to receive either a special PRO or infinitival subject case. Other analyses have proposed that the subject position of

a non-finite clause is not necessarily a caseless position (at least for English), in that subjects of infinitives (like PRO) can receive default accusative case in their base position, and so should still be able to undergo movement (Schütze 2001).

Following evidence from partial control and idiom tests, this section has proposed that subject gapped degree phrases can be understood as having a subject PRO. This proposal helps us understand how subject gapped degree phrases interact with case theory. However, if this proposal is correct, then subject gapped degree phrases are also somewhat unusual PRO constructions, in that they involve a PRO that moves to the edge of their embedded clause.

This analysis expands the idea of the movement operations that PRO can engage in. Particularly, subject gapped degree phrases show us that in at least some circumstances, PRO can undergo movement to the edge of its phase. This aligns with previous arguments that PRO can undergo other types of movement (e.g., Radford 2004 for an argument that PRO can undergo EPP movement to spec-TP).

Under an analysis where PRO does not receive Case, there is some debate about whether or not PRO can undergo movement to non-case positions (e.g., spec-CP or spec-DegP). This is a Case theory question, along the lines of (57b), and supported by examples like (83), taken from Chomsky (1981:176).

(83) a. *the man [$_S$ that [$_S$ you tried [$_S$ [$_S$ t to win]]]]
b. *the man [$_S$ that [$_S$ I wonder [$_S$ what [$_S$ t to see]]]] (Chomsky 1981:176 exx 19)

However, the ungrammaticality of (83) can also be explained by anti-locality. The structure in (84) show an updated representation of (83a) and its grammatical object gap correlate. Following standard assumptions about raising and control (e.g., Landau 2003) and Chomsky (1981), the structure in (84a) assumes that control predicates embed full CPs, and so CP_2 is the phase layer for the embedded clause. Ā-moving PRO successive cyclically to the edge of CP_1 would require moving PRO anti-locally from spec-TP_1 to spec-CP_1. The same anti-local movement would not be necessary to move the object null operator successive-cyclically to spec-CP_2 in (84b)[29].

[29]Again, the structures in (i) propose the canonical control clause structure where a verb embeds a full CP; if we can analyze control clauses are bare TPs, there is not a clear anti-locality analysis of the facts in (i).

(84) a. *the man [$_{CP_2}$ PRO that [$_{TP_2}$ you T [$_{vP}$ <you> tried [$_{CP_1}$ <PRO> C [$_{TP_1}$ <PRO> to win]]]]]
b. *the man [$_{CP_2}$ OP that [$_{TP_2}$ you T [$_{vP}$ <you> tried [$_{CP_1}$ <OP> C [$_{TP_1}$ PRO to win <PRO>]]]]]

Distinguishing between PRO and null operators

The previous discussion proposed that subject gapped degree phrases should be understood as having a subject PRO, while object gapped degree phrases should be understood as null operator constructions. In addition, the PRO in subject position of subject gapped degree phrases is proposed to be able to move to the embedded phase edge, spec-DegP. This is not the kind of movement that PRO is traditionally proposed to undergo. This raises the question of whether or not it is theoretically meaningful to distinguish between PRO and null operators in subject and object gapped degree phrases. This section argues that, while subject gapped degree phrases show that PRO can perhaps engage in a wider array of behaviors than previously assumed, it still makes sense to treat PRO and null operators as separate entities within the theory. To make this point, this section will review both the similarities and differences between null operator and PRO, and will argue that the two elements show more differences than similarities.

Similarities between PRO and OP: At a high level, PRO and OP, the two (English) silent and independent DPs, appear similar. Most obviously, null operators and non-arbitrary PRO (outside of partial control contexts) must be semantically bound to an argument that is structurally above their embedded clause. Additionally, both elements—when they occur in embedded clauses and are bound to a higher argument—can undergo movement within their embedded clause. The literature also argues that both null operators and PRO are semantically bound to their antecedent via movement.

This movement is described as follows (see Chapter 2 for an expanded discussion of null operator binding): Null operators are proposed to move from a lower position in an embedded clause (usually an object position) to a higher position at the edge of the embedded clause of type <t>, resulting in the creation of a type <e,t> node that can be bound with the higher clause via some semantic operation, such as predicate abstraction or COMPOSE (e.g., Nissenbaum & Schwarz). Bound null operators are proposed to be semantically vacuous when they enter the derivation and to receive an obligatory interpretation after semantic binding. Heim and Kratzer (1998:226), following May (1977), propose a

nearly identical life cycle for PRO. They propose that non-arbitrary PRO enters the derivation as "an empty pronoun" and is bound to its antecedent after movement. Like null operators, PRO moves to a node of type <t>, creating a type <e,t> node that can bound to a higher clause by predicate abstraction (or some similar semantic operator). However, while null operators and PRO are proposed to bind with their antecedents in parallel ways, they are not the only arguments proposed to undergo this type of binding. For example, Heim & Kratzer propose that overt pronouns also bind with their antecedents in parallel ways.

Differences between PRO and null operators: PRO and null operators also show a number of differences, however. First, PRO can be either referential or arbitrary, as in (85). It is not immediately clear that null operators have an arbitrary variant. While examples like (86a-b) appear to show something that resembles OP_{arb}, resultative tests like (86c) suggest that these silent objects are actually implicit arguments, and null full DPs.

(85) PRO_{arb} sentences
   a. PRO_{arb} training dogs takes patience.
   b. PRO_{arb} pushing other students is unacceptable.

(86) OP_{arb} sentences
   a. Watch out for this dog, he bites OP_{arb}.
   b. I don't like Billy, he punches OP_{arb}.
   c. *I don't like Billy, he punches OP_{arb} black and blue.

Additionally, PRO occurs much more freely than null operators. For example, unlike PRO, null operators cannot occur in any clause embedded by a control predicate. This is shown in (87).

(87) *John tried for Mary to talk to _.

Additionally, PRO and null operators differ in their case requirements. As previously discussed, PRO does not behave like overt DPs with respect to case. However, null operators do. For example, object null operators need to receive case either from a transitive verb or a preposition, as in (88).

(88) a. Which guest did Mary ignore [$_{CP}$ OP after talking to _]
   b. *Which guest did Mary ignore [$_{CP}$ OP after talking _]

Finally, as discussed in Section 5.2.5, null operators are ungrammatical in partial

control constructions, unlike PRO.

All this suggests that PRO and null operators are distinct elements. The literature also has traditionally treated PRO and null operators as distinct entities. While this is not empirical evidence for a distinction, it's worth noting that null operators and PRO are traditionally claimed to occur in distinct contexts. For example, PRO is traditionally argued to only occur in subject position (see Martin 2001 for arguments against allowing PRO to occur in non-subject positions), while null operators are disproportionately often proposed to occur in non-subject positions[30]. Null operators, also, are argued to occur in Ā-movement constructions: parasitic gaps (Nissenbaum 2000), relative clauses, and non-improper movement analyses of *tough* constructions (Hicks 2009). Similarly, PRO is not regularly argued to undergo Ā-movement. However, there is no reason to believe that null operators cannot undergo A-movement, or that PRO cannot undergo Ā-movement.

Comparing the properties and distributions of PRO and null operators suggests that there is a theoretically meaningful distinction between the two elements. Particularly, analyzing subject gapped degree phrases as having a subject PRO and object gapped degree phrases as having an object null operator helps account for why the two constructions show distinct behaviors in some regards.

Conclusions

The contrast between subject *tough*-constructions and subject gapped degree phrases shows the signature of anti-locality. Subject *tough* constructions, the structures that would require a short movement step from spec-TP to spec-CP are ungrammatical. However, once the movement chain is lengthened, from spec-TP to spec-DegP in subject gapped degree phrases, the movement chain is licensed again. Because the embedded clause in gapped degree phrases contains a DegP layer above the CP, both subject and object gapped degree phrases are grammatical. Because the embedded clause in *tough* constructions involves a bare CP, only object *tough* constructions are possible.

This section also proposed a distinction between subject and object gapped degree phrases: while object gapped degree phrases involve a null operator that

[30]With the exception of some analyses of subject relative clauses, I am not aware of any other structures proposed to contain a subject null operator. However, I am also not aware of any arguments that propose that null operators are banned from occurring in subject positions

undergoes movement to the phase edge, subject gapped degree phrases involve a PRO that follows a parallel movement path. This difference between subject and object gapped degree phrases helps us understand how subject gapped degree phrases interact with case theory, and explains why subject and object gapped degree phrases pattern differently with respect to partial control.

3.5.3 Matrix subject *wh*-questions

Matrix subject questions and matrix non-subject questions differ in a number of ways. First, *do*-support is prohibited in subject questions (89), but required in non-subject questions (90):

(89) Subject question:
a. Who saw John?
b. *Who did _ see John?

(90) Object question:
a. *Who John saw _?
b. Who did John see _?

Second, parasitic gaps are not licensed in subject questions, but are licensed in non-subject questions (Engdahl 1983):

(91) a. *Who _ hired Mary [without her talking to *pg*]? (subject)
b. Who did Mary hire _ [without talking to *pg*]? (object)

A third related difference comes from Chung and McCloskey (1983), who observe that subject relatives show weaker *wh*-island effects than non-subject relatives. The bracketed clause in (92a) is a subject relative, and in (93a) the bracketed clause is a non-subject relative. In the (b) examples, *which* is extracted from the relative clause. Comparing (92b) and (93b), subject relatives do not give rise to *wh*-island effects where non-subject relatives do; extraction of *which* is grammatical in (92b), but ungrammatical in (93b)

(92) a. Paul and Stevie were the only ones [*who* wanted to record that song].
b. Isn't that the song which Paul and Stevie were the only ones [*who* wanted to record <which>]?

(93) a. Paul and Stevie were the only ones [*who* George would let <*who*> record that song].
b. *Isn't that the song which Paul and Stevie were the only ones [who George would let <*who*> record <which>]?

However, a complication to this argument comes from examples like (94), which involve extracting a non-nominal/non-argument. In this context, there is no

distinction between subject and object relative clauses.

(94) a. Paul and Stevie were the only ones [*who* wanted to record that song with a synthesizer].
b. *Isn't that how Paul and Stevie were the only ones [*who* wanted to record that song <how>]?

(95) a. Paul and Stevie were the only ones [*who* George would let <*who*> record that song].
b. *Isn't that how Paul and Stevie were the only ones [who George would let <*who*> record <how>]?

An anti-locality account

As briefly discussed in Section 2, the analysis of subject *wh*-questions cannot be straightforward if anti-locality is correct. If *who* in (89a) underwent *wh*-movement, it would move from spec-TP to spec-CP with no intervening projection as in (96), and (89a) would not respect anti-locality, rendering it ungrammatical.

(96) [$_{CP}$ who C [$_{TP}$ <who> saw John]]

We suggest that an anti-locality configuration is avoided by leaving subject *who* in situ. If (96) involved *wh*-in-situ, then the ungrammaticality of (89b) and (91a) and the grammaticality of (92b) are all predicted.

With regard to *do*-support, English is a residual V2 language Holmberg (2010), where movement of an element to spec-CP in matrix clauses is typically accompanied by T-to-C movement. If subject *who* in (89) does not move to spec-CP, it is not surprising that T-to-C movement is not triggered and that *do*-support does not occur.

The parasitic gap data are similarly explained. It is well known that Ā-movement is required for parasitic gaps to be licensed. In Nissenbaum's (2000) analysis, Ā-movement is involved in creating a host for the adjunct containing the parasitic gap that is of an appropriate semantic type to compose with the adjunct. If *who* does not undergo Ā-movement in (91a), semantic composition fails, and the sentence is ungrammatical. This would also explain why subject gapped degree phrases are able to license parasitic gaps in (69).

The relative clause data, (92)-(93), are also explained by this analysis. If *who* is in situ in (92b), the subject relative is not a *wh*-island, and (92b) is grammatical.[31]

[31]Though, see Pesetsky and Torrego (2001) (2001) and Ginzburg and Sag (2001) for a counterar-

Neutralization

As previously seen, the introduction of an adverb (within an AdvP) along the clausal spine between TP and CP is enough to break an anti-local configuration. In the case of subject *wh*-questions, the addition of such an adverb appears to allow *wh*-movement to occur. Subject *wh*-questions that involve the adverb *for all intents and purposes* can license *do*-support and parasitic gaps (98).

(97) a. *Who does _ serve as president?
b. Who does **for all intents and purposes** _ serve as president?

Although the judgment is not clear for all speakers, several informants report that a non-emphatic reading of *do* is possible in (97b), unlike in (97a), indicating that when *for all intents and purposes* is present, the subject can undergo movement from spec-TP to spec-CP, and T-to-C movement is triggered.

The parasitic gap facts, originally observed by Longobardi (1985), are shown below. In (98), the subject can undergo non-anti-local movement across the adjunct clause containing a parasitic gap.

(98) Who [without Mary talking to *pg*] _ hired Jill nonetheless?

The adjunct containing the parasitic gap is plausibly attached between TP and CP. The position of *who* to the left of the adjunct indicates that it has moved from spec-TP to spec-CP, and because the adjunct is present, this movement respects anti-locality. With Ā-movement taking place, the parasitic gap in the adjunct is licensed. The grammaticality of (98) is thus consistent with our proposal.

3.6 Motivating anti-locality

This chapter has shown that anti-locality can account for a number of unrelated subject/non-subject asymmetries, both within English and across other languages. However, while anti-locality appears to be an active force in the grammar, it is not clear what motivates anti-locality—that is, it is not clear *why*

gument to subject questions being *wh*-in-situ from the distribution of *the hell*: *the hell* can generally attach to a *wh* only if has moved to spec-CP, e.g. (i) vs. (ii), and is possible in matrix subject questions, (iii).

(i) What the hell did John steal _? (ii) *John stole *what the hell*? (iii) *Who the hell* stole the jewels?

anti-locality exists as a restriction within the grammar. This is a particularly important question because the forces that have been proposed to motivate previous anti-locality constraints—Grohmann's PF constraint governing clause domain formation, Abels' definition of relationships between heads and their complements/specifiers—don't appear to obviously influence spec-to-spec anti-locality, or the alternations it predicts. One possibility is that spec-to-spec anti-locality is related to Contiguity Theory (Richards 2016), which proposes a prosody-driven account for why certain movement operations must occur over at least a specified distance. However, this question is left open for future research.

An important attribute of anti-locality is that it only applies to Ā-movement operations. A-movement operations, for example short verb movement from spec-vP to spec-TP, can be very local. This analysis also assumes that there is some fundamental difference between A and Ā-movement operations. Future research should explore this distinction further. It is, for example, possible that this distinction—the fact that anti-locality limits Ā but not A-movement operations—is not an arbitrary fact of the theory, but instead speaks to the distinction between A and Ā-movement more generally. While this theory could not explain *why* there is a distinction between A and Ā-movement, it could potentially help us understand the distinction between the two operations. Essentially, spec-to-spec anti-locality might be understood as a property of Ā-movement, in that all Ā-movement operations must be anti-local. Considered this way, anti-locality could potentially be thought to sharpen the long-standing observation that, while A-movement has to take place locally, Ā-movement tends to take place over longer distances—for example, the principle that underlies anti-locality might be able to somehow be connected to the same principles that limit long-distance A-movement (e.g., relativized minimality). At the very least, sensitivity to spec-to-spec anti-locality could also be understood as a diagnostic for distinguishing A and Ā-movement operations from each other, similar to weak crossover tests. It's also worth asking if we can understand Ā-movement as something that *must* be somehow definable as "long-distance movement", in a way analogous to our understanding that only Ā features can trigger Ā-movement. In this theory of movement, the feature that motivate A-movement operation would matter, but so would the structural relation between the head and the links in that movement chain.

3.7 Conclusions

This chapter has introduced and discussed spec-to-spec anti-locality, a constraint that limits certain Ā-movement paths that are *too short*–that is, this chapter has introduced a ban on very short Ā-movement chains, focusing on subject Ā-movement chains that originate in spec-TP and end in spec-CP. This chapter has shown that spec-to-spec anti-locality is operative in English, as illustrated by the sequence of examples given in Section 5, and is very likely operative in other languages. Section 4 discussed the ways that spec-to-spec anti-locality could provide a way to unify a series of unrelated subject/non-subject asymmetries across three unrelated languages (Imbabura Quechua, Hebrew, and Tarafit Berber). Finally, this section has addressed a puzzle introduced in Chapter 2: it has provided an explanation for why subject gapped degree phrases are possible while subject *tough*-constructions are prohibited.

Chapter 4

Acquiring *tough* and raising constructions

4.1 Introduction

This chapter discusses an experiment designed to help understand how children acquire *tough* and raising constructions. Using an experimental paradigm pioneered by Naigles *et al* (1995) and adopted for acquisition work by Becker et al. (2012), this experiment is the first to directly compare how children perceive the grammaticality of different types of raising and *tough* constructions. Particularly, this study uses the combination of a child's reaction time and expected response rate[1] as a proxy for a grammaticality judgment. These grammaticality judgments help us understand how children represent both *tough* and raising constructions, and allow us to compare child and adult grammars more directly.

The results of this study show that children do not have an adult-like representation of either *tough* or raising constructions. Importantly, however, these representations are not non-adult-like in the same way. Particularly, children's grammars are *too permissive* with respect to raising constructions and *too restrictive* with respect to *tough* constructions. While adult grammars allow subject raising and ban object raising, the child grammar allows both subject and object raising. With *tough* constructions, though the adult grammar allows for object *tough* movement but bans subject *tough* movement, the child grammar bans

[1]Other acquisition studies call this measure "accuracy." However, as will be discussed, "accuracy" is not a straight-forward concept in this study. Instead of looking at whether or not children are able to answer a question "accurately," this study will look at whether child answers pattern with adult answers. If children respond to a question with the same answer that we expect an adult English speaker would give, that response will be classified as an expected response.

both subject and object *tough* movement.

The results of this study can be summarized as follows: children find subject and object raising constructions grammatical and subject and object *tough* constructions ungrammatical. Some of these individual judgments align with adult grammaticality judgments (for example, both children and adults find subject *tough* movement ungrammatical), but these results, when considered together, show a non-adult grammar. Two child grammaticality judgments are unlike adult grammaticality judgments—children judge adult-ungrammatical object raising constructions as grammatical and adult-grammatical object *tough* constructions as ungrammatical—and two child grammaticality judgment align with adult-grammaticality judgments—both children and adult finds subject *tough* constructions ungrammatical, and believe that subject raising constructions are grammatical. Crucially, child and adult grammaticality judgments do not align across raising, *tough* movement, or object movement. The results of this study are given in (1).

(1) Child and adult grammaticality judgments

Constructions	Child	Adult
Subject raising	✓	✓
Object raising	✓	*
Subject *tough*	*	✓
Object *tough*	*	*

This chapter will discuss how we can interpret these results in the context of the previous acquisition literature on both raising and *tough* constructions, as well as how we can understand child representations of *tough* movement and raising within the current theoretical model of syntax.

4.1.1 Comparing *tough* and raising constructions

Tough and raising constructions have a number of properties in common: they are both infinitival embedding constructions that allow a matrix subject to associate with an embedded gap site, and both constructions are acquired later in childhood. However, when the two constructions are directly compared, a number of differences also emerge: Raising constructions are A-movement constructions that do not allow for object extraction from their embedded clause (2). The literature has argued that this restriction on object raising is due to relativized minimality restrictions that prevent the object from A-moving across

the subject. *Tough* constructions however are Ā-movement constructions[2] that do not allow subject extraction from the embedded clause (3). Chapter 3 of this dissertation has proposed that this asymmetry is due to an anti-locality constraint that disallows short Ā-bar movement from spec-TP to spec-CP.

(2) Raising constructions

a. Mary is likely [_ to see John]. (subject raising is ✓)
b. *Mary is likely [for John to see _] (object raising is ✗)

(3) *Tough*-constructions

a. *Mary is tough [_ to see John] (subject *tough*-movement is ✗)
b. Mary is tough [for John to see _] (object *tough*-movement is ✓)

Because English allows for a null PRO to serve as the subject of some embedded clauses, English also displays so-called symmetric *tough*-constructions, which lack any overt argument in the embedded clause, as in (4). Because English lacks an object-equivalent of PRO, however, symmetric raising constructions are not possible (5).

(4) Symmetric *tough*-construction
Jonathan Franzen is tough [PRO to talk about _]

(5) Symmetric raising construction (impossible in adult English)
*Jonathan Franzen is likely [_ to talk about PRO]

The *tough* construction acquisition literature has paid particular attention to the case of symmetric *tough*-constructions. Specifically, it has been known since Chomsky (1969) that children are able to associate symmetric *tough*-constructions like (4) with either of the scenarios outlined in (6). This is not adult-like behavior. Adult English speakers are only able to associate the symmetric *tough* construction in (6) with the scenario described in (6b).

[2]As will be discussed in Section 2.2 (and as was discussed in the introduction to this dissertation), the literature on *tough* movement is controversial and there is no universally accepted theory of *tough* movement. According to one theory of *tough* movement, the improper movement theory, *tough* constructions are like raising constructions in that they contain an A-movement step (raising constructions, however, are not proposed to involve an Ā-movement step under this analysis). However, all analyses of *tough* constructions agree that *tough* constructions contain at least an Ā-movement chain within their embedded clause.

(6) The dog is tough [_ to see _]

a. *Subject gap interpretation*: The dog is vision-impaired.
Analogous to: *It is tough for the dog to see* ***anything***

b. *Object gap interpretation*: The speaker cannot easily see the dog.
Analogous to: *It is tough for* ***anyone*** *to see the dog*

While there is a significant literature surrounding children's representation of adult-grammatical symmetric *tough* constructions like (4), there has not been much explicit investigation into how children judge adult-ungrammatical subject *tough* constructions like (3a). Similarly, there is not a robust literature surrounding how children judge adult-ungrammatical symmetric raising constructions like (5), or adult ungrammatical object raising constructions like (2b). If child grammaticality judgments do not match adult grammaticality judgments, we need to ask what a child's syntactic representation of of a particular constructions might look like. The question, then, is not just 'does a the child grammar contain a representation of an adult-grammatical construction' but also 'how might a child represent an adult-grammatical construction, given their judgments.' The current study has gathered child grammaticality judgments for subject raising (adult grammatical), object raising (adult ungrammatical), subject *tough* (adult ungrammatical) and object *tough* (adult grammatical) constructions. In this way, the results of this study are able to paint a more complete picture of the child grammar, compared to the previous literature.

4.1.2 Child grammaticality judgments and the child grammar

Discussions of how children represent specific linguistic phenomena are inherently ambiguous. Most studies can only provide information about whether children can consistently answer a question involving a certain syntactic structure the way that an adult speaker would. This is not the same as providing a model of the structure that the child invokes when they process a construction. Additionally, experiments that only look at whether a child can understand adult-grammatical questions may not provide a complete picture of how a children understands a syntactic operation. For example, in such experiments is often very different to determine if non-adult-like behavior is the result of the child finding the construction ambiguous or ungrammatical. For this reason, this study focuses on child grammaticality judgments directly. However, the data from this study is still somewhat open to interpretation. Because child judgments

are determined by proxy, child grammaticality judgments are in many ways more "fuzzy" than adult-grammaticality judgments. One consequence of this fact is that child grammaticality judgments must be gathered from a statistically significant number of children (compared with adult-grammaticality judgments, where the response of a single adult participant is often assumed to be representative of their dialect). Additionally, child-grammaticality judgments come with less meta-data. In an adult grammaticality judgment, the adult participant can often be asked directly about the locus of ungrammaticality. These kinds of follow-up questions aren't possible in child studies. As a result, when this chapter discusses the child grammar directly—for example, in terms of the syntactic structures or the operations available to children—the claims will less definitive that those of a traditional theoretical syntax paper. The discussion here will instead focus on structures that are consistent with the child-grammaticality judgments available, while acknowledging that in many cases there may be more than one structure that is consistent with the data, and that the data may not be robust enough to distinguish between competing theories.

4.1.3 Secondary goals of this chapter

The main goal of this paper is to provide child grammaticality judgments for subject raising (child and adult grammatical), object raising (child grammatical and adult ungrammatical), subject *tough* movement (child and adult ungrammatical) and object *tough* movement (child ungrammatical and adult grammatical) constructions and to discuss what these judgments say about the relationship between the child and adult grammars. However, these results can also engage with other aspects of the literature. Specifically, these results can also be used to help distinguish between two competing hypotheses that have arisen within the *tough* and raising acquisition literature. Additionally, these results, when taken in the context of the broader acquisition literature, may be able to help distinguish between competing theoretical models of *tough* movement, though this analysis is much more tentative and preliminary.

Distinguishing between competing hypotheses

Within the *tough*-construction acquisition literature, there are two distinct theories about why children do not have an adult-like representation of *tough*-constructions. Because the previous literature is most immediately focused

on determining which possible scenarios a child might be able to associate with an adult-grammatical symmetric *tough* construction, both hypotheses are consistent with the literature surrounding the acquisition of *tough* constructions. As previously mentioned, the majority of the *tough* acquisition literature is focused on how children engage with symmetric *tough* constructions like (6). They key tension in the literature surrounds whether children find constructions like (6) ambiguous or ungrammatical. One hypothesis, advanced by Wexler (2012), predicts that children find sentences like (6) ungrammatical, and so do not have access to any either an object *tough* (adult grammatical) or a subject *tough* (adult ungrammatical) reading of (6). Therefore, to be able to provide a response to a stimulus like (6), children must make use of some repair or reanalysis strategy. The other hypothesis, advanced by Anderson (2005), proposes that children have access to both an object *tough* and subject *tough* interpretation of (6), and find both these interpretations grammatical.

Importantly, both hypotheses also extend to the raising data: Wexler predicts that children will find subject and raising constructions ungrammatical, while Anderson predicts that children will find subject and object raising constructions grammatical. Because this study is set up to directly solicit child grammaticality judgments for subject and object raising constructions, it also provides a way to test the predictions of these competing hypotheses across multiple constructions, in order to understand which hypothesis has the greater predictive power. The specifics of each hypothesis and additional details about how this experiment is able to distinguish between their predictions is given in Section 2.

Distinguish between theories of *tough* movement

Tough movement is a controversial topic within the syntax literature. The controversy centers around the question of whether *tough* movement involves a null operator that Ā-moves from within the embedded vP to the edge of the embedded clause where it is then semantically bound to the matrix subject (the null operator analysis) or whether *tough* movement involves an overt DP that first Ā-moves from the embedded vP to the edge of the embedded clause and then A-moves to the matrix subject position, a process known as improper movement (the improper movement analysis). For a fuller discussion of theories of *tough* movement, see Section 2.2 of this chapter as well as Chapters 1, 2 and 3 of this dissertation. Understanding child grammaticality judgments for *tough* constructions may additionally help distinguish between syntactic theories of

tough movement.

Acquisition data has the potential to help distinguish between theoretical models of *tough* movement because each theory of *tough* movement requires a different syntactic skill set. For children to be able to find *tough* constructions grammatical under the null operator theory, they must be able to do the following: Ā-move an argument, and bind a covert argument in the embedded clause to the matrix subject. As Section 2 of this chapter will discuss, these are skills that children already have: they demonstrate an adult-like understanding of *wh*-questions (a classic Ā construction) and also demonstrate an adult-like understanding (at least some) null operator relative clause constructions (which involve binding and Ā-moving a null operator), as well as a mastery of control (a syntactic process that involves binding a covert DP to a higher argument, parallel to null operator binding). An improper movement analysis would require a larger syntactic toolbox. Given children's existing syntactic capabilities, we predict that they should be as adept at Ā-moving null operators and binding them to their matrix antecedents. If children find *tough* constructions grammatical, the results of this study might be understood as evidence that suggests that the null operator theory is a better way to understand *tough* constructions in the grammar.

An improper movement analysis of *tough* movement requires children to have a slightly larger syntactic toolbox. In order to represent an improper movement derivation of a *tough* construction, children must be able to Ā and A-move an element successively. This kind of successive movement is called improper movement and it not usually allowed in the grammar. For children to able to represent *tough* constructions as improper movement constructions, they must have already learned that *tough* constructions are an exception to a general ban on improper movement.

Because there are, to the best of my knowledge, no specific studies regarding whether or not children understand that there is a ban on improper movement (with some possible exceptions), the results of this study are harder to take as evidence in support of an improper movement theory of *tough* constructions. That being said, if children find all *tough* constructions ungrammatical, that could potentially be taken as evidence against a null operator theory of *tough* movement, as their syntactic toolbox predicts that they will be able to construct and process a null operator *tough* construction. Between the controversy surrounding the syntax of *tough* movement is a two way debate, evidence against a null operator approach can potentially also be understood as evidence in favor

or an improper movement theory of *tough* constructions.

It's important to note, however, that this experiment was not uniquely designed to differentiate between competing syntactic theories of *tough* movement, and any findings that suggest one approach over another should be taken as very tentative.

4.1.4 Organization

The remainder of this chapter is organized as follows: Section 2 summarizes the *tough* and raising literature (both the theoretical literature and the acquisition literature), providing background on both previous experiments surrounding children's representations of *tough* and raising constructions and fleshing out the competing theories of *tough* movement. Section 2 also provides a fuller account for the Wexler and Anderson hypotheses, as well as their predictions. Section 3 describes my own experiment and its design. Section 4 presents the results of the experiment and Section 5 discusses the best way to interpret those results. Section 6 concludes the chapter.

4.2 Previous literature

This section discusses (a selection of) the previous literature surrounding *tough* and raising constructions, with a focus on the acquisition literature for both those topics. This section also briefly reviews the two competing theoretical accounts for the structure of *tough* constructions, though it does not go into details of this debate (for a more thorough discussion of the controversy surrounding the syntax of *tough* constructions, see Chapter 1, Section 2 of this dissertation). This section also includes a fuller discussion of Wexler's and Anderson's hypotheses about the acquisition of subject and object movement constructions. Discussions of individual experiments will focus on how the results of those experiments are ambiguous enough to be taken as support for either Anderson's or Wexler's theory of acquisition. This section will also show how studies that ask child participants to answer yes/no questions about which scenarios they can associate with a certain sentence are particularly likely to yield ambiguous results. The main goals of this discussion are to situate unfamiliar readers with the *tough* and raising acquisition literature and to argue that the current methodology—which solicits grammaticality judgments directly—is better suited to achieve results that

provide a clearer insight into the workings of the child grammar.

4.2.1 Competing hypotheses of *tough* acquisition

Two different hypotheses have emerged in the literature to explain the delay in acquisition of *tough* constructions. Importantly, these hypotheses are primarily concerned with how children are able to establish filler gap relationships, and so both hypotheses make predictions about how children will acquire raising constructions, as well. This subsection discusses both these hypotheses in detail.

Anderson's hypothesis

Anderson (2005) claims that children find both subject and object *tough* and raising constructions grammatical. Anderson proposes that children are late to acquire *tough* constructions because they believe that symmetric *tough* constructions like (6) are ambiguous between two interpretations. One interpretation is the expected adult-like interpretation (6b). The syntactic structure that would accompany this interpretation is the adult-like representation of *tough* movement. Here, Anderson predicts that children are able to establish a filler-gap relationship between the matrix subject and the embedded object (either viA-movement of the subject itself or a null operator DP that is later bound to the subject). The second interpretation is a non-adult-like subject gap interpretation (6a). Here, a structural analysis is slightly more complicated, as the analysis is no longer constrained by the adult grammar. At its most general, Anderson's hypothesis predicts that children will be able to access *some* possible structure that aligns with the (6a) interpretation of (6)[3]. Following the discussion of *tough* constructions and gapped degree phrases in Chapters 2 and 3 of this dissertation, however, this chapter will elaborate on Anderson's predictions. This chapter will predict that, if children are able to access both interpretations of symmetric *tough* constructions like (6), that they have access to both subject and object *tough* movement. For this to be possible, children would have had to acquire an adult-like representation of *tough* movement, but not the spec-to-spec

[3]For example, Anderson's hypothesis is equally compatible with (a) a theory where children understand object *tough* movement as true *tough* movement and understand subject *tough* movement as raising (meaning that, in some sense, children do know that subject *tough* movement is not permitted in the grammar) and (b) a theory where children understand both object and subject movement as *tough* movement (meaning that children do not know that subject *tough* movement is not permitted in the grammar).

anti-locality principle that bans subject *tough* movement.

Anderson's predictions also extend to raising constructions. Specifically, Anderson proposes that children will find both subject and object raising constructions grammatical. As with *tough* constructions, Anderson does not make an explicit proposal for what syntactic structures children will use to represent raising constructions. As with *tough* constructions, this chapter will elaborate on Anderson's predictions. This chapter will predict that, if children find both subject and object raising constructions grammatical, this is because that have not acquired the subject/non-subject asymmetry inherent to raising. Section 5 of this chapter will further discuss this prediction and argue that it does not make sense to analyze children's representations of raising another way. However, it's still worth noting that Anderson's hypothesis is compatible with multiple representations of raising. For example, Anderson is compatible with a model where children do not distinguish between *tough* and raising constructions (and so assign subject raising and subject *tough* constructions that same syntax), as well as models where children represent *tough* and raising constructions distinctly. This ambiguity emerges because the core predictions of Anderson's hypothesis are child grammaticality judgments, not the structures that underly them.

Wexler's hypothesis

In direct opposition to Anderson, Wexler's hypothesis claims that children will find all four constructions (subject and object *tough* and raising constructions) ungrammatical. Wexler proposes that children do not arrive at a consistent and adult-like interpretation of (symmetric) *tough* and raising constructions because they find all *tough* and raising constructions ungrammatical—and thus impossible to parse (not ambiguous). Importantly, Wexler proposes a reason for this child ungrammaticality (a delay in the acquisition and understanding of phases, see Section 2.2.2), but does not explicitly discuss what repair strategy (or strategies) children might utilize when asked to respond to a question that they find ungrammatical. Because of this, the Wexler hypothesis is compatible with an extremely wide array of possible interpretation strategies.

At one end, these strategies barely make reference to linguistic structure: children could answer ungrammatical questions totally randomly, given only their experience in previous, unrelated conversations (allowing them to establish, for example, that the researcher is requesting a *yes* or *no* answer. This expectation could be gleaned from the researcher's intonation and the children's experience

with previous practice questions). At the other extreme, compatible strategies can make reference to (possibly untestable) theories of child processing or specific (and also possibly untestable) theories of child syntax. For example, Wexler's hypothesis would be compatible with a theory where children deal with ungrammaticality like trained linguists. Under this analysis, children would recognize that subject *tough* constructions involved moving an element from an embedded clause into a matrix clause and would be able to identify this movement chain as ungrammatical. As a result, they might be able to access the grammatical syntactic structure that is the "closest" to the ungrammatical structure—in the case of subject *tough* construction, a bi-clausal construction with an expletive *it* as its subject and a tough predicate in the matrix clause—and answer given the truth conditions of that construction. Wexler describes this process in terms of an analogy: he compares child processing of ungrammatical sentences to how an L2 language speaker might process a sentence in their second language if they knew the words that the construction contained but had not yet learned the grammatical structure (though the process that occurs for L2 learners in this context is not defined).

This chapter will pick the following repair strategy to account for how children deal with child ungrammatical *tough* and raising constructions: Hirsh, Orfitelli & Wexler's "deletion hypothesis". Under the deletion hypothesis, children are able to "delete" (or ignore) words that they cannot understand in the context of a particular construction. Thus, if a child is presented with a sentence like *John is tough to see Mary* and they have acquired a representation of transitive sentences but not *tough* constructions, they would be able to delete (or ignore) the word *tough* (and any associated function words) from the linear word order and re-process the sentence (now *John sees Mary*). When the child has performed enough deletions to arrive at a familiar and grammatical structure, they process and respond to the truth conditions of the resulting sentence. Importantly, the deletion hypothesis only proposes that children can delete from the linear word order. It does not, for example, propose that children can move or rearrange elements within the sentence or otherwise alter the word order[4]. This hypothesis also assumes that, like adults, children can correct for small errors (like agreement mismatches) on the fly and so could still understand the truth conditions of a

[4]It's worth noting that this hypothesis does not address how children identify candidates for deletion, particularly if they encounter multiple unfamiliar elements. This chapter will assume that this identification process, whatever it is, is stable and consistent across both children and constructions.

transitive sentence with, for example, a subject/verb agreement mismatch.

One important aspect of the deletion hypothesis is that it makes different predictions for subject gap and object gap constructions. This is because subject movement constructions will retain the standard transitive clause DP-V-DP (SOV) word order even after movement and deletion. This means that children should be able to arrive at an interpretable transitive representation of a subject *tough* or raising construction. Object movement constructions, however, would show a word order of DP-DP-V (ambiguous between SOV and OSV) after movement and deletion. This word order, in addition to being ungrammatical, is also ambiguous. It is not clear which DP is the subject of the verb and which is the object. A child could interpret the first DP as the subject (because it is the first DP in the sentence) or a child could interpret the second DP as the subject (because it immediately precedes the verb). We thus do not necessarily predict that children will be able to arrive at an interpretable transitive representation of object *tough* or raising constructions under the deletion hypothesis.

4.2.2 *Tough* movement literature

Acquisition literature

This section provides an overview of six significant studies about children's acquisition of *tough*-constructions. This section is by no means an exhaustive overview of the current literature, but rather highlights the studies that were particularly relevant or influential to the current project. This literature review starts with a summary of Carol Chomsky's (1969) research, which originally revealed the delayed acquisition of *tough* constructions. This is followed by a very brief review of Cromer (1970), who addresses a number of design flaws from Chomsky's original experiment. Importantly, Cromer (1970) was able to show that Chomsky's results were not the result of an experimental artifact and represent a real delay in child language acquisition. The next project discussed is Armbuster (1981), who was the first to suggest that children can legitimately associate two different contexts with a single symmetric *tough*-construction. Anderson (2005) is, in many ways, a follow-up to Armbuster. Anderson (2005) was the first study to ask children to not only chose whether or not a scenario can be associated with a symmetric *tough* construction, but also to justify their decision. Becker *et al* (2012) shows that the methodology proposed in Naigles et al (1995)—which uses the interaction of accuracy rates and reaction times as a proxy

for grammaticality judgments—can be successfully adapted to child acquisition studies. The literature review ends with Wexler (2012)—this is not an experimental paper. Rather, Wexler (2012) offers a proposal for why children show a delay in the acquisition of *tough* and raising constructions.

Chomsky 1969 & Cromer 1970: The first acquisition study that asked how children can interpret *tough* constructions was conducted by Carol Chomsky. Forty children, aged 5 to 9, were shown a blindfolded doll, named Cathy (the rest of her face was exposed) and asked the following question:

(7) Is the doll easy to see or hard to see?

Children who responded that the doll was "hard to see" (i.e., a non-adult interpretation) were asked the follow-up question in (8). Children who answered that the doll was "easy to see" (i.e., an adult-like interpretation) were asked the follow up question in (9).

(8) Could you make her easy to see?

(9) Could you make her hard to see?

Sample responses to Chomsky's follow up questions are given in (10)-(11); (10) comes from a child (age 5;0) who found the doll "hard to see" while (11) comes from a child (age 8;8) who found the doll "easy to see." In the following examples, *italics* indicate Chomsky's responses and prompts and child text is not emphasized.

(10) *Is this doll easy to see or hard to see?*
Hard to see
Could you make her easy to see?
(Pause)
Is there a way you could make her easier to see?
Yeah, you could untie that.
Why is she easy to see now?
'Cause I took this off
When she was hard to see, who couldn't see?
(He points to doll.)
Her?
Yeah
And now that she's easy to see, who's doing the seeing?
She is.
And she's easy to see. Could you tell me that? She's easy to see. Can you say

that?
She is easy to see.

(11) *Is this doll easy to see or hard to see?*
Easy.
Could you make her hard to see? Can you think of a way?
In the dark.

Of the participant pool, 26 children (65%) gave an adult-like response to the target question and 14 (35%) gave a non-adult-like response. Responses correlated with age. Almost all 5 year olds (78%) gave non-adult-like responses. Performance was mixed between children 6-8 (68% adult-like). By 9 years old, 100% of participants gave adult-like responses to the stimulus question.

Chomsky was the first to note the delayed acquisition of symmetric *tough* constructions. This study, however, is not enough to paint a complete picture of how children represent *tough* constructions. It only shows that children do not represent *tough* constructions the way that adults do; it cannot provide insight into whether or not children find *tough* constructions grammatical or what structures children can access when they interpret a *tough* construction. Chomsky's study also suffers from a number of experimental design flaws – for example, her experiment lacked control sentences, including baseline biclausal sentences with either different kinds of embedding predicates (e.g., control predicates) and *tough* predicates with expletive subjects and no overt movement. Additionally, Chomsky's context could have been confusing to children. For example, obscuring the doll's eyes is still obscuring a portion of the doll, and therefore making the doll more difficult for the child to see in its entirety.

Cromer (1970) set out to reduplicate Chomsky's work while simultaneously addressing many of these potential design flaws. Cromer's design placed a particular focus on understanding whether children struggled with all biclausal sentences and whether children believed that obscuring the doll's eyes constituted obscuring the doll, thereby making her more difficult to see. Broadly, Cromer's (1970) results support Chomsky's conclusion that not only is the acquisition of *tough* constructions delayed, but also that children are able to associate a single symmetric *tough* construction with at least two possible scenarios. Importantly, Cromer was also able to establish that this delay in acquisition is not due to the child belief that blindfolding the doll also constituted obscuring the doll or because children were generally delayed in their acquisition of all biclausal constructions.

Interestingly, Cromer in fact found a strong correlation between the acquisition of *tough* constructions and control predicate structures, though the same effect was not found between acquisition of control predicate structures and *tough* constructions. Essentially, all children in Cromer's study who only assigned an adult-like interpretation to *tough* constructions also only assigned an adult-like interpretation to control constructions. However, the reverse was not true. Many children who displayed an adult-like understanding of control constructions did not consistently provide adult-like answers to *tough* construction target stimuli. In other words, Cromer shows that control constructions are systematically acquired before *tough*-constructions.

This finding is interesting for any theory of how children judge the grammaticality of *tough* and raising constructions. Control involves semantically binding an overt argument in the matrix clause with a null argument in the embedded clause, not establishing a filler-gap dependency. This sequence of acquisition suggests that children do not necessarily struggle with semantically binding arguments across clause boundaries, and that the locus of a child's non-adult-like representation of *tough* constructions stems from having to establish a filler-gap dependency. That is, these results begin to suggest that children know that *tough* and raising constructions are both movement constructions. This also suggests that children represent *tough* constructions as control as fundamentally different constructions. However children are representing *tough* constructions, they understand them as distinct from control constructions.

Armbuster (1981): The studies by Chomsky (1969) and Cromer (1970) were not designed to understand the nature of children's delay in acquiring *tough* constructions. Armbuster (1981) was the first experiment designed to try to understand *why* children could give non-adult-like responses to *tough*-constructions. Importantly, Armbuster interpreted Chomsky & Cromer's results as suggesting that children found symmetric *tough* constructions ambiguous between a subject and object gap interpretation. In effect, Armbuster was the first study that was designed to show that children could associate symmetric *tough* constructions with two unique scenarios. It was also the first to attempt to offer clearer evidence that children could associate symmetric *tough* constructions with either a(n adult ungrammatical) subject *tough* interpretation of an (adult grammatical) object *tough* interpretation.

Specifically, Armbuster was the first study to present children with scenarios

that could be consistent with the truth conditions of either an adult-interpretation of a symmetric *tough*-construction (an object gap interpretation) or a non-adult-like interpretation of a symmetric *tough* construction (a subject gap interpretation)[5]. Importantly, Armbuster was also the first study to include overt *for*-clauses in a selection of the experimental stimuli, though Armbuster's study only looked at adult-grammatical object gap *tough*-constructions with overt embedded subjects. Because Armbuster's study only looked at how children responded to object gap *tough*-constructions with an overt embedded subject, this portion of the study cannot shed additional light on how children represent the relationship between subject and object *tough* constructions in their grammar.

In Armbuster's study, 28 children, ages 5;7 to 6;4, were asked to complete a truth-value judgment task. The children were asked to respond to various scenarios involving Kathy (a doll), Harry (a green stuffed hippo toy), a blindfold that could be placed on either Kathy or Harry, and a partially-opaque 8 in glass panel sprayed with blue enamel. The panel was difficult, but not impossible, to see through. In one condition, the panel was placed in front of (a non-blindfolded) Kathy. In another condition, Kathy was blindfolded and not obscured by the glass panel. Thus, Armbuster included both a condition where it was difficult for Kathy *to see* and a condition where it was difficult for Kathy *to be seen*. In both of these conditions, 21/28 children (75%) judged Kathy as "hard to see."

This data is most easily understood to suggest that children find symmetric *tough* constructions grammatical, and that children are able to access a grammatical representation of both subject and object *tough* constructions. This is consistent with Anderson's hypothesis. Because children surveyed had access to both an adult-like and a non-adult-like interpretation of symmetric *tough* constructions. However, Armbuster's data is still ambiguous, and is also in line with Wexler's hypothesis that children find all *tough* constructions ungrammatical. This data could suggest that children have no access to a grammatical structure for *tough* constructions, and so employ some repair strategy, when presented with a target stimulus *tough* construction. This is because a specific repair or reparsing strategy is not mandated by the Wexler hypothesis and at least some possible repair strategies could produce results that

[5]Determining the "correct" interpretation for an (adult) ungrammatical construction is not a trivial task. This chapter will assume that the "correct" or "expected" interpretation in these cases corresponds with the (expected) interpretation of the closest related (adult) grammatical construction—in the case of subject *tough* constructions, this would be a construction where an expletive *it* serves as the matrix subject of a *tough* predicate which embeds a gapless clause.

are consistent with Armbuster's data.

Crucially, nothing in Armbuster's data makes it clear that children were interpreting stimulus questions as grammatical. If children found any of Armbuster's stimuli ungrammatical, it is also possible that different children would have responded with different repair strategies. If the same repair strategy is not selected by all children, this might explain why only three quarters of the participant pool provided consistent results. Additionally, because Armbuster only asked for yes/no responses from children, children did not actually need to arrive at a grammatical parse of the target question before providing an answer. They could have simply recognized that they were being asked a yes/no question (e.g., from intonation or extrapolating from the practice items) and answered accordingly. The imbalance between answers—the fact that *yes* answers answers in Armbuster's study were significantly more common than *no* answers—could reflect the so-called "yes bias," or the fact that children prefer answering questions with *yes*, compared to *no*, in experimental settings, all else being equal (Blom and Unsworth 2010).

Anderson (2005): Much like Armbuster (1981), Anderson (2005) presented children with scenarios that were either consistent with the truth conditions of a subject or object gap *tough*-construction. However, instead of simply asking children to answer a yes/no question, Anderson also asked children to justify their responses to a target question. The justification component of Anderson's experiment makes results less likely to showcase artifacts of the "yes bias." In Anderson's experiment, 44 children, ages 3;4 to 7;5 participated in a truth value judgment task. Children were told a story which was simultaneously acted out with toys. A puppet then summarized the story using a symmetric *tough* construction. Stories set up contexts that were compatible with both the adult-like (matrix subject is associated with the embedded object gap) and non-adult-like (matrix subject is associated with the embedded subject gap) interpretations of each *tough* construction. However, for each interpretation, the truth conditions were flipped (e.g., the sentence would be true on an object *tough* interpretation but false on a subject *tough* interpretation). After answering the target question, children were asked to explain why they believed their answer was correct. An example of one of the stories presented is given below:

(12) *Toy props*: a pig, a dog, a cat, a soccer ball, a goal, and a slide.
Narrator: This is a story about a dog and a pig who are playing in the park.
Dog: I can teach you how to play football, pig. Would you like that?
Pig: But you're a dog, not a football player. How can you teach me how to play football?
Dog: Watch! I push the football with my nose. Then I run with the ball and push it into the fence. There! I scored a goal. Now you do the same thing, pig.
Pig: OK. Like this? (Pig pushes ball into fence with his nose.) Yea! I scored a goal too. Thanks for teaching me how to play football, dog. Now I'll teach you how to play on the slide. Just watch. You go up the steps like this and then slide down this end. Whee! Your turn, dog.
Dog: Like this? (The dog tries to go up the slide from the wrong end.)
Pig: No, dog, that's not right. Go round to the steps and try again.
Dog: (noticing cat) Hey, is that a cat over there? Forget about the slide, I'm going to chase that cat! (Dog runs after cat. Cat makes 'meow' noise.)
Puppet: I know what happened in that story. The dog was difficult to teach.

Because the dog in the story both taught something to the pig (how to play soccer) and was taught something by the pig (how to use the slide), the puppet's target question is compatible with either an adult-like or non-adult-like interpretation of the *tough* construction *'The dog was difficult to teach.'* On an adult-like interpretation, the sentence is *true* (because the pig wasn't able to teach the dog how to play on the slide). On a non-adult-like interpretation, the sentence is *false* (because the dog had no difficulty teaching the pig how to play soccer)[6]. Sample answer justifications from Anderson's data set (for the story in (12)) are given in (13). Note that Anderson never asked children if they believed the target *tough* construction was both true and false (i.e., had two possible interpretations), nor did any child ever volunteer that observation.

(13) a. Because the dog ran after the cat! (object interpretation; true)
b. Because the pig scored a goal! (subject interpretation; false)

While all children were able to give an answer explaining their truth value judgment, not all the answers were analyzable. Some children gave answers that did not sufficiently reference the context of the sentence, e.g., explaining that they

[6]Anderson's results come from comparing how many justifications align with a child's *yes/no* answer. Because Anderson's stimuli were balanced, there were an equal number of scenarios where the *yes* answer corresponded to an adult-like interpretation and a non-adult like interpretation. This allowed for the possibility that children might first try to interpret a stimulus as "true."

made their choice *'because he [the puppet] was wrong,'* or *'because he said it right.'*

Across 11 of the 12 contexts in Anderson's experiment, children's justifications supported their truth value judgments at a rate significantly above chance (81.39%), for both adult-like and non-adult-like interpretations of the *tough*-construction. Additionally, Anderson notes that, until age 6;5, children are essentially at-chance regarding whether they will assign an adult-like (46.43%) or a non-adult-like (53.57%) interpretation to the target *tough* construction. After age 6;5, children consistently (82.3%) assign an adult-like interpretation to *tough* constructions.

Anderson notes that a six year old with a 82.3% adult-like response-rate to symmetric *tough* constructions can be understood as having a fully adult-like comprehension of *tough* constructions. This behavior is not distinct from the behavior observed when Anderson ran her study on an adult control group. The adult control group only gave a "correct" justification for their interpretation of the stimulus 90% of the time. When asked to explain their answers for the remaining 10% of trials, adults gave explanations that did not reference the stimulus question. For example, an "incorrect" explanation to the story above would be (along the lines of) "False. The dog just got distracted." Though this is a reasonable response to the stimulus, Anderson would not mark it "correct" because it does not explicitly support either subject or object gap interpretation of the stimulus question.

Importantly, while the results of Anderson's own study are certainly compatible with the Anderson hypothesis, they also compatible with at least one version of Wexler's hypothesis. Particularly, if children do not have a standardized, consistent and grammatical structure that they can map *tough* constructions onto, it is possible that children are not actually answering the *tough* construction that they are asked. For example, it is possible that children are actually responding to some reparse or reanalysis of the stimulus, or that children are simply describing some set of truth conditions from the story and explaining whether or not these truth conditions are met.

Becker *et al* 2012: Becker, Estgarribia & Gylfadottir (2012) is important to the present study for two reasons. First, it was the first acquisition study to adopt a methodology from Naigles et al and show that it was possible to elicit the equivalent of a grammaticality judgment from children. Second, this experiment firmly established that children are able to distinguish *tough* constructions from control constructions (and, as a result, *tough*-predicates from control predicates).

The first point is important for understanding the methodology of the main experiment in this chapter. The second point is important for understanding how children judge the grammaticality of *tough* and raising constructions compared to control constructions. Becker *et al* was the first experiment to show that a methodology originally proposed by Naigles et al. (1995) to elicit grammaticality judgment from neurologically atypical adult participants could be extended to child language acquisition studies. Naigles *et al* (1995) proposed that the intersection of reaction times are accuracy rates could be used as a proxy for grammaticality judgments in adult participants. That is, when grammatical and ungrammatical constructions were compared, participants in Naigles's study showed a baseline accuracy rate and reaction time for grammatical constructions, and a deviant reaction time and accuracy rate for ungrammatical constructions. Becker *et al* (2012) showed that this methodology produced similar results in neurologically typical child participants, leading them to use reaction time and accuracy as a proxy for grammaticality judgments in their study, as well. While this methodology is not able to make predictions about the behavior of individual children, it does extend to children as a class. This study was also the first to show that it was possible to ask children adult-ungrammatical sentences as part of an acquisition experiment.

Specifically, Becker *et al* (2012) used this technique to investigate how children understood the theta-grids of *tough* and control constructions. For their study, 40 children, age 4;3 to 7;10 years old, watched a short video, which they were then asked about. Each child was asked two true/false questions, only one of which was adult grammatical. For each question, the answer and response time was measured. Based on the methodology of Naigles, *et al* (1995), Becker *et al* predicted that children would be slower to answer ungrammatical questions than their grammatical counterparts; Becker *at al* also predicted that children would have adult-like theta-grids for both *tough* and control constructions. Sample questions from Becker *et al* are given in (14) and (15). In (14), Becker *et al* predicted that children would respond in a way that indicated a grammatical judgment (*yes/no* answers consistent with adult answers and a reaction time consistent with a child's baseline reaction time) for (14a) and an ungrammatical judgment for (14b) (no consistent *yes/no* answer pattern and a deviant reaction time). The same predictions were made for the *tough* predicate stimuli in (15) (note that these examples involve *tough* predicates but not *tough* movement).

(14) a. Is it hard to move a piece of plastic?
b. *Is the nurse hard[7]?

(15) a. *Is it[8] afraid to fight the dinosaur?
b. Is the nurse afraid?

As predicted, children spent considerably more time (had a higher reaction time) answering adult ungrammatical questions, compared with adult grammatical questions, across adjective types. Across adult ungrammatical constructions, accuracy rates were also lower.

The experiment was also repeated using novel adjectives (e.g., "daxy" and "greppy"). These adjectives were presented to the child in a context meant to disambiguate whether they were *tough* or control adjectives[9]. Sample stimulus sentences for the novel adjectives were: *'An apple is very **daxy** to draw,'* and *'I'm sure Mr. Farmer would be **greppy** to help.'* Here, Becker *et al* predicted that children would be able to distinguish which novel adjectives were intended to be novel *tough* predicates and novel control predicates. Additionally, children were expected to be able to distinguish grammatical uses of the novel *tough* predicates (e.g., *daxy*) in *tough* contexts from ungrammatical uses of the novel *tough* predicates in control contexts (and visa versa for the novel control predicates, e.g, *greppy*).

Again, children were noticeably faster and more accurate when responding to adult grammatical *tough-* and control constructions, compared to adult ungrammatical *tough-* and control constructions, even when the *tough* and control predicates were novel. Throughout their experiment, Becker *et al* did not find these results to vary by age. This suggests that children have acquired at least some of the complex structure of *tough* constructions by age 4. Specifically, this data suggests that children understand that *tough*-predicates, unlike control predicates, do not assign a θ-role to their surface subjects, while control predicates do. This requires, at a minimum, understanding that control and *tough* adjectives have at least some properties that are distinct from those of control predicates.

[7]It was not clear to me from Bekcer *et al* what the expected "correct" response to the adult ungrammatical example in (14b) was. My best interpretation, given that they predicted that children would find (14b) ungrammatical, was that they predicted a mix of *yes/no* responses and predicted that *yes/no* response would not correlate with which video was shown.

[8]from Becker *et all* (2012:3): "N.B. *it*=expletive." How an expletive interpretation of the pronominal subject was enforced in the experiment is not stated

[9]This context included another video where scenes where acted out before stimulus sentences (e.g., *'An apple is very **daxy** to draw,' 'I'm sure Mr. Farmer would be **greppy** to help.'*) were presented to the child.

This also helps us understand how, if children find any *tough* movement constructions grammatical, *tough* constructions would be represented in the child grammar. Particularly, it shows that a hypothesis where children misinterpret *tough* constructions as control constructions is not tenable, as children understand that *tough* and control constructions are fundamentally different things.

Wexler (2012): Though it is not an experimental paper, Wexler (2012) provides a theory for *why* children do not demonstrate an adult-like performance on *tough* constructions until approximately age eight. Wexler 2012 proposes that children cannot understand *tough* constructions because of the way that they represent phasehood—essentially, Wexler (2012) adopts the *universal phase requirement*, defined in (16). Additionally, Wexler claims that *tough* construction must involve a defective phase which children cannot process. This leads their derivations to crash and results in children judging all *tough* constructions as ungrammatical.

(16) *Universal phase requirement* (Wexler 2004): Children (up to age 8) take *all* vP and CP to define phases, rendering passives, unaccusatives and (subject-to-subject) raising structures ungrammatical[10].

Importantly, Wexler's theory—that the universal phase requirement underlies the child ungrammaticality of *tough* constructions—requires A-movement theory of *tough* constructions, and is not compatible with a null operator theory of *tough* constructions. In Wexler (2004), the universal phase requirement is used to explain why children are late to acquire specific A-movement constructions (e.g., verbal passives or subject to subject raising over an experiencer), and Wexler (2012) states that, in order for the universal phase requirement hypothesis to take effect, *tough* constructions must involve and element that undergoes A-movement into the matrix subject position from within the embedded clause[11]. Without this

[10]Under a Wurmbrand-style contextually specified definition of phasehood, this could rephrases as roughly "the child grammar does not allow for defective phases."

[11]Specifically, Wexler adopts Hicks' (2009) theory of *tough* movement, where a complex operator, containing both a null operator and a phonologically overt DP, Ā-moves to the left edge of the embedded clause. For a full review of Hicks' analysis of *tough* movement, see Chapter 1 of this dissertation. Very briefly, in Hicks' system: the overt DP is sub-extracted from within a complex operator, allowing it to A-move into the matrix clause. This avoids A-movement chain that meets the technical definition of improper movement—the same element is not A-moved after Ā-movement since the DP contained within the operator is technically not the same thing as the operator itself–but it is unclear whether there is evidence for this kind of movement or complex operator elsewhere in the grammar. This is particularly true since this kind of sub-extraction is sharply ungrammatical in non-tough movement, as in (ii), compared to a baseline in (i).

(i) It seems that [Mary loves [pictures of President Obama]].

A-movement step, there would be no phase that children could fail to process. Thus, evidence for Wexler's theory of *tough* acquisition can potentially be understood as support for A-movement approach to *tough* constructions, and evidence against Wexler's theory could also potentially be understood as evidence against A-movement approach to *tough* constructions.

Importantly, Wexler (2012) does not make any predictions for the structures that children will utilize in processing ungrammatical *tough* constructions; Wexler (2012) is limited to proposing a mechanism to explain the delay in *tough* acquisition (which he also proposes explains the delay in the acquisition of raising). All Wexler (2012) predicts is that children will find all *tough* and all raising constructions ungrammatical. Similarly, Wexler (2012) does not discuss what strategies children will use in order to be able to respond to a child-ungrammatical question in an experimental setting.

The syntax of *tough*-acquisition

While *tough* constructions are often discussed within the theoretical syntax literature (cf. Hartman 2012, Hicks 2009, and the introduction to this dissertation), there is very little focus on the ungrammaticality of subject *tough* constructions. Instead, the bulk of the literature focuses on how to represent the movement path within object *tough*-constructions. Often, the ungrammaticality of subject *tough* constructions is reduced to the claim that DPs cannot Ā-move from a caseless position—and the subject position in a nonfinite embedded clause is not a case position. Evidence for this claim comes from the fact that overt nominals in embedded subject positions must be assigned case by a preposition in order to be licensed. However, as discussed in Chapter 3 of this dissertation, there are some examples that suggest that the impossibility of subject *tough* constructions reduces to something more complicated than a general restriction on Ā-movement from a non-case position. Note, in (17), that the embedded subject cannot undergo *tough* movement even after being assigned case by a preposition[12]. Chapter 3 proposes that the ungrammaticality of subject *tough*-constructions is the result of a spec-to-spec anti-locality constraint that restricts short subject movement from spec-TP to spec-CP.

(ii) *Which president seems [[pictures of _] for Mary to love _?]

[12]Note that stranding a case-assigning preposition after Ā-movement in English is normally acceptable for non-subjects, as in *Who did Mary talk to* _?

(17) *Who(m) is it possible for _ to see Mary?

As noted in Chapters 1, 2 and 3 of this dissertation, there are two competing theories about the structure of *tough* constructions in the literature. Under one theory, *tough* movement is an example of null operator movement (c.f., Chomsky 1977, among others). According to this theory, *tough* constructions involve a null operator that Ā-moves from the embedded object position to the left edge of the embedded clause. After Ā-movement, the null operator is then semantically bound to the matrix subject, as in (18). See Chapter 2 Section 4.2 of this dissertation for an overview of this specific analysis.

(18) Null operator *tough* construction structure
[$_{DP}$ Jonathan Franzen]$_i$ is tough [$_{CP}$ OP$_i$ for [$_{TP}$ Anneke to think about <OP$_i$>]]

Under the second theory *tough* constructions involve a single overt DP which moves, in multiple steps, from the embedded object position to the matrix subject position (c.f. Hartman 2012, among others). Under this analysis, the overt embedded object DP first Ā-moves to the left edge of the embedded clause before A-moving to the matrix subject position, as illustrated in (19). This sequence of movement operations is known as "improper movement," so-called because it is often ungrammatical. This is illustrated in (20), where an element first Ā-moves to the left-edge of the embedded clause and then A-moves into the matrix clause. A key question for this the improper movement theory of *tough* movement is why improper movement is possible in *tough* constructions when it is not allowed in structures like (20).

(19) **Improper movement *tough* construction structure**
[$_{DP}$ Jonathan Franzen]$_i$ is tough [$_{CP}$ <DP$_i$> for [$_{TP}$ Anneke to think about <DP$_i$>]]

(20) **Improper movement is not generally grammatical**
[$_{DP}$ Which guest]$_i$ was believed [$_{CP}$ <DP$_i$> that [$_{TP}$ Anneke thought about <DP$_i$>]]

4.2.3 Raising literature

Theoretical literature

Because the analysis of raising is not hotly controversial, this section will largely dedicate itself to discussing the acquisition literature. To briefly review the theoretical analysis of raising: raising predicates are clausal embedding predicates that assign a theta-role to the clause that they embed, but not the argument that appears in their matrix subject position. Raising predicates can embed either finite or non-finite clauses. When raising predicates embed non-finite clauses, however, the subject of the non-finite embedded clause must A-move to the matrix subject position in order to receive Case. This movement is raising.

Because raising is an instance of A-movement, it is subject to relativized minimality (Rizzi 1990). Relativized minimality, which bans any movement operation that would require one argument to A-move "across" another, is standardly assumed to explain why object raising constructions are adult-ungrammatical. The grammatical subject raising example in (21a) does not violate relativized minimality. The ungrammatical raising example in (21b), however, does involves the embedded object moving "across" the embedded subject (which receives case from the preposition *for*), a relativized minimality violation[13].

(21) a. Jonathan Franzen is likely [_ to hate himself]
b. *Jonathan Franzen is likely [for [$_{DP}$ **Anneke**] to hate _]

Within the theoretical raising literature, however, there is one area of controversy. Certain English raising predicates, like *seem*, can license an optional experiencer argument, as in (22a). This experiencer argument is proposed to intervene between the matrix subject position and the subject of the embedded clause. Because of this, Relativized Minimality predicts that the subject of the embedded clause will not be able to raise to the matrix subject position[14]. This prediction is

[13]The example in (21b) could also be considered ungrammatical because it is A-movement from a case-marked position.

[14]The English facts are actually slightly more complicated than this, as the experiencer DP does not technically c-command the embedded subject. However, the experiencer does *act* as though it c-commands in, for instance, binding examples like (i).

(i) *It seems to her$_i$ that Anneke$_i$ will hate Jonathan Franzen.

Additionally, these same experiencers *do* case defective intervention effects when licensed by other English raising predicates, as in (ii), and in other languages (c.f. Hartman 2012 for an overview). For

not borne out (22b).

(22) a. It seems to Anneke that Jonathan Franzen is a terrible writer.
b. Jonathan Franzen seems to Anneke to be a terrible writer.
(predicted * by relativized minimality)

Interestingly, while certain English raising predicates allow this apparent relativized minimality violation, this is not a robust cross-linguistic effect. For example, Romance languages like Italian and French famously do not allow any raising to take place across an oblique experiencer, as shown in (23). Why some English raising verbs appear to allow these kinds of relativized minimality violation is a subject of debate within the raising literature, and largely tangential to the current experiment.

(23) a. Il semble (au garçn) qu'elle a du talent
EXPL seem (to.the boy) that-she has of-the talent
'It seems to the boy that she has talent.'
b. Elle semble (*au garçn) avoir du talent
She seems (*to.the boy) to.have of-the talent
'She seems to the boy to have talent.'
(French, Hartman 2013 citingMcGinnis 2003)
c. Sembra (a Maria) che Gianni é stanco
seems (to Maria) that Gianni is tired
'It seems to Marie that Gianni is tired.'
d. Gianni sembra (*a Maria) essere stanco
Gianni seems (*to Maria) to.be tired
'Gianni seems to Maria to be tired.'
(Italian, Hartman 2013 citing Boeckx 2008)

Raising and A-movement Acquisition literature

Oftitelli (2012) notes that there is a dichotomy within the acquisition of A-movement: children do not show any delays in acquiring *some* types of A-movement but show significant delays in acquiring other types of A-movement. For example, children are not delayed in acquiring the ability to move subjects from their thematic position within the verbal domain to the

these reasons, I will assume that these experiencers are still expected to cause defective intervention effects.

(ii) a. It appears [to me] [that Anneke hates Jonathan Franzen]
b. *Anneke used/happened/was (un)likely [to me] [_ to hate Jonathan Franzen]

canonical subject position, spec-TP (Stromswold 1996). This suggests that children do not have any general difficulty acquiring A-movement operators. However, specific A-movement constructions do show acquisition delays. For example, children are famously delayed in their acquisition of verbal passives (Slobin 1966).

The body of literature surrounding the acquisition of raising is not as broad as the body of literature surrounding the acquisition of passives (Orfitelli 2012), but there is still dissent within the acquisition of raising literature. In particular, some studies (e.g., Becker 2006) present evidence that children do not show any delay in the acquisition of raising, and report consistently adult-like performance on raising constructions from as early as age three. Other studies, however, show that children have a general delay in acquiring raising (Hirsch, Orfitelli & Wexler 2008). Finally, a third group of studies (e.g., Orfitelli 2012) suggest that children only struggle with raising constructions that involve an oblique experiencer, such as (22b). Though the raising acquisition literature is largely limited to English, there is some additional evidence that children are delayed in their acquisition of raising in other languages, as well (see, e.g., Koring 2007 for Dutch). Within the portion of the literature that argues that children do not have an adult-like representation of raising, there is no real consensus about why children's acquisition of raising is delayed.

Three raising acquisition studies are particularly relevant to the current study: Becker (2006), Hirsch, Orfitelli & Wexler (2008) and Orfitelli (2012). The first of these studies, Becker (2006), argues that children have acquired raising by approximately age four. Becker's study involved a truth value judgment task which asked children whether a certain scenario was compatible with a raising constructions that was used to describe it. For example, one scenario involved putting a toy dog with white fur under a purple lamp and asking children "*Does the dog seem purple?*" Because children had previously seen that the dog's fur was white, Becker predicted that children would only answer "yes" if they understood the stimulus as a raising constructions. Essentially, Becker claimed that the only way children could rectify the facts that the dog *had* purple fur that *looked* purple was to analyze the stimulus question as a raising construction. This prediction was borne out, and children in Becker's study were able to successfully differentiate whether the dog was or seemed to be purple. This result suggests that children have an adult-like representation of raising constructions.

Hirsch, Orfitelli & Wexler (2008) propose that Becker's results were

experimental artifacts. Particularly, Hirsch, Orfitelli & Wexlers propose that if children "ignore the raising verb entirely and interpret the sentence as copular, incorrectly answering 'false' to 'The dog seemed to be purple"' if they do not represent raising constructions in an adult manner (Orfitelli 2012). This hypothesis is based on the fact that the English copular is ambiguous between a stage and individual level interpretation. Thus it is possible that children were actually interpreting the stage individual-level copular sentence 'The dog is (always) purple' as an individual level copular constructions meaning 'The dog is (currently) purple'. If children were performing this re-analysis, their behavior in Becker (2006) would still be predicted.

Orfitelli (2012) suggests that Becker's results were, in fact, representative of how children understand raising. Over the course of seven different experimental studies, Orfitelli proposes that children have acquired A-movement, and so can also raise subjects. However, Ortifelli also shows that children are not able to raise a subject argument over an intervening oblique experiencer—or, at the very least, children do not show adult-like performance on these constructions. Orfitelli believes that this delay is due to to the fact that children have, at this point, not learned that relativized minimality violations are possible. In other words, Orfitelli proposes that the child grammar is too restrictive to allow for the relativized minimality violation required in these types of raising constructions. However, another interpretation is possible: children could be mis-analyzing the oblique experiencer as the subject of the embedded clause and interpreting the subject raising constructions as object raising constructions. If children were analyzing raising constructions this way, we would also not predict them to show adult-like behavior.

4.2.4 Combining the syntax and acquisition literatures

Looking at the experimental literature in combination with the theoretical literature helps us understand which syntactic operations children have already acquired, and which operations children systematically struggle with. Importantly, most individual acquisition experiments are only designed to distinguish whether or not children have an adult-like understanding of a particular construction. If the results of a study suggest that children can, in fact, respond to a particular construction in an adult-like way, we can assume that they are able to represent that construction the same way that adults do. However, if

children provide non-adult-like responses, we do not have any information about the structures that they are accessing when they parse a construction.

This is largely a result of experimental design: during experiments, children are normally only asked yes or no questions. If a child's yes/no responses match the expected adult response, we can assume that the child is using the same syntax as an adult. Because we have access to adult interpretations, we can also extrapolate a child's interpretation. However, if a child does not give an adult like response, we don't have a lot of additional information. Particularly, we cannot know from accuracy data alone whether children show non-adult-like performance because (a) they have access to some structural representation that adults do not or (b) they do not have access to any structural representation, and so are answering the stimulus question either randomly or making use of some repair strategy.

This is largely the difference between the Anderson and Wexler hypotheses. Anderson proposes that children show non-adult-like performance on *tough*-constructions because they believe that symmetric *tough*-constructions are ambiguous. This hypothesis predicts that children will show a non-adult-like understanding of *tough*-constructions. This chapter has extended Anderson's predictions, and proposes that, if Anderson's hypothesis is correct, it also predicts how children will understand the semantics of symmetric *tough*-constructions. As a result, Anderson's hypothesis, I believe, also predicts the (type of) syntactic representation children would need to possess to be able to understand symmetric *tough* constructions as ambiguous[15]. Wexler proposes that children show non-adult-like performance on *tough*-constructions because they find all *tough*-constructions ungrammatical. This hypothesis also predicts that children will not show adult-like behavior. Wexler does not propose *how* children will interpret *tough*-constructions, or what structures they might rely on when trying to answer stimulus questions about an ungrammatical question.

The previous sections reviewed the theoretical and experimental literature surrounding *tough* and raising constructions. This section will look at which theoretical tools children appear to have access to, given the acquisition literature. Assembling a better picture of which structures children can represent in an adult-like way will help us better understand what the child grammar might look like. This, in turn, will also help us understand how to interpret child response

[15]Particularly, a syntax where *tough*-movement is possible from either the embedded subject position or the embedded object position

data, which might initially appear ambiguous between a grammatical judgment of an ambiguous structure and an ungrammatical judgment.

As discussed in the introduction to this chapter, this step may be able to help us distinguish between models of *tough* movement. The remainder of this section will discuss this claim in more detail. The core claim of the remainder of this section is the following: Under the null operator theory, *tough*-movement involves being able to do the following syntactic operations: allow a silent null operator argument to function as the object of an embedded clause, Ā-move that element to the left-edge of the embedded clause, and semantically bind that null operator to the matrix subject. Under the overt DP theory, *tough*-movement involves being able to do the following syntactic operations: Ā-move an overt embedded object to the left-edge of the embedded clause and then A-move the same argument into the matrix clause from the embedded clause. Because each theory of *tough*-movement requires a slightly different syntactic skill set, each theory makes slightly different predictions in term of child language acquisition. Particularly, the literature predicts that children will find *tough* constructions grammatical under a null operator analysis. The predictions for the improper movement analysis are slightly less straightforward, however, as there have not been any independent acquisition studies regarding how children acquire improper movement restrictions generally.

Null operator hypothesis predictions

In order for children to be able to represent a null operator *tough* movement structure, they would need to be able to do the following tasks: Ā-move a null operator within the embedded clause and then bind that null operator to the matrix subject. Given the skills already available to children, children are predicted to find *tough* constructions grammatical under a null operator theory of *tough* movement.

This section will focus on the acquisition of Ā and null operator constructions. Broadly: children have acquired both Ā-movement and null operator binding before they have acquired an adult-like representation of *tough* movement. Specifically: Children acquire Ā-movement operations later than they acquire Control constructions (Friedmann and Lavi 2006), but these structures are still acquired before *tough* constructions. For example, children are able to produce adultlike relative clauses by age 5 (Friedmann & Lavi 2006, citing de Villiers et al. 1994, Friedmann and Novogrodsky 2004 and McKee et al. 1998, a.o.), with some

studies showing that children begin producing relative clauses as early as 2;6 (Friedmann & Lavi 2006, citing Berman 1997, Crain et al. 1990, a.o.). Wexler (1992) has argued that a child's slight delay in acquiring relative clauses and other Ā-movement constructions involves a difficulty moving an argument to a non-argument position.

Many theories analyze relative clauses as null operator constructions (e.g., Chomsky 1993 for a theoretical analysis, Klein 1993 for an acquisition analysis). That children can construct and understand relative clauses before they demonstrate a mastery of *tough* movement suggests that children acquire the ability to Ā-move and bind null operator arguments before they acquire *tough* movement.

This claim is supported by examining how children acquire *wh*-movement, relative to A-movement. Friedmann & Lavi compared to order of acquisition of unaccusative A-movement to three types Ā-movement operations: topicalization, subject relative clauses and object relative clauses. While their results show that unaccusative A-movement is acquired before the Ā-movement operations, 71% of Friedmann & Lavi's participants, all younger than 3;10 (and some as young as 2;2) showed adult-like production of the three Ā-movement structures. These results show that while the acquisition of Ā-movement is slightly delayed, it is still acquired before *tough* constructions.

Given this, under a null operator theory of *tough* movement, we predict that children will find subject and object *tough* constructions grammatical. This is because child have already acquired the skills needed to bind both subject and object null operators in relative clause constructions (Friedmann & Lavi), as well as more general Ā-movement operations. This proposal is not particularly surprising when the control data is considered. Recall that children acquire control constructions much earlier than they acquire *tough* constructions. This suggests that children have acquired the general ability to bind covert elements to overt DP before they have acquired *tough* constructions.

Improper movement predictions

The predictions of the improper movement theory of *tough* movement are less straightforward. In order for children to be able to represent an improper movement *tough* structure, they would need to already have acquired A and Ā-movement operations. In addition, they would need to be able to A-move an argument that has already undergone Ā-movement. There is, to the best of my

knowledge, no literature specifically dedicated to whether or not children believe that improper movement is possible within the grammar.

Because of this, either child grammaticality judgment could be taken as support for an improper movement analysis of *tough* constructions. If the child grammar is *too flexible*—that is, if the child grammar has fewer restrictions than the adult grammar, and additional restrictions are learned over the course of acquisition—we would predict that children would not yet have acquired a general ban on improper movement (to which *tough* movement would need to be an exception), and so would find both *tough* constructions grammatical. If the child grammar is *too restrictive*—that is, if the child grammar have more restriction than the adult grammar, and exceptions to these rules are learned over the course of acquisition—we would predict that children would not yet have learned that *tough* movement is an exception to a larger ban on improper movement. This would predict that children will find *tough* construction ungrammatical. If the child grammar is not uniformly too flexible or too restrictive—if acquisition involves learning a mix of new restrictions and exceptions—we cannot make a prediction regarding how children will judge *tough* constructions unless more research is done into how children conceive of improper movement more generally.

For this reason, if children judge *tough* construction as grammatical, we cannot use the results of this study to differentiate between a null operator and improper movement approach to *tough* movement. This is because a grammatical judgment is potentially predicted by each approach. If children, however, find *tough* constructions ungrammatical, we can take that as evidence *against* a null operator analysis of *tough* movement. This is not the same as an argument for the improper movement analysis[16], but it does suggest an avenue for future research.

[16]This is particularly true given that the structure of *tough* constructions might be additionally complex, even under a null operator analysis. See Chapter 1 of this dissertation for a more thorough discussion of the complications of both a null operator and improper movement approach to *tough* movement.

4.3 Experiment

This experiment used reaction time and expected response data[17] to determine children's grammaticality judgments for subject and object *tough* and raising constructions. To do this, children were told a story about two characters (Jim and Julia) who were playing a game. In the story, one character was all essentially guaranteed to win the game (e.g., *Jim and Julia are racing each other! Julia has always been faster than Jim, and yesterday Jim hurt his ankle. Today he can barely run at all.*). While the children heard the story, they were shown an image that depicted the same scenario (e.g., Julia about to cross the finish line on a track while Jim, far behind her, lay on the track in pain). The child was then asked a question about the story/image pair—either a baseline question involving an embedded tough or raising predicate with an expletive matrix subject and no gap site in the embedded clause (e.g., *Will it be tough for Jim to beat Julia?*) or a stimulus question with either a subject or an object gap (e.g., *Will Julia be tough for Jim to beat?*). All of the baseline questions were adult grammatical, while only half of the test questions were adult grammatical. The child's response ("yes" or "no") and their reaction time were measured. Following Becker *et al* (2012) and Naigles *et al* (1995), this experiment used the combination of expected response rate and reaction time as a proxy for grammaticality judgments. This allowed the experiment to investigate whether (i) children have an adult-like understanding of *tough* and raising constructions and (ii) whether children have an adult-like understanding of restrictions on object raising and subject *tough* movement.

4.3.1 Participants

Twenty eight typically developing children between the ages of 4:5 to 7:7 (mean age: 5:6) were included in the final sample. There were 13 females and 15 males. To ensure that children in the final sample had acquired all the necessary vocabulary items and were able to process biclausal sentences, children were

[17] Because this study asks children to answer questions that are adult ungrammatical, "accuracy" (the term used in most acquisition studies that solicit yes/no questions from children) is a somewhat inappropriate term. It is not clear that there is an accurate answer to a question that is not well-formed. If children find an adult-grammatical question ungrammatical, it is also not clear that is it reasonable to assume that an accurate question truly exists for the child. For this reason, this study uses the term expected response in place of accuracy. Expected response is defined as the expected adult answer to a question that is adult grammatical, and the expected adult response to the closest grammatical rephrasing of an adult-ungrammatical question, if such a representation exists.

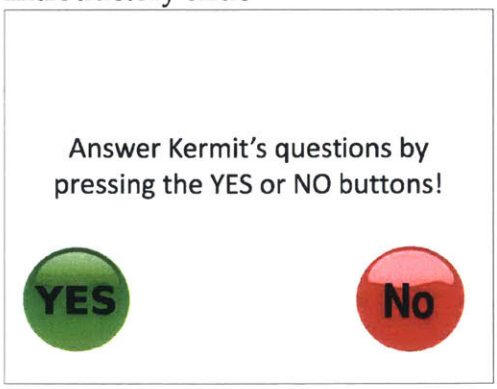
asked a total of 8 baseline questions. Four baseline questions included an embedded *tough*-predicate with an expletive matrix subject (e.g., *Will it be tough for Jim to beat Julia?*) and four included an embedded raising predicate and an expletive matrix subject (e.g., *Is it likely for Jim to beat Julia?*). To be included in the final sample, children needed to answer at least 5/8 of the baseline questions correctly, missing no more than 2 baseline questions in either category. Thirty three children participated in the study, and five were excluded by this metric. The mean age for female participants was 5:4 and the mean age for male participants was 5:8. Nine participants were younger than 5:0, eleven participants were between 5:0 and 6:0, and eight participants were older than 6:0. Children were recruited from Boston, Massachusetts and were acquiring English as their native language, with less than 50% exposure to another language.

4.3.2 Materials

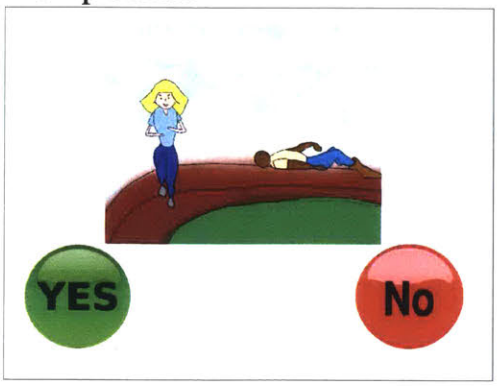
Visual stimuli

Children were shown a powerpoint presentation consisting of 17 slides. The first slide, shown in (24), introduced the children to the experiment. The remaining 16 slides each contained an image featuring two characters, Jim and Julia, playing a different game. Each slide also contained a green "YES" button and a red "NO" button, which played a recording of the words *yes* and *no*, respectively, when pressed. Children were asked to press the buttons in order to answer questions about the images. An example of one of the 16 image slides is given in (25). For a full inventory of images used in this project, see Appendix A. All images were accompanied by auditory stimuli.

(24) **Introductory slide**



(25) **Example slide**



Auditory stimuli

Each story was accompanied by a story about two characters, Jim and Julia, which a researcher read to the child. The child was then asked a yes/no question about the story/image pair. Though different researchers conducted the experiment at different locations across Boston, all researchers were native speakers of American English with American accents. A sample story is given in (26), and the full list of possible questions that could be asked about that story is given in (27). For a full list of stories and question matrices used in this project, see Appendix B.

(26) **Example story**

Jim and Julia are racing each other. Julia has always been faster than Jim. Yesterday, Jim was practicing for the race, and he hurt his ankle. Today, he's running even slower than usual!

(27) **Question Matrix**

Condition	Yes-response expected	No-response expected
Raising baseline	Is it likely for Julia to beat Jim?	Is it likely for Jim to beat Julia?
Raising subject	Is Julia likely to beat Jim?	Is Jim likely to beat Julia?
Raising object	*Is Jim likely for Julia to beat?	*Is Julia likely for Jim to beat?
Tough baseline	Will it be easy for Julia to beat Jim?	Will it be easy for Jim to beat Julia?
Tough subject	*Will Julia be easy to beat Jim?	*Will Jim be easy to beat Julia?
Tough object	Will Jim be easy for Julia to beat?	Will Julia be easy for Jim to beat?

Each story had twelve possible questions associated with it: a set of six "yes" questions—questions for which the expected answer to the question was "yes"—and six "no questions"— questions for which the expected answer was

"no". Each set of questions (the "yes" set and the "no" set) consisted of a baseline (no movement), subject- and object-movement for each verb condition (*tough* and raising predicates). Each set of six questions contained two adult-ungrammatical questions (one subject *tough* construction and one object raising construction). Expected response for adult-ungrammatical questions was determined by the truth conditions of the corresponding adult grammatical baseline question (e.g., the expected answer to **Is Jim likely for Julia to beat?* corresponded to the answer to the question *Is it likely for Julia to beat Jim?*).

Apparatus and procedure

Apparatus: The experiment was presented to the children on a MacBook Pro laptop computer. Children were seated approximately 1 foot from the MacBook screen and were asked to indicate their answers to yes/no questions by using two interactive buttons on each powerpoint slide. Both buttons emitted recordings of a male American-English speaker. Children were asked to use the buttons, instead of saying *yes* or *no*, because many children (particularly children under 5) chose to answer the yes/no questions non-verbally if possible (e.g., by shaking or nodding their head).

Procedure: Two researchers were present for each run of the experiment. One researcher spoke to the child's legal guardian (explaining the experiment, providing the consent form), while the other researcher introduced themselves to the child. Researchers introduced the child to the laptop and the powerpoint presentation, including the response-buttons. Children were also introduced to a Kermit the Frog puppet, with the following script:

(28) **Introductory script**

Thanks for agreeing to play a game with us! [Researcher shows the child Kermit] This is my friend—do you know who it is? [Waits for child's response] It's Kermit the Frog! Here's the game we're going to play. I'm going to show you and Kermit some pictures, and then tell you some stories. I need you to listen to the stories, ok? See, Kermit's not a very good listener—I bet you're a really good listener! Because Kermit's not a very good listener, sometimes he'll ask you questions about the stories you just heard. Do you think you can answer Kermit's questions?

If the child indicated that they were willing to play the game, the researcher had

the Kermit puppet ask the children several simple yes-or-no questions (*Are you 5 years old?, Do you have any siblings?*) to familiarize the child with the experimental set up. The child was instructed to answer Kermit's questions by pressing the response buttons. No child needed more than three training questions in order to become comfortable using the response buttons and the laptop. During the experiment, children were presented with 16 story/image pairs. For each story/image pair, the researcher preceded the target question by checking in with the child to ensure they understood the story. The check-in script was as follows:

(29) **Check-in script**

a. Ok! Did you understand that story? (If the child indicated that they did not understand the story, the researcher could repeat the story up to one time.)

b. Kermit, do you have any questions about the story?

Kermit would then ask the child the target question (e.g., **Will Jim be easy to beat Julia?*). The child would answer the question by pressing one of the response buttons. Regardless of the child's answer, after each response Kermit would thank the child for helping him by telling the child: *Ok! Thanks! Now I understand.*

Each child was asked 16 stimulus questions: eight questions contained a matrix raising predicate and eight questions contained a matrix *tough* predicate. For each predicate type, children were asked four baseline (no movement) questions and four movement questions. The raising and *tough* predicate pairs chosen for this experiment had roughly analogous semantics (e.g., *likely* and *easy*; *unlikely* and *tough*). Additionally, raising/*tough* pairs were matched with individual stories (e.g., the *race* story was always paired with the predicates *likely* and *easy*). For a full list of raising/*tough* predicate pairs and the stories those predicates were paired with, see Appendix B.

After every other story, Kermit also asked the child a filler question. These filler question were yes/no questions about the image depicted on the slide (e.g., *Is Julia wearing a red shirt?*). Each image/story pair had a corresponding filler question associated with it.

Subject & Object Participant Groups: Participants were randomly divided into subject and object groups. Subject group participants were asked 16 total questions, eight baseline questions (four raising baseline questions and four *tough* baseline questions) and eight subject gap questions—four subject *tough* questions

and four subject raising questions. The four subject *tough* questions were adult-ungrammatical. Because participants were randomly assigned to one of these groups, individual participants saw both raising and *tough* constructions, but only one type of movement (either subject or object movement). All children were asked 12 adult-grammatical stimulus questions and four adult-ungrammatical stimulus questions. All filler questions were adult-grammatical.

Blocking within Participant Groups: The *tough* and raising conditions were blocked within a trial—for example, a child would be asked all eight raising constructions questions (four baseline questions and four gap stimuli) first, followed by all eight *tough* construction questions (four baseline questions and four gap stimuli). Blocking was done to prevent children from generalizing the structure of an individual *tough* or raising predicate across multiple predicates (e.g., to prevent children from assuming that, because *easy* and *likely* have similar meanings, they might also have similar structures). Across both *tough* and raising blocks, the baseline questions were always asked before the movement questions. Block order was randomized across participants.

Stories, Questions, & Predicates: Stories/image pairs were divided into four groups, classified by the *tough*/raising predicate that corresponded with the story/image pair. For each predicate pair, the *tough* and the raising predicate had similar semantics. Children encountered each *tough* and raising predicate twice during the course of the experiment; once in a baseline question and once in a gap question. Children were asked one item per condition. For each story/image pair, the child was asked a question from the question grid associated with the story/image pair (cf. (27)). All experimental factors—expected response (yes/no, 2 each per 4 question condition), subject/object blocking, which predicate type the child was first exposed to (*tough* or raising), story order—were pseudo-randomized and counter-balanced across trials. For a sample experimental script, see Appendix A.

4.3.3 Coding

Two researchers reviewed the audio recordings of each trial, and recorded the child's response ("yes" or "no") and reaction time (in milliseconds) for each question. For each trial, the age (in months) and gender of the child was also

recorded, along with information about that particular trial (whether the child was in the subject or object group, which predicate type the child saw first).

Reaction time data was coded in milliseconds. Expected response data was treated as a binomial. Expected responses from children were given the score 1, and unexpected responses were given the score 0. For all adult-grammatical questions, the expected response was the response that aligned with an adult interpretation of the sentence. For all adult-ungrammatical sentences, the expected response was the response that aligned with the truth conditions of the corresponding baseline question. Because gap site was the locus of ungrammaticality in all adult-ungrammatical conditions, this was the least biased option for determining expected responses for adult-ungrammatical constructions.

In the statistical analysis, gap type (no gap, subject gap, object gap) was Helmert coded. This allowed the model to make two comparisons: (i) baseline condition to all movement conditions (regardless of gap type or gap site) and (ii) subject and object movement within movement conditions. Predicate type (*tough* vs raising) was sum coded, with *tough* predicates assigned -1 (and raising predicates assigned 1).

4.3.4 Predictions

As discussed in Section 2, non-adult like child responses to *tough* and raising constructions are consistent with both syntactic ambiguity and ungrammaticality. The differences between the predictions of the Anderson and Wexler hypotheses reflect this. Anderson, who believes that children find symmetric *tough* constructions ambiguous between two possible interpretations, predicts that children will find all movement constructions grammatical. Wexler, however, who believes that children find symmetric *tough* constructions ungrammatical, predicts that children will find none of the movement conditions grammatical.

This section focuses on understanding what grammatical and ungrammatical judgments are predicted to look like in this study. The results of the experiment show that children find both subject and object raising constructions grammatical, but also that children find both subject and object *tough* constructions ungrammatical. This result is not predicted by either Anderson or Wexler. Taken as a whole, the data does not align with either theory. Taken by itself, the raising data aligns with Anderson's predictions and the *tough* data, when taken by itself,

aligns with Wexler's predictions. However, it is not clear that either subset of the results should be considered in isolation.

Grammaticality and Anderson's predictions

The current methodology makes clear predictions for what a child grammatical judgment would look like. Particularly, it predicts that there should be no significant difference between the ways that children respond to subject gap, object gap and baseline questions within the experiment. Children who successfully complete the baseline task should be able to accurately answer all movement constructions with the same (high) degree if accuracy, regardless of adult grammaticality.

If children find all movement constructions grammatical, there should not be a statistically significant different between subject and object gap constructions across predicate type. Additionally, reaction times for all movement constructions should be "fast," meaning that they should not be significantly different from baseline reaction time (abstracting away from any process cost required from movement more generally; a cost which would be predicted to be apparent across both subject and object gap conditions). A summary of these predictions is given in (30).

(30) Anderson's predictions (for children who pass the baseline condition)

Condition	Expected Response Rate	Predicted RT
Baseline	very high (80-100%)	baseline reaction time
Subject	very high (80-100%) no significant different from baseline	baseline (+ movement cost) no significant different from object
Object	very high (80-100%) no significant different from baseline	baseline (+ movement cost) no significant different from subject

If children find all *tough* and raising constructions grammatical, they must also be able to map those constructions onto a syntactic structure. One possibility for how this happens is as follows: if children find both subject and object *tough* constructions are grammatical, they might believe that both subject and objects can be *tough* moved. Essentially, children may have an adult-like representation of object *tough* movement and believe (unlike adults) that subject can move the same way. Similarly, if a child finds both subject and object raising constructions

grammatical, they may have an adult-like representation of subject raising and believe (unlike adults) that objects can also be raised. If children find all four constructions grammatical, additionally, it may be the case that child mis-analyze *tough* constructions as raising constructions (or raising constructions as *tough* constructions) and so assume a single syntactic representation for both predicate types.

If children find both subject and object *tough* constructions grammatical, we can also assume that children find symmetric *tough* constructions like (6) grammatical as well. Particularly, we can assume that children are able to access the adult-like object *tough* interpretation (assuming a null PRO subject in the embedded clause). We would not, however, have evidence that children are able to access a subject-gap interpretation, as proposed by Anderson (2012) and outlined in (31). This is because the example in (31) involves an additional complication: a transitive verb that lacks an overt object. Because we do not know how children judge transitive sentences that lack an overt object, we don't know how children would judge examples like (31).

(31) The pig_1 is tough [<DP> to teach (PRO)]
Interpretation: The pig has difficulty teaching

Similarly, if children find both subject and object raising constructions grammatical, we predict that children would find adult-ungrammatical symmetric raising constructions grammatical and would at least be able to access an object gap interpretation. However, we do not know whether or not children would find adult-ungrammatical symmetric raising constructions ambiguous between a subject and object gap interpretation.

Ungrammaticality and Wexler's predictions

The current methodology makes slightly less clear predictions about how children will behave when making ungrammatical judgments. This is because, unlike grammaticality, which aligns with only set of behaviors, ungrammaticality can manifest itself in a number of ways. The possibilities are as follows: A child might answer ungrammatical questions quickly, but with a low expected response rate. This behavior might arise if a child is unable to parse a given structure and simultaneously believes that the most important part of the task is to respond quickly to the test question, leading them to make a random guess. A

child might also respond to an ungrammatical question slowly but with a high expected response rate. This behavior might arise if the child is not responding to the test question itself, but is instead answering a grammatical reparse or rephrasing of the stimulus which happens to have the same truth conditions as the adult-grammatical (equivalent of the) test question. Finally, a child might respond slowly (i.e., high reaction time) with a low expected response rate. This behavior might arise if the child first tries to understand the test question and then fails to arrive at any possible reparse (and so makes a random guess) or arrives as a reparse that does have the same truth conditions as the adult-grammatical (equivalent of the) test question. Importantly, all possible ungrammatical judgment behaviors are distinct from grammatical judgment behaviors. Particularly, children are not expected to respond quickly and with a high expected response rate to test questions that they find ungrammatical. The possible reaction time and expected response rate results for an ungrammatical judgments are summarized in (32).

(32) Ungrammatical judgment behavior[18]

	Possibility 1	Possibility 2	Possibility 3
movement condition RT compared to baseline	no significant different	significant difference significantly longer RT	significant different significantly longer RT
movement condition expected response rate	at-chance	at-chance	at change
behavior	guessing no (re)analysis	reanalysis fails/does not match adult interpretation	reanalysis matches adult interpretation

It is not expected that all children will show the same ungrammatical judgment behavior. That is, different children are expected to use different repair strategies to respond to the same ungrammatical questions. This methodology does not, and cannot, predict how an individual child will respond to an ungrammatical question, however it can predict how children will respond to ungrammatical questions *as a class*. Overall, given a truly random sample, children should show lower expected response rates and higher reaction times when responding to ungrammatical questions, compared to grammatical baselines.

Wexler's hypothesis predicts that children will find all movement constructions (subject *tough*, object *tough*, subject raising, object raising) ungrammatical. If children find a construction ungrammatical, we don't have reliable information about which interpretations children can assign to the construction, or what syntactic structures a child is accessing in order to produce a response. If, as Wexler predicts, children find all movement constructions ungrammatical, however, we can also assume that children find (adult grammatical) symmetric *tough* constructions and (adult ungrammatical) symmetric raising constructions ungrammatical as well.

[18]One logical possibility, a very fast (low) reaction time and a high expected response rate is excluded from this chart. This is because movement has been shown to incur a processing cost (e.g., the scrambling effect, Tamoaka *et al*) and so movement constructions are predicted to have longer reaction times compared to non-movement constructions. Additionally, studies like Becker *et al* (2012) show that children can process biclausal constructions containing a *tough* predicate before they have acquired *tough* movement. Thus, we do not expect that any child who finds the baseline questions ungrammatical with find the movement constructions grammatical

4.4 Results

Overall results, for both expected response rates and reaction time, are summarized in Figure 1. This figure shows the mean expected response rates (for both "yes" and "no" answers) and reaction time for each condition as well as the delta reaction time—the difference between the mean baseline reaction time and the mean movement condition reaction time—for each movement condition. Rows that correspond to adult ungrammatical constructions are in **bold**.

Figure 1: Overall Accuracy and RT Results

Gap	Predicate Type	Accuracy (%)	Mean RT (sec)	Δ RT (sec)
None	Raising	85.71	3.034745	n/a
None	Tough	85.11	2.412840	n/a
Subject	Raising	84.62	2.808269	-0.31959615
Subject	**Tough**	**77.37**	**2.647870**	**-0.05522222**
Object	**Raising**	**82.50**	**2.782925**	**-0.18355000**
Object	Tough	43.90	3.004024	0.70290244

Baseline accuracy is high for both raising and *tough* constructions. Moving from baseline to subject gap conditions, expected response rate dips slightly for both constructions, with a slight but not significant decrease for *tough* constructions. There is also a slight, but not significant, increase in reaction times for both *tough* and raising subject movement constructions. Overall, both subject *tough* and subject raising constructions have expected response rates and reaction times that are comparable to their baselines.

Object gap constructions show a different pattern. Object raising constructions look like subject raising constructions: there is no significant difference, in terms of either expected response rate or reaction time. However object *tough* constructions do not pattern with the rest of the data. Object *tough* reaction times are significantly longer than reaction times for baseline *tough* constructions. This difference in reaction time is already enough to suggest an ungrammatical judgment. Additionally, the expected response rate for object *tough* constructions was roughly at-chance, suggesting that children were guessing. This was significantly different from the subject raising constructions. This low expected response rate also suggests an ungrammatical judgment.

4.4.1 Expected response data

The expected response data was fit to a generalized fixed-effects mixed logistic regression model using maximum likelihood, with participant and item (i.e., the story that was presented to the child) as random effects. Participant random effects corresponded to a random slope determined by predicate type. Item random effects corresponded to a random slope determined by predicate type and gap type (e.g., subject x *tough*). This was the full random effect structure supported by the data. Fixed effects were predicate type (*tough* predicate vs raising predicate; sum coded) and gap type. Gap type was Helmert coded. Expected answers were given a score of 1 and unexpected answers were given a score of 0. Helmert coding allowed the model to compare two things: baseline vs gap conditions generally (no movement vs movement) and, within gap conditions, subject gaps vs object gaps.

There was a marginally significant decrease in expected response rate when gap (movement) and baseline (no movement) conditions were compared (p = 0.059) across *tough* and raising constructions. There was no significant effect of gap site (p=0.5) across the raising and *tough* data. There was a significant decrease in expected response rate for *tough* constructions, compared to raising constructions, across all conditions (p=0.023). Post hoc analysis revealed that the marginally significant effect of movement was driven by this *tough* movement effect. Additionally, within the *tough* movement condition, there was a significant decease in expected response rate (p=0.005) for object *tough* constructions, compared to subject *tough* constructions. There was no parallel effect within the raising data. The analysis is summarized in Figure 2.

Figure 2: Accuracy: statistical analysis

Factor Label	Description	Co-efficient	Standard Error	Significance
Intercept		1.5503	0.2738	≈ 0
Gap1	object (-) vs subject gap	-0.202	0.3325	0.5428
Gap2	gaps (-) vs baseline conditions	-0.3167	0.1675	***0.0587*** | <.01
PredType	tough (-) vs raising	-0.4839	0.2127	***0.0229*** | <.05
Pred*Gap	tough movement vs baseline	-0.3741	0.2465	0.1291
	tough object vs subject	-0.3402	0.1211	***0.00498*** | <.005

4.4.2 Reaction Time

Reaction time data was fit by REML to a generalized linear mixed effects model with the same fixed and random effects as the participant data. Reaction time data was not modeled on the raw reaction times themselves, but on the delta reaction time scores: the difference between the baseline reaction time and the movement condition reaction time, by predicate, by participant. All children saw each *tough* and raising predicate twice, once in a baseline condition and once in A-movement condition. For each child, their delta reaction time per predicate was computed by subtracting the baseline condition reaction time from the movement condition reaction time. Delta reaction times were determined for each movement/baseline pair, regardless of the child's response to either condition. Delta reaction time was 0 for all baseline conditions. Delta reaction times were used instead of raw reaction times to reduce variance across participants (some children had slower reaction times across the board). This also made it easier to directly compare the four movement conditions. Additionally, the use of delta reaction times makes changes in reaction time between the baseline and movement conditions more obvious to the reader. Importantly, results did not differ when the analysis was run on the raw and delta reaction time data.

Reaction time was coded in milliseconds. Reaction time data was also trimmed by participant; for each participant, the reaction time cutoff point was the number represented by their mean reaction time +/- twice the standard deviation. While this number differed considerably between participants, the average cutoff point across all participants was approximately 7.25 seconds[19].

Reaction time data mirrored expected response data in two ways. First, there was a significant different for reaction time ($p=0.0042$) between *tough* and raising movement stimuli across the board. This is the same as the effect of *tough* movement that was present in the expected response data. This means that all *tough* constructions showed significantly slower reaction times, compared to raising constructions. Second, within *tough* constructions, object *tough* constructions showed significantly higher reaction times, compared to subject *tough* constructions ($p=0.01$). When baseline and gap conditions are compared, there was no general effect of movement on reaction time. Figure 3 summarizes the statistical analysis for reaction time.

[19]No data points were cut because the reaction time was too fast.

Figure 3: Reaction time: statistical analysis

Factor	Description	Co-efficient	Standard Error	Significance
Intercept		0.0161	0.1493	≈ 0
Gap1	object vs subject gap	0.2134	0.1772	0.1193
Gap2	gap vs baseline	0.0097	0.0798	0.9045
PredType	tough vs raising	0.3133	0.0999	***0.0041*** | <.05
Pred*Gap	tough gap vs baseline	0.0587	0.1292	0.6534
	tough object vs subject	0.1576	0.0569	***0.01009*** | =.01

Unlike the expected response data, the reaction time data show no marginal effect of movement. However, like the expected response data, there is still a significant effect of *tough* movement in the reaction time data. This supports the claim that the main effect of movement in the expected response data was actually driven by the *tough* movement data. Additionally, the reaction time data show the same strongly significant effect of object *tough* movement as the expected response data. Both the expected response and reaction time data show that children struggle most with object *tough* constructions.

4.5 Discussion

Overall, the data clearly show two things. First, children distinguish *tough* and raising constructions as separate syntactic constructions. This is clear because there is a significant difference of both expected response rate and reaction time between subject and object *tough* constructions and there is no difference of either expected response rate or reaction time between subject and object raising constructions. Additionally, this is clear because there is an effect of movement for *tough* constructions but no parallel effect of movement for raising. Second, the data show that children find both subject and object raising constructions grammatical. This is clearly the case when we compare the raising data to the predictions in Section 3. We know from previous studies and the current study that children found the baseline raising examples grammatical. The fact that children's responses—in terms of either reaction times or expected response rates—do not differ between raising constructions (with movement) and the raising baseline conditions indicates a grammaticality judgment.

Importantly, this result does not suggest that children have an adult-like

representation of raising. Children do not share the adult judgment that object raising is ungrammatical. In this way, children's representation of raising is fundamentally different from adults'.

The *tough* data is less straightforward. From the data, it is immediately clear that children find object *tough* constructions ungrammatical. This is clear because children display an at-chance expected response rate and delayed reaction time when they answer test questions containing an object *tough* construction. These are signatures of an ungrammaticality judgment. The subject *tough* movement data is more ambiguous. Overall, there is a main effect of *tough* movement for both expected response rate and reaction time. In terms of this effect, subject and object *tough* constructions pattern together. This could mean that children find both subject and object *tough* constructions ungrammatical. However, when object *tough* and object raising constructions are directly compared, there is still a significant effect of both expected response rate ($p=0.007$) and reaction time ($p=0.04$). This result is not found when subject *tough* and subject raising constructions are compared ($p>0.05$ for both expected response rate and reaction time), despite the fact that there is a decrease in expected response rate between subject *tough* and subject raising constructions. This result can be interpreted as showing that children actually find subject *tough* constructions grammatical. Under this interpretation, the main effect of *tough* movement could actually be driven by object *tough* construction data.

This section will argue that children, in fact, find both subject and object *tough* constructions ungrammatical. In particular, this section argues that the object *tough* data shows us what happens when children encounter an ungrammatical construction and are unable to utilize a successful repair strategy. After a processing delay, children produce a random guess as their answer, resulting in a delayed reaction time and an at-chance expected response rate. The subject *tough* data, however, show what happens when children encounter an ungrammatical construction which they are able to reanalyze. Importantly, this section will focus on why a successful repair strategy is only available for subject *tough* construction and why children struggle to reanalyze object *tough* constructions. Regardless of whether children find subject *tough* constructions grammatical or ungrammatical, however, this data does support the conclusion that children do not have an adult-like representation of *tough* movement.

This data is not consistent with either Wexler's or Anderson's predictions. The raising data alone, which shows that children find both subject and object raising

constructions grammatical, supports Anderson's predictions, however the *tough* data does not. The *tough* data alone, assuming an analysis where children find both subject and object *tough* constructions ungrammatical, is consistent with Wexler's predictions, but the raising data is not. This is an unexpected result, and this section will also discuss how to understand the results of this experiment in relation to the literature.

In discussing why children treat *tough* constructions differently from raising constructions—and particularly why children are able to represent both subject and object raising constructions while they struggle with *tough* constructions—this section will also suggest that the data can potentially be used to help distinguish between competing theoretical models of *tough* movement. The data do not support a null operator analysis of *tough* constructions, and so can potentially be interpreted to support an improper movement model of *tough* constructions. However, any attempt to tie experimental results to a theoretical model of syntax will remain tentative until experiments that are specifically designed to directly compare competing theoretical models are performed.

4.5.1 Raising results

This section discusses how to interpret the raising data from this study. The raising results are particularly striking. These results support the previous literature which suggests that children are able to raise subjects as early as age four. These results also help us understand why children show a non-adult like performance on raising constructions that involve an oblique experiencer. If children believe that both subjects and objects can undergo raising, they might misinterpret oblique experiencers as the subject of the embedded clause and then struggle to understand if the matrix subject should be co-referential with the embedded subject or object. This study also suggests that children do not have an adult-like representation of raising because they overgeneralize the number of arguments that can undergo raising, not because children have difficulty establishing filler gap relationships in raising constructions. The fact that children believe that objects can raise over subjects potentially suggests that children have not yet acquired relativized minimality effects.

Children can raise subjects and objects

The data from this experiment support the conclusion that children can raise arguments from both the embedded subject and object position. Some previous studies have suggested that children do not have an adult-like representation of raising, and this data is consistent with that general result—allowing for grammatical object raising is decidedly non-adult-like. The fact that children were able to answer subject and object raising questions as quickly and accurately as the raising baseline questions suggests that children do not face any additional difficulty when responding to subject and object raising constructions, compared to the baseline constructions.

It is important to ask what, exactly, children are doing when they respond to a subject or object raising construction. It may or may not make sense to describe what children are doing with these constructions as "raising"—after all, the fact that object raising is not possible in the adult grammar is a key part of the definition of and theory of raising. However, from the child response data, we know that children believe that subject and objects of the clauses embedded by certain modal predicates can move into the matrix clause. The semantics of these constructions align with the semantics of adult raising constructions.

We also know that children are able to establish a filler-gap relationship across a clause boundary, and that children are able to interpret a modal operator as taking scope over an embedded clause without assigning a Θ-role to the matrix predicate. We have access to this level of semantic detail because (a) we know that children can parse the baseline constructions and (b) the individual predicates chosen for this study did not have a lexical semantics that could assign a Θ-role to a type <e> argument. If children were, for example, mis-analyzing raising constructions as control constructions, we would expect to see ungrammatical response behaviors when they encounter something like "Julia is unlikely." This is consistent with Becker *et al*'s (2012) result, which showed that children are able to distinguish between movement and control constructions and that children have an adult-like understanding of the Θ-grids for control predicates.

Similarly, the results of this study are not consistent with children using the deletion hypothesis to reanalyze the raising constructions as monoclausal transitives (Hirsch, Orfitelli & Wexler 2008). Were this the case, we would predict lower expected response rates for subject raising constructions (where children would be able to arrive at a transitive interpretation but would not be able to incorporate the modal semantics of the raising predicate into the transitive

interpretation) and an at-chance expected response rate for object raising constructions (where children would not be able to arrive at a transitive interpretation). While it is not clear, exactly, how children represent subject and object raising constructions (with object raising posing the biggest mystery), it does appear the children are representing raising constructions as some kind of A-movement construction. The following discussions address each of these points in greater detail.

Children differentiate raising and control

An important result of Becker *et al* is that children understand that control predicates must assign a Θ-role to their matrix subject and that other adjectival predicates, like *tough* predicates, do not[20]. The results of this study suggest that children can also differentiate between the Θ-grids of raising and control predicates. This data suggests that children understand that raising predicates do not assign a Θ-role to their matrix subject.

This study used raising predicates that, on an adult interpretation at least, did not have a semantics that allowed them to assign a Θ-role to their matrix subject. Adults cannot grammatically parse a sentence like "Julia is likely" or "Jim is predicted." Children show an adult-like understanding of the baseline raising constructions, which involve an expletive subject and a gapless embedded clause. Because children can grammatical represent the Θ-grids of raising predicates in baseline construction, and because children find non-baseline raising

[20] *Tough* predicates are regularly assumed not to assign a Θ-role to their matrix subjects (e.g., Polinsky 2013, see also the introduction to this dissertation). However, other research has noted that *tough* constructions do have a Θ-role that they can discharge to an optionally licensed experiencer argument (Postal 1971, Pesetsky 1995). Particularly, in cases where *tough* constructions do not license an experiencer argument, the choice of *tough* predicate does appear to influence the interpretation of the matrix subject, as in (i). However, (ii) shows that this interpretation is clearly an implicature.

(i) That mountain is tough to climb (implies the mountain is somehow "tough," perhaps rugged or steep)

(ii) That mountain is tough to climb because, even though the slope is gently, it's normally very crowded.

However, the discussion of whether *tough* predicates are able to assign a Θ-role to their matrix subject is further complicated by idiom facts, as in (iii) and (iv).

(iii) Headway is easy to make on problems like this. (suggests no Θ-role)

(iv) *Heed is important to pay to such warnings. (suggests a Θ-role)

constructions grammatical, we can understand the data as suggesting that children understand the Θ-grids of raising predicates across both movement and baseline constructions. The fact that children show such a high expected response rate—85% and 86% on the subject and baseline raising constructions respectively—provide support of this claim, as it show that children are extremely adult-like for subject raising and raising baseline conditions.

The object raising data are slightly less straightforward because object raising constructions are adult-ungrammatical. However, because the expected responses to object raising constructions were determined by the semantics of their corresponding baseline constructions, the raising object data also suggest that children believe that the relationship between object raising constructions and their corresponding baselines was no different than the relationship between subject raising constructions and their corresponding baselines.

Additionally, the expected responses for the raising questions was randomized and counterbalanced. If children interpreted raising predicates as assigning a Θ-role to their matrix subjects, we would not predict that children would consistently provide the same responses to pairs like (33), which have the same expected response but different matrix subjects. Because factors like gap condition and the present/absence of negation in the test question had no effect on object raising responses, we can assume that children understand the Θ-grids of raising predicates.

(33) a. Jim is not likely to beat Julia.
b. *Julia is not predicted for Jim to beat.

Children do not delete raising predicates

Hirsch, Orfitelli & Wexler (2008) propose that children process raising constructions by simply deleting the raising predicate from the linear word order before responding to the transitive sentence that the raising predicate embeds. Hirsch, Orfitelli & Wexler (2008) claim that a child cannot grammatically parse a question like *"Does the dog seem to have blue fur?"* and so reanalyze the sentence into a simple transitive by deleting the raising predicate and any associated function words, resulting in the transitive *"Does the dog have blue fur?"* After this process, the child answers according to the truth conditions of the resulting sentence.

With subject raising, it is very difficult to determine whether children are

utilizing the deletion strategy. This is because raising sentences often have truth conditions that are very similar to those of the transitive clauses that they embed; a dog can *have* blue fur and *seem to have* blue fur simultaneously.

While the stories used in this study tried to control for this possibility, three of the four raising predicates selected had very similar modal semantics (*likely, expected, predicted*). These are raising predicates that, like *seem*, occur in raising constructions that can share truth conditions with the transitive sentences that they embed. The fourth raising predicate, *unlikely*, does not have this property. If Hirsch, Orfitelli, & Wexler's deletion hypothesis was correct, we would predict to see a difference in expected response rates between raising constructions with *unlikely* (81% expected response rate) and the other raising predicates (*likely*=85%, *expected*=81%, *predicted*=92%). This is not the case, which suggests that children are not utilizing the deletion hypothesis when answering subject raising questions. However, the dataset is still relatively small and so it is still somewhat difficult to draw conclusions from the subject raising data alone.

The object raising data, however, are straightforwardly incompatible with the deletion hypothesis. This is because object raising constructions cannot be transformed into grammatical transitive constructions via deletions of the raising predicate alone[21]. Deleting the raising predicate from the object raising sentence **Is Jim likely for Julia to beat?* yields the adult ungrammatical sentence *"*Jim Julia beat."* We do not expect children to be able to interpret an ungrammatical sentence like *"*Jim Julia beat."* Indeed, if children were re-analyzing object raising constructions this way, we would predict that object raising constructions would show long reaction times (reflecting the cost of reanalysis) and at-chance expected response rates. We predict at-chance expected response rates because sentences like *"*Jim Julia beat"* should be very difficult for children to interpret. This difficulty is expected because there is nothing to help children determine whether *Jim* (the first DP in the sentence) or *Julia* (the DP that directly precedes the verb) is the subject of the sentence. Because of this ambiguity, we predict either that children will guess which DP should be interpreted as the subject, and so will only give an expected response when they correctly guess that *Julia* should be interpreted as the subject or, that children will routinely answer "yes" to all questions (the yes-bias). Because expected response rates were consistent across

[21]Recall that, under the deletion hypothesis, children can only delete an embedded verb and surrounding functional material from the linear word order. Children are not proposed to be able to delete arguments or rearrange other words

object gap and baseline movement constructions, and because no individual child answered "yes" to to each object raising question, there is no evidence that children were using a deletion strategy to parse object raising constructions.

4.5.2 Tough results

This section discusses how to interpret that *tough* results. Specifically, this section proposes that the data support an analysis where children find both subject and object *tough* constructions ungrammatical. The *tough* construction data, then, paint a picture that is very different from that of the raising data. If children find both subject and object constructions ungrammatical, then children believe that neither subjects nor objects can undergo *tough* movement. This is unlike children's view of raising. One important question that emerges from this is why children represent *tough* and raising constructions so differently. Importantly, this result also supports the claim that children are able to distinguish *tough* and raising predicates.

The proposal that children find both subject and object *tough* constructions ungrammatical comes from the fact that children's reactions to subject and object *tough* constructions directly align with Hirsh, Orfitelli & Wexler's deletion hypothesis predictions and are also consistent with ungrammaticality judgments. This section will discussed how the data support a child ungrammaticality judgment for both subject and object *tough* constructions. Following this, the remainder of this section will discuss how to understand the results of this study within the previous literature and why this data can be used as evidence against a null operator theory of *tough* movement.

Children find both subject and object *tough* constructions ungrammatical

While the data clearly show that children find object *tough* constructions ungrammatical, the subject *tough* data is initially more difficult to interpret, particularly since the result of the statistical analysis could be consistent with either a model where children find subject *tough* constructions ungrammatical (i.e., the most straightforward interpretation of the main effect of *tough* movement) or grammatical (i.e., the most straightforward interpretation of the lack of any significant difference between subject *tough* and subject raising constructions).

Understanding the object *tough* data is important to understanding the subject

tough data. Recall that object *tough* constructions showed significantly longer reaction times and an at-chance expected response rate. This is exactly the behavior predicted for object raising constructions under Hirsh, Orfitelli & Wexler's deletion hypothesis. If children were using the deletion hypothesis as a repair strategy after encountering a child-ungrammatical object *tough* construction, it would be expected that children would reanalyze object *tough* constructions like *"*Will Jim be easy for Julia to beat?"* as the ungrammatical transitive *"Jim Julia beat."* Children are predicted to show an at-chance expected response rate for these ungrammatical transitive constructions, as they have no way to understand which DP they should interpret as the subject or object of the sentence. These sentences are also predicted to show a longer reaction time, to reflect the cost of reanalysis. The object *tough* data clearly aligns with these predictions.

The subject *tough* data is also consistent with the predictions of the deletion hypothesis. Recall that children produced the expected responses to *tough* subject questions roughly 75% of the time. Like with the raising predicates chosen for this study, three of the *tough* predicates included in this study (*simple, easy, possible*) had modal semantics that allowed for subject *tough* construction to have truth conditions that were fairly similar to the truth conditions of the embedded transitive clause. For example, the expected answer to *"Will it be easy for Jim to beat Julia?"* (the closest grammatical correlate of the adult ungrammatical subject *tough* construction *"Will Jim be easy to beat Julia,"* which arguably does not have any truth conditions to adult English speakers) is likely the same as the expected response to the question *"Will Jim beat Julia,"* assuming no change in context.

However, the fourth *tough* predicate used for this study (*tough*), does not have this property. As a result, the deletion hypothesis predicts that children's expected response rate will differ between questions that used the predicate *tough* and the other three predicates. This prediction is born out: expected response rates for *simple, easy* and *possible* were 75%, 72% and 78%, respectively. The expected response rate for questions involving *tough* was 61%, lower than the other three predicates and not far from chance[22].

This supports an analysis where children find both subject and object *tough*

[22]If this analysis is correct, it might stand to ask how children acquire the meaning of the predicate *tough* at all. One possibility is that children acquire the meaning of this predict through the expletive-subject use of *tough*, as well as perhaps straightforwardly adjectival uses of *tough* (e.g., *"This problem is tough."* I leave the analysis of how and when children acquire the semantics of *tough* to future research, however.

constructions ungrammatical. Children clearly find object *tough* constructions ungrammatical; children are not able to reanalyze object *tough* constructions as transitive sentences and so answer object *tough* constructions slowly and at-chance. The fact that children are able to reanalyze subject *tough* constructions as grammatical transitive clauses initially obscures the fact that children find subject *tough* constructions ungrammatical. However, closer investigation of the data shows that the main effect of *tough* movement is real: children find both types of *tough* construction ungrammatical.

Children are not re-analyzing *tough* as control or raising

The data clearly show that children have different representations of *tough* and raising constructions. This means that children should understand that *tough* and raising predicates have different properties. Because children simultaneously believe that subjects and objects can raise and that neither subjects nor objects can *tough* move, it does not makes sense to interpret these results as suggesting that children conflate *tough* and raising predicates, or cannot distinguish between raising and *tough* movement.

Similarly, the data also show that children are able to distinguish between *tough* and control constructions. If the results of Becker *et al* (2012) can be taken seriously, and we have evidence that children have acquired an adult-like representation of control before they have acquired *tough* constructions, we do not have reason to believe that children are conflating or confusing *tough* and control predicates. Additionally, if children believed that *tough* constructions had the same structure as control constructions, we would at least predict that children would judge subject *tough* constructions as grammatical, as the child grammar allows for subject control.

The data are least compatible with an OP theory of *tough* movement

As discussed in Section 4.2, the results of this experiment are not consistent with a null operator theory of *tough* movement. This is because a null operator theory of *tough* movement predicted that children would find *tough* constructions grammatical, as children have already acquired the tools to engage in both subject and object *tough* movement by the time they are participating in the study. Specifically, children have already acquired Ā-movement and the ability to bind null operators with their matrix antecedents. The fact that children cannot use

tough movement suggests that *tough* movement involves at least some additional syntactic operation.

An improper movement theory of *tough* movement makes much broader predictions. Particularly, because there has been no research into how children acquire a ban on improper movement, an improper movement analysis of *tough* movement could be consistent with either a grammaticality judgment or an ungrammaticality judgment. Because of this, while the results of this study favor an improper movement approach to *tough* movement over a null operator analysis, it is difficult to meaningfully say that the results of this study are predicted by an improper movement theory of *tough* constructions.

4.5.3 Connecting with the previous literature

Connecting with the raising literature

The results of this study show that children find both subject and object raising constructions grammatical and both *tough* constructions ungrammatical. This result is not predicted by either Wexler's or Anderson's hypothesis. Importantly, Anderson's and Wexler's hypotheses were formed based on limited acquisition data. While this study looked at child ungrammaticality judgments for four types of movement constructions, previous studies only looked at whether or not children gave adult-like responses to adult-grammatical *tough* and raising constructions. Child response data alone is useful for determining whether or not children have an adult-like representation of a given construction, but it not sufficient to access child grammaticality judgments.

The results of previous studies have clearly shown that children are very delayed in their acquisition of *tough* constructions (the results of raising acquisition studies are mixed), but further interpretation is difficult. For example, the *tough* data from previous studies is consistent with both Anderson's hypothesis (grammaticality and an ambiguous representation of symmetric *tough* constructions) and Wexler's hypothesis (ungrammatical and an ungrammatical representation of symmetric *tough* constructions). Without additional information, it is impossible to distinguish between these models.

This study is consistent with the previous acquisition literature in that it confirms that children have non-adult-like representations of *tough* and raising constructions. These results also help illuminate why the *tough* and raising acquisition literatures have traditionally been concerned with separate questions.

Specifically, the raising literature has traditionally debated *whether* children have an adult-like representation of raising while the *tough* literature has debated *why* children do not possess an adult-like representation of *tough* movement.

The results of this study help make sense of the raising literature. Recall that the results of this study show that children can raise subjects. This clarifies why previous studies that have focused on unambiguous subject raising constructions have shown that children can raise subjects. Studies that have shown that children can raise subjects have also argued that children have an adult-like representation of raising. Similarly, studies that use stimuli that could be ambiguous to a child—for example, stimuli that make use of an optionally transitive embedded clause verb, like *John is likely to win* or stimuli that make use of raising predicates that optionally license matrix level experiencers (which a child may interpret as an embedded subject) like *seem*—show non-adult-like performance. This results of this study help clarify why children do not show an adult-like representation of raising in these contexts: they believe that both subjects and objects can raise, and so interpret raising stimuli as ambiguous between a subject gap and object gap interpretation.

The results of this study also help make sense of the *tough* acquisition literature. Recall that this study has shown *why* children do not treat *tough* constructions like adults do: because they do not believe that subjects or objects can *tough* move. Importantly, these results also help provide a context for why so many studies present results that suggest that children can access subject gap interpretations for symmetric *tough* constructions. The deletion hypothesis predicts that children will reanalyze symmetric *tough* constructions like *"Is Cathy easy to see?"* as *"Cathy see(s)?"* If a study makes use of *tough* predicates like *easy* or *possible*, we additionally predict that in many cases the truth conditions of the resulting transitive (*"Cathy sees."*) will be consistent with the truth conditions of the subject *tough* interpretation of the construction (*"Is it easy for Cathy to see *(anything)?"*).

Connecting with the passive literature

It also stands to question whether or not these results can tell us anything about how children represent and understand verbal passive constructions. While there are a number of English passive constructions, this section will focus on the English short passive (no *by*-phrase) and long passive (containing a *by*-phrase), as illustrated in (34):

(34) a. Anneke criticized Jonathan Franzen. (active)
b. Jonathan Franzen was criticized. (short passive)
c. Jonathan Franzen was criticized by Anneke (long passive)

The acquisition of the passive is famously delayed (see Slobin 1966, Turner and Rommetveit 1967 for a evidence of a comprehension delay; Wells 1979 for evidence of a natural production delay; Hayhurst 1967, Maratsos and Abramovitch 1975 and Baldie 1976 for evidence of an elicited production delay). The vast majority of acquisition studies note that children are delayed in their mastery of both short and long passives (Gorden and Chafetz 1990, Hirsch and Wexler 2006, O'Brien et al. 2006, though c.f., Fox and Grodzinsky 1998 for an alternative proposal), and a variety of theories have been proposed to account for this delay. While a full account of the passive acquisition literature is beyond the scope of this discussion, see Orfitelli (2012) and the references therein for a more thorough overview of this literature.

Passives have also been argued to be syntactically complicated constructions (Chomsky 1981, Baker 1988, Jaeggli 1986, Collins 2005, among many others). Importantly, most analyses of the passive propose that passive constructions are derivationally related to their corresponding active constructions. Two of the many facts that passive analyses must account for are (a) that the canonical subject can be "demoted" to a(n optional) *by*-phrase and (b) that the object can A-move to the canonical subject position.

Different theories of the passive approach this problem differently. For example, a classic approach (e.g., Chomsky 1981, Jaeggli 1986) is to propose that the passive morphology itself can absorb the case-marking properties and agent Θ-role from the verb, contributing to the demotion of the subject. An alternative approach (Collins 2005), suggests that passive argument ordering can be derived via smuggling. Regardless of theory, a critical question in any analysis of the passive is determining a mechanism to allow the object to move across the thematic subject (position) to the canonical subject position without causing a relativized minimality violation.

While such a mechanism is vital to the adult grammar, the results of this experiment suggest that, perhaps, such a mechanism is not needed in the child grammar. Perhaps the results of this study might be understood to suggest that children show non-adult-like performance on passives because they don't need to demote or smuggle in order to A-move the object into the canonical subject position. If the raising data from this study can be extended to general

A-movement operations, children should be able to A-move passive objects across subjects without difficulty, unlike an adults. If this is the case, it might be additionally possible that an economy constraint could limit children's abilities to perform unnecessary syntactic operations. Essentially: if children can just A-move a passive object, they don't need to go to the trouble of smuggling. If children can A-move objects across subjects, they might have a relatively simple representation of the passive. It's additionally possible that, if children are able to access a simple passive structure, their grammars might prohibit them from access a more complicated passive representation[23].

Clearly, this discussion of how the raising data from this study might impact analyses of passive acquisition is, at this point, entirely hypothetical. It's entirely possible that the child grammar maintains passive movement and raising as very distinct movement operations. However, connecting the results of this study to the broader body of acquisition research is a worthwhile goal and should be pursued in future research.

4.5.4 Future research

This experiment has helped clarify how children judge the grammaticality of subject and object *tough* and raising constructions. It has additionally helped illuminate the results of the previous acquisition literature. However, there is still work to be done in order to fully understand the acquisition of *tough* and raising constructions.

While we know that children find subject and object raising constructions grammatical, future research needs to be done to understand how children represent the truth conditions of subject and object raising constructions. The result of this study suggest that children's interpretations of subject raising constructions align with adult interpretations, but this finding should be investigated directly. Additionally, since children do not have an adult-like representation of raising, we cannot know for certain how children represent and understand the syntax of raising constructions. For example, there is a mystery surrounding why children believe that objects can raise past subject, violating relativized minimality. Future work should directly investigate how children understand relativized minimality and the syntax of raising constructions. As previously noted, future research should also investigate what these results can

[23]Thanks to Danny Fox for bringing this possibility to my attention.

tell us about how children acquire passive constructions.

Additionally, it's worth noting that this study treated grammaticality judgments as binary. However, it is also possible treat grammaticality as a spectrum. Particularly, some English speakers report that, while both object raising and subject *tough* constructions are ungrammatical, object raising constructions are less sharply ungrammatical than subject *tough* constructions. To the best of my knowledge, it is not understood why grammaticality judgments can contrast like this, and there is no literature surrounding how children differentiate between constructions that are mildly and sharply ungrammatical.

Finally, this study suggests that the improper movement analysis of *tough* constructions better predicts the fact that children find all *tough* constructions ungrammatical, compared to the null operator analysis of *tough* constructions. This conclusion, however, comes largely because the results of this study are not consistent with the predictions of the null operator theory of *tough* movement—the improper movement theory of *tough* movement makes predictions that are too broad to be meaningful, because the acquisition of improper movement is not clearly understood. Future research should probe the question of whether language acquisition studies can differentiate between syntactic theories more directly. This study was not designed to directly compare competing syntactic theory of *tough* movement. For the claim that acquisition data favors an improper movement analysis of *tough* movement to be meaningful, follow up work must be done to compare these two syntactic analyses directly.

4.6 Conclusions

This study looked at how children judge the grammaticality of subject and object raising and *tough* constructions. To do this, this study made use of an experimental paradigm that used the combination of reaction time and expected response rate as a proxy for child grammaticality judgments. Importantly, this study did not limit itself to investigating adult-grammatical constructions.

The results of this study show that children find both subject and object raising constructions grammatical and both subject and object *tough* constructions ungrammatical. This result was not predicted by either of the leading hypotheses within the literature. One such hypothesis, from Wexler (2012), predicted that children would find all four construction ungrammatical. The other hypothesis, from Anderson (2005), predicted that children would find all four constructions

grammatical.

That being said, the results of this study can still help us understand the previous acquisition literature. With respect to raising, this study shows that—while children do not have an adult-like representation of raising—children do interpret subject raising constructions in a way that is consistent with how adults interpret (subject) raising constructions. This explains why studies that focus on constructions that are only compatible with a subject raising interpretation suggest that children have an adult-like representation of raising, while studies that utilize potentially ambiguous raising constructions do not.

The *tough* data clearly shows that children find object *tough* construction ungrammatical. This chapter has argued that children utilize Hirsh, Ortifelli & Wexler's deletion hypothesis in order to assign some interpretation to an ungrammatical *tough* construction. This hypothesis predicts that children will have an easier time interpreting subject gap constructions compared to object gap constructions, and this result is born out within the *tough* data. The fact that children are utilizing the deletion hypothesis to process *tough* constructions also clarifies why children might appear to have access to a subject gap interpretation of symmetric *tough* constructions. Perhaps most importantly, this study showed that it is possible to solicit child grammaticality judgments, and that we can use those judgments to better understand the child grammar. Especially, this study showed that results based on child grammaticality judgments can be much less ambiguous than results based on response results alone.

Finally, this study also raised new questions for future research. First, future work is needed to understand how children might represent the syntax of subject and object raising constructions. The relationship between raising and passive constructions in the child grammar is also an open—and potentially extremely interesting—question. The raising facts are particularly interesting given that children can distinguish between *tough*, raising and control predicates, as well as between A and Ā-movement.

Additionally, future research should look at *why* children cannot *tough* move. One hypothesis is that children cannot *tough* move because they have not yet learned that *tough* movement is an exception to a broader ban on improper movement, but this hypothesis is still very tentative. Particularly, future works need to be done to see if these results really can support adopting an improper movement theory or *tough* movement over a null operator theory of *tough* movement.

Chapter 5

Conclusions

5.1 Theoretical conclusions

This dissertation has proposed that a range of subject/non-subject asymmetries can be traced back to an independent Ā movement constraint in the grammar, spec-to-spec anti-locality. Spec-to-spec anti-locality restricts "short" Ā movement. Brillman & Hirsh's definition of anti-locality is given in (1). This definition bans Ā movement that does not cross either an intervening specifier operations, including subject Ā movement from spec-TP to an immediately adjacent spec-CP.

(1) **Spec-to-Spec Anti-Locality**
Ā-movement of a phrase from the specifier of XP must **cross** a specifier projected by a maximal projection other than XP. Movement from position α to β **crosses** γ if and only if γ dominates α but does not dominate β

While spec-to-spec anti-locality does not make an explicit reference to subjecthood, it underlies subject/non-subject asymmetries because subjects—by virtue of being the highest argument in the clause, and by virtue of independent constraints on successive-cyclic movement and phasehood—are the arguments that are most likely to enter in an Ā movement configuration banned by anti-locality (e.g., subject *wh*-movement from an embedded spec-TP to an embedded spec-CP). In this way, spec-to-spec anti-locality, when combined with other, independent systems in the grammar, can unify a disparate set of subject/non-subject asymmetries.

Chapters 2 and 3 proposes that spec-to-spec anti-locality can explain why subject *tough*-movement is prohibited (2a) while object *tough* movement is grammatical (2b), as well as the relationship between *tough* constructions and

gapped degree phrases. Recall that, unlike *tough* constructions, gapped degree phrases are compatible with both subject and object gaps (3).

(2) *Tough* movement

a. *Anneke is tough _ to think about Jonathan Franzen.
(subject *tough* ✗)

b. Jonathan Franzen is tough for Anneke to think about _.
(object *tough* ✓)

(3) Gapped degree phrases

a. Anneke is *too smart* _ to read about Jonathan Franzen.
(subject gap ✓)

b. Jonathan Franzen is *too inconsequential* for Anneke to read about _.
(object gap ✓)

This conclusion stems from an analysis where gapped degree phrases have a *larger* embedded clause structure, compared with *tough* constructions. While *tough* constructions involve an embedded CP clause, gapped degree phrases involve an embedded DegP clause—that is to say, the phase boundary of the embedded clause is a DegP layer that sits directly above the CP. The presence of this DegP layer allows the subjects of gapped degree phrases to undergo non-anti-local movement from spec-TP to spec-DegP (across the CP layer, Deg head, and the covert evaluator argument introduced in spec-DegP). *Tough* constructions cannot license subject movement because that movement chain would involve anti-local movement from the embedded spec-TP to the embedded spec-CP. This is the case under either a null operator or improper movement analysis of *tough* constructions.

Analyzing gapped degree phrases as having a *larger than usual* embedded clause structure also explains the aspects of gapped degree phrases that have seemed mysterious within the previous literature: (i) the fact that (object) gapped degree phrases share a number of syntactic behaviors with *tough* constructions and (ii) the fact that (object) gapped degree phrases appear to have a semantics very similar to that of parasitic gaps. Syntactically, analyzing gapped degree phases as having a DegP as their embedded clause allows for the proposal that (object) gapped degree phrases contain an instance of (something like) *tough* movement contained entirely within their embedded clauses. Semantically, this proposal allows the semantic operation COMPOSE—the same operation that Nissenbaum (2000) and Nissenbaum & Schwarz (2011) argue takes place in

parasitic gaps—to occur, binding the matrix subject to the embedded gap site. This analysis also provides an example of what a theory of syntax stands to gain from adopting Wurmbrand's (2017) proposal that phase boundaries are contextually determined. Wurmbrand's theory of contextually determined phasehood dovetails with a theory where there is no one-size-fits-all analysis for embedded clause sizes, and embedded clauses are allowed to range in size from (smaller than) TPs to (larger than) CPs.

Chapter 3 argued that spec-to-spec anti-locality can explain subject/non-subject asymmetries outside of the *tough* movement and gapped degree phrase facts. Specifically, this chapter discussed how spec-to-spec anti-locality could unify subject/non-subject asymmetries within English, including the famous *that*-trace effects (and perhaps the wider suite of English comp-trace effects, more generally), as well as a range of facts across unrelated languages. Specifically, Chapter 3 proposed that spec-to-spec anti-locality can explain the unavailability of subject *wh*-movement in Imbabura Quechua, resumptive pronoun asymmetries in Hebrew and Berber anti-agreement facts (when combined with a theory of anti-agreement as true agreement, following Baier 2016). This chapter also worked to explain the differences between the theories of spec-to-spec anti-locality proposed by Erlewine (2016) and Brillman & Hirsh (*to appear*), as well as how spec-to-spec anti-locality can interact with other facets of the grammar.

Importantly, Chapter 3 stressed the fact that spec-to-spec only affects Ā movement operations. While Chapter 3 did discuss what independent principles might underly spec-to-spec anti-locality, it is still not entirely clear why a spec-to-spec anti-locality constraint might exist within the grammar. Importantly, spec-to-spec anti-locality can only operate in a grammar that is sensitive to the distinction between A and Ā movement. To that end, it is possible that spec-to-spec anti-locality represents something more fundamental in the distinction between A and Ā movement. According to the analysis in Chapter 3, all Ā movement operations are sensitive to spec-to-spec anti-locality, much like all A-movement operations are sensitive to certain locality constraints. The most conservative interpretation of this fact allows sensitivity to spec-to-spec anti-locality to be added to the list of tests that can be deployed to distinguish A from Ā movement chains. A more aggressive interpretation of this fact might allow for the proposal that spec-to-spec anti-locality may be able to help us understand how A and Ā movement operations are fundamentally different from

each other—for example, perhaps Ā probes cannot enter in Agree relationships with local or adjacent heads, while A probes can only enter in relationships with local or adjacent heads. If such a line of analysis is on the right track, it's possible that spec-to-spec anti-locality and A-movement locality constraints (e.g., relativized minimality) could be motivated by the same principles. Under this theory of syntax, both the features that motivate A-movement operation and the structure of the resulting movement chain are be crucial in determining the grammaticality of a given movement chain. However, while these might be fruitful avenues for understanding why spec-to-spec anti-locality is active in the grammar, such analysis is left for future research.

5.2 Experimental conclusions

The experimental portion of this dissertation looked at whether adult subject/non-subject asymmetries are mirrored in the child grammar. To do this, Chapter 4 presented an experiment that looked at how children judge the grammaticality of subject and object raising and *tough* constructions. This study used an experimental paradigm that used the combination of reaction time and expected response rate as a proxy for child grammaticality judgments. Perhaps most importantly, this study showed that it is possible to solicit child grammaticality judgments, and that we can use those judgments to better understand the child grammar. This study also argued that results based on child grammaticality judgments can be less ambiguous than results based on response results alone. In this way, a crucial result of this chapter is the proposal that the current methodology—which has only been used a handful of time across language acquisition experiments—should be added to the language acquisition experimental toolbox, and is capable of producing clear results that sharpen our ability to compare the child and adult grammars.

More concretely, the results of this study showed that children do not conceive of subject/non-subject asymmetries the same way that adults do. While adults find subject raising and object *tough* constructions grammatical and subject *tough* and object raising constructions ungrammatical, children find both subject and object raising constructions grammatical and subject and object *tough* constructions ungrammatical. In the context of the previous acquisition literature, this is a surprising result. While there are two competing theories on how children might represent the grammaticality of *tough* and raising constructions,

neither predict this set of results. However, there are other aspects of the acquisition literature that the results of this experiment can help illuminate.

With respect to raising: This study shows that, while children do not have an adult-like representation of raising, children do interpret subject raising constructions in a way that is consistent with how adults interpret (subject) raising constructions. This explains why studies that focus on constructions that are only compatible with a subject raising interpretation suggest that children have an adult-like representation of raising, while studies that utilize potentially ambiguous raising constructions do not. Perhaps the most surprising result in this study is the discovery that children believe that object raising is grammatical, contrary to what theories of relativized minimality predict.

That children believe object raising is possible is potentially important for analyses of how children represent passive constructions. One possible extension of this study is the proposal that children do not have an adult-like representation of the passive because they do not need to engage in complicated syntactic operations (e.g., smuggling) in order to A-move the object past the subject. However, future research is needed to determine whether or not this hypothesis has merit.

With respect to *tough* constructions: The data clearly shows that children find object *tough* construction ungrammatical. While the subject *tough* data is initially more difficult to interpret, Chapter 4 has argued that children utilize Hirsh, Ortifelli & Wexler's deletion hypothesis in order to assign some interpretation to an ungrammatical subject and object *tough* construction. This hypothesis predicts that children will have an easier time interpreting subject gap constructions compared to object gap constructions. Importantly, this result is born out within the *tough* data. The fact that children are utilizing the deletion hypothesis to process *tough* constructions also clarifies why children might appear to have access to a subject gap interpretation of symmetric *tough* constructions.

The *tough* movement data also tentatively favors an improper movement theory of *tough* movement over a null operator theory of *tough* movement. This is because a null operator analysis of *tough* movement predicts that children will find (at least) object *tough* constructions grammatical[1], though the specific predictions of an improper movement theory of *tough* movement are more difficult to pin down, given the dearth of research available surrounding how

[1] Abstracting away from the question of whether or not children have not yet acquired the spec-to-spec anti-locality constraint that bans subject *tough* movement.

children represent improper movement chains more generally. However, future work is needed to confirm or refute this hypothesis.

This study also raised new question for future research. While the results of this study suggest that children and adults represent subject raising constructions in parallel ways, future studies to are needed to determine exactly how children represent the syntax of object raising constructions. This study also raises the question of how children distinguish between A and Ā movement, given that children believe that object raising is grammatical. Additionally, future research should look at *why* children cannot *tough* move. If these results really do favor an improper movement analysis of *tough* movement, one possibility might be that children cannot *tough* move because they have not yet learned that *tough* movement is an exception to a broader ban on improper movement. However, this hypothesis is still extremely tentative.

Appendix A

Tables

A.1 Experimental Design Items

A sample trial of the experiment, complete with 16 store/image/question triplets is given below. Filler questions have been removed from this script.

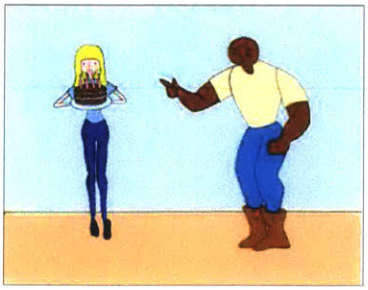
Sample Script: Subject Movement Trial

---

(1) **Raising block**

No Movement condition

Raising Baseline 1:
IMAGE:



STORY: Jim and Julia are playing a game where they try to boop each other on the nose, like this [demonstrates on Kermit]. The only problem is that Julia is holding

a birthday cake, and she can't let go, or the cake will fall!
QUESTION: Is it likely for him to boop her on the nose?

Raising Baseline 2:
IMAGE:

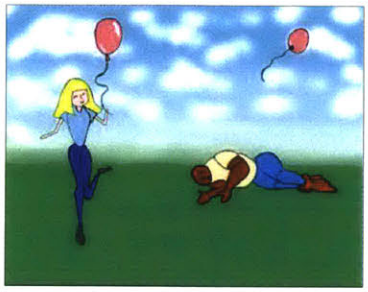
STORY: Jim and Julia are trying to find each other in a dark room. Jim has a cool flashlight. But Elizabeth's walking around with her eyes closed and can't see a thing!
QUESTION: Is it unlikely for him to find her?

Raising Baseline 3:
IMAGE:

STORY: Jim and Julia are trying to scare each other with their toys. Julia has on a really scare ghost costume! But poor Jim only has a flower. QUESTION: Is it expected for her to scare him?

Raising Baseline 4:

IMAGE:

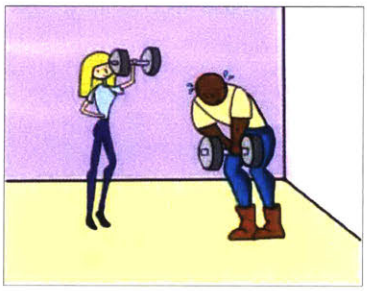


STORY: Jim and Julia are playing a game and trying to hit each other with a balloon. Jim stopped under a tree to take a quick nap, and accidentally let go of his balloon. Then Julia found him! QUESTION: Is it predicted for him to hit her with a balloon?

Subject movement condition

Raising Subject Movement 1:

IMAGE:

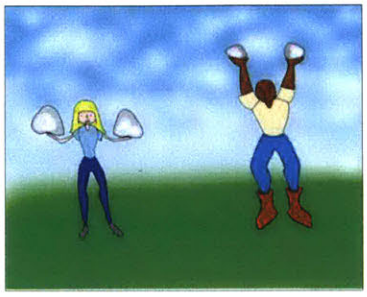


STORY: Jim and Julia are playing a game to see if they can pick the other up. Julia's really strong! She can lift anything. But Jim hurt his arm, and can't even pick up a pencil right now.

QUESTION: Is she likely to pick him up?

Raising Subject Movement 2:

IMAGE:

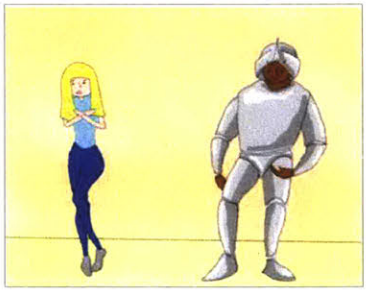


STORY: Jim and Julia are playing a game to see who is stronger. Both of them hold their arms out like this [demonstrates], and see and can see who can outlast the other ['Do you know what that word means?']. Julia's really sore from a game she played yesterday, and holding her arm out like this is really hard. She can barely do it!

QUESTION: Is he unlikely to outlast her?

Raising Subject Movement 3:

IMAGE:



STORY: Jim and Julia are trying to tickle each other! But Julia's a little nervous, because Jim is wearing a suit of armor!

QUESTION: Is she expected to tickle him?

Raising Subject Movement 4:
IMAGE:

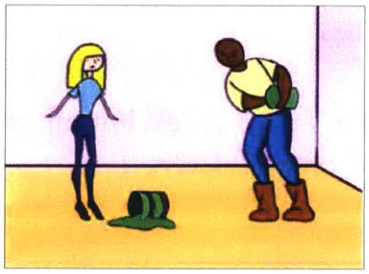
STORY: Jim is walking around with a blindfold! He doesn't know this, but Julia is also taking a nap in the field–and Jim's about to walk right where she's sleeping!
QUESTION: Is he predicted to bump into her?

---

(2) **Tough block**

No Movement condition

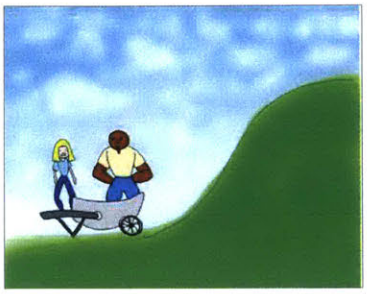
Tough Baseline 1:
IMAGE:



STORY: Jim and Julia are playing a game where they try to splash each other with paint. Jim has a whole can of paint to splash Julia with, but Julia accidentally spilled all her paint before the game ever started.
QUESTION: Will it be easy for him to splash her?

Tough Baseline 2:

IMAGE:

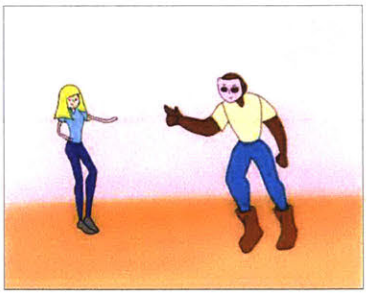


STORY: Jim and Julia and playing a game where they try to push each other up a hill in a wheelbarrow. Julia is really light, but Jim is very, very heavy. Julia doesn't think she can do it!

QUESTION: Will it be tough for him to push her up the hill?

Tough Baseline 3:

IMAGE:

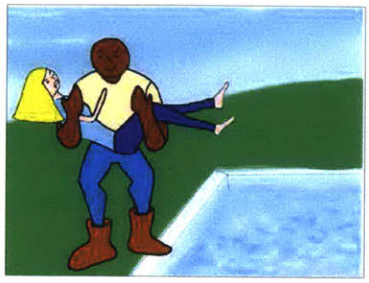


STORY: Jim and Julia are both trying to pinch the other on the cheek, but Jim is wearing a big Halloween mask that completely covers his face!

QUESTION: Will it be simple for her to pinch him?

Tough Baseline 4:

IMAGE:



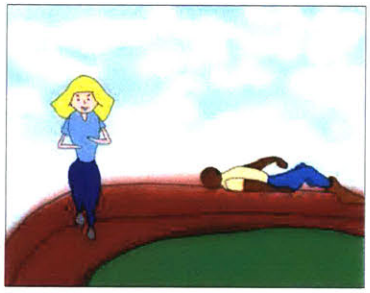
STORY: Jim and Julia are trying to throw each other in a pool! Jim's already picked Julia up, and is about to throw her in!

QUESTION: Will it be possible for him to throw her in the pool?

Subject Movement condition

Subject *Tough* 1:

IMAGE:

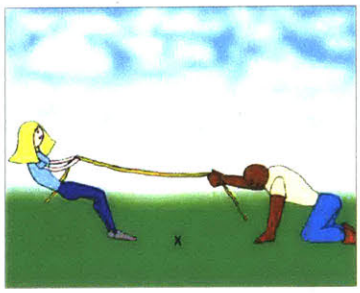


STORY: Jim and Julia are going to race each other. Julia has always been faster than Jim.Yesterday, Jim was practicing for the race, and he hurt his ankle. Today, he's running even slower than usual!

QUESTION: Will she be easy to beat him?

Subject *Tough* 2:

IMAGE:



STORY: Jim and Julia are having a tug-of-war. One of them will win when they pull on the rope so hard that the other crosses this spot (point to the line on the image). But Jim hurt his arm, and can't play very well! Plus, Julia is really strong!
QUESTION: Will she be tough to pull him over the line?

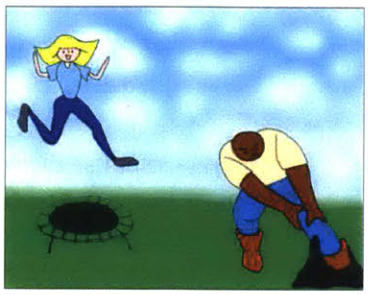
Subject *Tough* 3:

IMAGE:

STORY: Jim and Julia and playing a game where they try to touch each other with a stick. Jim's stick is very, very short, but Julia's stick is very, very long.
QUESTION: Will he be simple to touch her with a stick?

Subject *Tough* 4:

IMAGE:



STORY: Jim and Julia are playing leap frog, and trying to jump over each other. Julia is a great jumper, and even has a trampoline to help her out. Jim, however, has his feet trapped in some tar!

QUESTION: Will she be possible to jump over him?

A.2 Questions and predicate groups

A..1 *Tough*/raising pairs

The *tough*/raising pairs used in these stories are given in (3).

(3) *Tough*/raising predicate pairs

	Tough-predicate	Raising predicate
Pair 1	easy	likely
Pair 2	tough	unlikely
Pair 3	simple	expected
Pair 4	possible	predicted

A..2 Sample questions matrix

An example question matrix is given in (4), repeated from (27).

(4) **Question Matrix (beat)**

a. *Raising baseline*

i. Is it likely for Julia to beat Jim?	*Yes-response expected*
ii. Is it likely for Jim to beat Julia?	*No-response expected*

b. *Raising subject movement*

i. Is Julia likely to beat Jim?	*Yes-response expected*
ii. Is Jim likely to beat Julia?	*No-response expected*

c. *Raising object movement*

i. *Is Jim likely for Julia to beat?	*Yes-response expected*
ii. *Is Julia likely for Jim to beat?	*No-response expected*

d. *Tough baseline*

i. Will it be easy for Julia to beat Jim?	*Yes-response expected*
ii. Will it be easy for Jim to beat Julia?	*No-response expected*

e. *Tough subject movement*

i. *Will Julia be easy to beat Jim?	*Yes-response expected*
ii. *Will Jim be easy to beat Julia?	*No-response expected*

f. *Tough object movement*

i. Will Jim be easy for Julia to beat?	*Yes-response expected*
ii. Will Julia be easy for Jim to beat?	*No-response expected*

An addition question matrix is given in (5).

(5) **Question Matrix (beat)**

a. *Raising baseline*

i. Is it unlikely for Jim to pull Julia over the line? *Yes-response expected*

ii. Is it unlikely for Julia to pull Jim over the line? *No-response expected*

b. *Raising subject movement*

i. Is Jim unlikely to pull Julia over the line? *Yes-response expected*

ii. Is Julia unlikely to pull Jim over the line? *No-response expected*

c. *Raising object movement*

i. *Is Julia unlikely for Jim to pull over the line? *Yes-response expected*

ii. *Is Jim unlikely for Julia to pull over the line? *No-response expected*

d. *Tough baseline*

i. Will it be tough for Jim to pull Julia over the line?

Yes-response expected

ii. Will it be tough for Julia to pull Jim over the line?

No-response expected

e. *Tough subject movement*

i. *Will Jim be tough to pull Julia over the line? *Yes-response expected*

ii. *Will Julia be tough to pull Jim over the line? *No-response expected*

f. *Tough object movement*

i. Will Julia be tough for Jim to pull over the line?

Yes-response expected

ii. Will Jim be tough for Julia to pull over the line?

No-response expected

Baseline questions for the remaining 14 stories are given below.

(6) **Baseline questions for remaining items**

Item	baseline *raising* (yes)	baseline *tough* (yes)
splash	Is it likely for Jim to splash Julia?	Will it be easy for Jim to splash Julia?
pick up	Is it likely for Julia to pick Jim up?	Will it be easy for Julia to pick Jim up?
boop	Is it likely for Jim to boop Julia on the nose?	Will it be easy for Jim to boop Julia on the nose?
find	Is it unlikely for Julia to find Jim?	Will it be tough for Julia to find Jim?
outlast	Is it unlikely for Julia to outlast Jim?	Will it be tough for Julia to outlast Jim?
push	Is it unlikely for Julia to push Jim up the hill?	Will it be tough for Julia to push Jim up the hill?
touch	Is it expected for Julia to touch Jim with her stick?	Will it be simple for Julia to touch Jim with her stick?
pinch	Is it expected for Jim to pinch Julia?	Will it be simple for Jim to pinch Julia?
tickle	Is it expected for Julia to tickle Jim?	Will it be simple for Julia to tickle Jim?
scare	Is it expected for Jim to scare Julia?	Will it be simple for Jim to scare Julia?
hit	It is predicted for Julia to hit Jim with her balloon?	Will it be possible for Julia to hit Jim with her balloon?
understand	It is predicted for Jim to understand Julia?	Will it be possible for Jim to understand Julia?
bump	Is it predicted for Julia to bump into Jim?	Will it be possible for Julia to bump into Jim?
jump	Is it predicted for Julia to jump over Jim?	Will it be possible for Julia to jump over Jim?

Bibliography

Abels, K. (2003). *Successive cyclicity, anti-locality, and adposition stranding*. PhD thesis, University of Connecticut.

Aissen, J. (1992). Topic and focus in Mayan. *Language*, pages 43–80.

Akmajian, A. (1972). Getting *tough*. *Linguistic Inquiry*, 3:373–377.

Anderson, D. (2005). *The acquisition of* Tough*-movement in English*. PhD thesis, University of Cambridge.

Armbuster, T. (1981). How to influence responses to "Is Kathy easy to see?". *Journal of Psycholinguistic Research*, 10(1):27–39.

Baier, N. (2016). Unifying anti-agreement and *wh*-agreement. Manuscript., University of California, Berkeley.

Baier, N. (2017). Antilocality and antiagreement. *Linguistic Inquiry*, 48:367–377.

Baker, M. (1988). *Incorporation*. Chicago University Press.

Baldie, B. (1976). The acquisition of the passive voice. *Journal of Child Language*, 3(3):331–348.

Becker, M. (2006). There began to be a learnability puzzle. *Linguistic Inquiry*, 37:441–456.

Becker, M., Estigarribia, B., and Gylfadottir, D. (2012). *Tough*-adjectives are easy to learn. In *Supplemental Proceedsings of BUCLD 36*.

Berman, A. (1973). A constraint on *tough*-movement. In *Ninth Regional Meeting of the Chicago Linguistics Society*.

Berman, R. (1997). Early acquisition of syntax and discourse in Hebrew. In Shimron, Y., editor, *Psycholinguistic Studies in Israel: Language acquisition, Reading and Writing*, pages 57–100. Magnes Press.

Blom, E. and Unsworth, S., editors (2010). *Experimental Methods in Language Acquisition Research*, volume 27. John Benjamins.

Boeckx, C. (2008). Islands. *Language and Linguistic Compass*, 2(1):151–167.

Borer, H. (1984). Restrictive relatives in modern Hebrew. *Natural Language and Linguistic Theory*, 2(2):219–260.

Bresnan, J. (1977). *Transformations and categories in syntax*. Springer Netherlands.

Brillman, R. (2015). Improper movement in *tough* constructions and gapped degree phrases. In *Proceedings from PLC 38*, volume 21. University of Pennsylvania Working Papers in Linguistics.

Brillman, R. and Hirsh, A. (to appear.). An anti-locality account of English subject/non-subject asymmetries. In *Proceedings of the 50th Chicago Linguistics Society*.

Cable, S. (2010). *The grammar of Q: Q-particles, wh-movement and pied-piping*. Oxford.

Chomsky, N. (1977). On *wh*-movement. In Culicover, P., Wasow, T., and Akmajian, A., editors, *Formal Syntax*, pages 71–132. Academic Press, New York.

Chomsky, N. (1981). *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, N. (1993). A minimalist program for linguistic theory. In Hale, K. and Keyser, S. J., editors, *The View from Building 20*, pages 1–52. MIT Press.

Chomsky, N. (1995). *The Minimalist Program*. MIT Press.

Chomsky, N. (2000). Minimalist inquiries: the framework. In Martin, R., Michaels, D., and Uriagereka, J., editors, *Step by Step: essays on Minimalist Syntax in honor of Howard Lasnik*. MIT Press, Cambridge, MA.

Chung, S. (1994). *Wh*-agreement and "referentiality" in Chamorro. *Linguistic Inquiry*, 25:1–44.

Chung, S. (1998). *The Design of Agreement: Evidence from Chamorro*. University of Chicago Press.

Chung, S. and McCloskey, J. (1983). On the interpretation of certain island facts in GPSG. *Linguistic Inquiry*, 14(4):704–713.

Cinque, G. (1999). *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.

Cole, P. (1985). *Imbabura Quechua*. Israeli Hebrew.

Collins, C. (2005). A smuggling approach to the passive in English. *Syntax*, 8(2):81–120.

Coon, J. (2013). TAM split ergativity, part I. *Language and Linguistic Compass*, 7(3):171–190.

Coon, J. and Henderson, R. (2017). Adverbs and variability in Kaqchikel agent focus: A reply to Erlewine (2016). *Natural Language and Linguistic Theory*, 35(3):1–25.

Crain, S., McKee, C., and Emiliani, M. (1990). Visiting relatives in Italy. In Frazier, L. and de Villiers, J. G., editors, *Language processing and Language acquisition*. Springer.

Cromer, R. (1970). Children are nice to understand: Surface structure clues for the recovery of a deep structure. *British Journal of Psychology*, 61(3):397–408.

Culicover, P. (1993). Evidence against ECP accounts of the *that-t* effect. *Linguistic Inquiry*, 24.

de Villiers, J., de Villiers, P., and Hoban, E. (1994). The central problem of functional categories in the English sytnax of oral deaf children. *Constraints on Language Acquisition: Studies of atypical children*, pages 9–47.

Deal, A. R. (2015). Interaction and satisfaction in ϕ-agreement. Manuscript., University of California, Berkeley.

Deal, A. R. (2016). Interaction and satisfaction: a theory of agreement. Handout. Massachusetts Institute of Technology.

Diercks, M. (2010). *Agreement with subjects in Lubukusu*. PhD thesis, Georetown University.

Doherty, C. (1996). Clausal structure and the Modern Irish copula. *Natural Language and Linguistic Theory*, 14(1):1–46.

Doron, E. (1982). On the syntax and semantics of resumptive pronouns. *Texas Linguistics Forum*, pages 1–48.

Elouazizi, N. (2005). Anti-agreement effects as (anti-)connectivity. In *Proceedings of the 24th West Coast Conference on Formal Linguistics*, pages 120–128.

Engdahl, E. (1983). Parasitic gaps. *Linguistics and Philosophy*, 6:5–34.

Erlewine, M. (2016). Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language and Linguistic Theory*, 34:429–479.

Erlewine, M. (2017). Why the null complementizer is special in the English *that*-trace effect. In Halpert, C., van Urk, C., and Kotek, H., editors, *A Pesky Set: Papers for David Pesetsky*. MIT Working Papers in Linguistics, Cambridge, MA.

Fleisher, N. (2013). On the absense of scopre reconstruction in *tough*-subject a chains. *Linguistic Inquiry*, 44(321-332).

Fleisher, N. (2015). *Rare*-class adjectives in the *tough*-construction. *Language*, 91(73-108).

Fox, D. (1999). Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry*, 30(2):157–196.

Fox, D. and Grodzinsky, Y. (1998). Children's passive: a view from the *by*-phrase. *Linguistic Inquiry*, 29(2):311–332.

Fox, D. and Pesetsky, D. (2006). Cyclic linearization of shared material. talk given Berlin.

Friedmann, N. and Lavi, H. (2006). On the order of acquisition of A-movement, *Wh* movement and V-C movement. In Belletti, A., Bennati, E., Chesi, C., Di Domenico, E., and Ferrari, I., editors, *Language Acquisition and Development*, pages 211–217. Cambridge Scholar Press, Newcastle, UK.

Friedmann, N. and Novogrodsky, R. (2004). The acquisition of relative clause comprehension in Hebrew. *Journal of Child Language*, 31:661–681.

Georgopoulous, C. (1991). *Syntactic variables: Resumptive pronouns and Ā binding in Palauan*. Springer.

Ginzburg, J. and Sag, I. (2001). *Interrogative Investigations: the form, meaning, and use of English Interrogatives*. CSLI Publications.

Gorden, P. and Chafetz, J. (1990). Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. *Cognition*, 36(227-254).

Grohmann, K. K. (2000). *Prolific Peripheries: A Radical View From the Left*. PhD thesis, University of Maryland, College Park.

Grohmann, K. K., Drury, J. E., and Castillo, J. C. (2000). No more EPP. In Billerey, R. and Lellehaugen, B. D., editors, *Proceedings of the 19th West Coast Conference on Formal Linguistics*, pages 153–166, Somerville, MA. Cascadilla.

Hartman, J. (2012). (Non-)Intervention in A-movement: some cross-constructional and cross-linguistic consequences. *Linguistic Variation*, 11(121-148).

Harves, S. and Kayne, R. (2012). Having 'need' and needing 'have'. *Linguistic Inquiry*, 43(1):120–132.

Hayhurst, H. (1967). Some errors of young children in producing passive sentences. *Journal of Learning and Verbal Behavior*, 6(4):634–639.

Heim, I. and Kratzer, A. (1998). *Semantics in generative grammar*. Blackwell, Malden, MA.

Hicks, G. (2003). *So easy to look at, so hard to define:* tough *movement in the Minimalist framework*. PhD thesis, University of York.

Hicks, G. (2009). *Tough*-constructions and their derivations. *Linguistic Inquiry*, 40:535–566.

Hirsch, C., Orfitelli, R., and Wexler, K. (2008). When *seem* means *think*: the role of the experiencer-phase in children's comprehension of raising. In Belikova, A., Meroni, L., and Umeda, M., editors, *Proceedings of the 2nd Conference on Generative Approaches to Langauge Acquisition North America*, pages 135–146. Cascadilla.

Hirsch, C. and Wexler, K. (2006). By the way, children don't know *by*. In Bamman, D., Magnitskaia, T., and Zaller, C., editors, *BUCLD: Proceedings of the 30th Annual Boston University Conference on Language Development*, pages 249–261, Somerville, MA. Cascadilla.

Holmberg, A. (2010). Parameters in minimalist theory: The case of Scandinavian. *Theoretical Linguistics*, 36(1-48).

Hornstein, N. (2001). *Move! A minimalist theory of construal*. Blackwell, Oxford.

Jaeggli, O. A. (1986). Passive. *Linguistic Inquiry*, 17(4):587–622.

Kandybowicz, J. (2006). Comp-trace effects explained away. In Baumer, D., Montero, D., and Scanlon, M., editors, *Proceedings of the 25th West Coast Conference on Formal Linguistics*, pages 220–228, Somerville, MA. Cascadilla Proceedings Project.

Kayne, R. (2009). Need and light verbs. Handout. Universitá di Siena.

Keine, S. and Poole, E. (2016). Intervention in *tough* constructions revisted. *The Linguistic Review*.

Kinjo, K. (2017). An intervention approach to antiagreement. In Kaplan, A., Kaplan, A., McCarvel, M. K., and Rubin, E. J., editors, *Proceedings of the 34th West Coast Conference on Formal Linguistics*, pages 319–338, Somerville, MA. Cascadilla.

Klein, E. (1993). *Toward second language acquisition*. PhD thesis, Queens College and the Graduate School and University Center, City University of New York.

Koopman, H. (1983). ECP effects in main clauses. *Linguistic Inquiry*, 14(2):346–350.

Koopman, H. and Sportiche, D. (2007). The *que/qui* alternation: New analytical directions. Manuscipt.

Koring, L. (2007). Raising verbs in Dutch: structure and acquisition. Master's thesis, Utrecht University.

Landau, I. (2001). Control and extraposition: the case of super-equi. *Natural Language and Linguistic Theory*, 19(1):109–152.

Landau, I. (2003). Movement out of Control. *Linguistic Inquiry*, 34(3):471–498.

Landau, I. (2006). Severing the distribution of PRO from case. *Syntax*, 9(2):153–170.

Lasnik, H. and Fiengo, R. (1974). Complement object deletion. *Linguistic Inquiry*, 5(4):535–571.

Lasnik, H. and Stowell, T. (1991). Weakest crossover. *Linguistic Inquiry*, 22:687–720.

Longenbaugh, N. (2016). Non-intervention in *tough* constructions. In Hammerly, C. and Prickett, B., editors, *Proceedings of the 46th Annual Meeting of the North East Linguistic Society*, volume 2, pages 293–307.

Longobardi, G. (1985). Connectedness, scope, and C-command. *Linguistic Inquiry*, 16(2):163–192.

Maratsos, M. and Abramovitch, R. (1975). How children understand full, truncated and anomalous passives. *Journal of Learning and Verbal Behavior*, 14(2):145–157.

Martin, R. (2001). Null case and the distribution of PRO. *Linguistic Inquiry*, 32:141–166.

May, R. (1977). *The Grammar of Quantification*. PhD thesis, Massachusetts Institute of Technology.

McCloskey, J. (1990). Resumptive pronouns, Ā-binding, and levels of representation in Irish. In Hendrick, R., editor, *The syntax and semantics of the modern Celtic languages*, pages 199–248. Academic Press, San Diego, CA.

McCloskey, J. (2002). Resumption, successive cyclicity, and the locality of operations. In Epstein, S. D. and Seely, T. D., editors, *Derivation and Explanation in the Minimalist Program*, pages 184–226. Blackwell Publishing.

McGinnis, M. (2003). Variation in the phase structure of applicatives. In Rooryck, J. and Pica, P., editors, *Linguistic Variation Yearbook*. John Benjamins.

McKee, C., McDaniel, D., and Snedeker, J. (1998). Relative children say. *Journal of Psycholinguistic Research*, 27:573–596.

Meier, C. (2003). The meaning of *too, enough,* and *so...that*. *Natural Language Semantics*, 11(1):69–107.

Merchant, J. (2003). Remarks on stripping. Manuscript, University of Chicago, Chicago, IL.

Morgan, J. (1973). Sentence fragments and the notion 'sentence'. In *Issues in linguistics: Papers in honor of Henry and Renée Kahane*, pages 719–751. University of Illinois Press.

Naigles, L., Fowler, A., and Helm, A. (1995). Syntactic bootstrapping from start to finish with special reference to down syndrome. In Tomasello, M. and Merriman, W. E., editors, *Beyond Names for Things: Young Children's Acquisition of Verbs*, chapter 11, pages 299–330. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ.

Nissenbaum, J. (2000). *Investigations of covert phrase movement*. PhD thesis, MIT.

Nissenbaum, J. and Schwarz, B. (2011). Parasitic degree phrases. *Natural Language Semantics*, 19:1–38.

O'Brien, K., Grolla, E., and Lillo-Martin, D. (2006). Long passives are understood by children. In Bamman, D., Magnitskaia, T., and Zaller, C., editors, *Proceedings of the 30th Annual Boston University Conference on Language Development*, Somerville, MA. Cascadilla.

Oftitelli, R. (2012). *Argument Intervention in the Acquisition of A-movement*. PhD thesis, UCLA.

O'Herin, B. (2002). *Case and Agreement in Abaza*. SIL International.

Ordonez, F. (1995). The anti-passive in Jacaltec: a last resort strategy. *Catalan Working Papers in Linguistics*, 4(2):329–343.

Ouali, H. (2011). *Agreement, Pronominal Clitics, and Negation in Tamazight Berber*. Bloomsbury Academic.

Ouhalla, J. (1993). Subject-extraction, negation and the anti-agreement effect. *Natural Language and Linguistic Theory*, 11(477-518).

Pesetsky, D. (1995). *Zero Syntax: Experiencers and Cascades*. MIT Press, Cambridge, MA.

Pesetsky, D. (1998). Some optimality principles of sentence pronunciation. In Barbosa, P., Fox, D., Hagstrom, P., McGinnis, M., and Pesetsky, D., editors, *Is the best good enough?*, pages 337–384. MIT Press, Cambridge, MA.

Pesetsky, D. (2017). Complementizer trace effects. In Everaert, M. and van Riemsdijk, H., editors, *A Companion to Syntax, 2nd edition*. Blackwell, Oxford.

Pesetsky, D. and Torrego, E. (2001). T-to-C movement: Causes and consequences. In Kenstowicz, M., editor, *Ken Hale: A life in language*, pages 355–426. MIT Press.

Polinsky, M. (2013). Raising and control. In den Dikken, M., editor, *The Cambridge Handbook of Generative Syntax*, pages 41–63. Cambridge University Press.

Postal, P. (1971). *Cross-over phenomena*. Rinehart and Winston, New York.

Rackowski, A. and Richards, N. (2005). Phase edge extraction: A Tagalog case study. *Linguistic Inquiry*, 36(4):565–599.

Radford, A. (2004). *Minimalist syntax: Exploring the structure of English*. Cambridge University Press.

Rasin, E. (2017). Two types of resumptive prounous: a minimal account of Hebrew interpretative asymmetries. Manuscript, MIT.

Richards, N. (1997). *What moves where when in which language?* PhD thesis, Massachusetts Institute of Technology.

Richards, N. (2016). *Contiguity Theory*. MIT Press, Cambridge, MA.

Rizzi, L. (1990). *Relativized Minimality*. MIT Press.

Rizzi, L. and Shlonsky, U. (2007a). Strategies of subject extraction. In *Interfacting + recursion = language? Chomsky's minimalism and the view from syntax-semantics*. Walter de Gruyter.

Rizzi, L. and Shlonsky, U. (2007b). Strategies of subject extraction. In *Interfacting + recursion = language? Chomsky's minimalism and the view from syntax-semantics*. Walter de Gruyter.

Rosenbaum, P. S. (1967). *The grammar of English predicte complement constructions*. MIT Press, Cambridge, MA.

Ross, J. (1967). *Constraints on variables in syntax*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Schütze, C. T. (2001). On the nature of default case. *Syntax*, 4(3):205–238.

Sells, P. (1984). *Syntax and semantics of resumptive pronouns*. PhD thesis, University of Massachusetts, Amherst.

Shlonsky, U. (1992). Resumptive pronouns as a last resort. *Linguistic Inquiry*, 23:443–468.

Shlonsky, U. (2004). Enclisis and proclisis. In Rizzi, L., editor, *The Structure of CP and IP*, pages 329–353. Oxford University Press, Oxford.

Sigurdhsson, H. A. (2008). The case of PRO. *Natural Language and Linguistic Theory*, 26(2):403–450.

Slobin, D. (1966). Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of Verbal Learning and Verbal Behaviors*, 5:219–227.

Stromswold, K. (1996). Does the VP-internal subject stage really exist? In *21st Annual Boston University Conference on Language Development.*

Turner, E. and Rommetveit, R. (1967). Experimental manipulation of the production of active and passive voice in children. *Language and Speech*, 10:169–180.

van Urk, C. and Richards, N. (2013). Two components of long-distance extraction: Successive cyclicity in Dinka. available online: lingbuzz/001766 (to appear in *Linguistic Inquiry*).

Željko Bošković (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.

Watanabe, A. (1996). *Case Absorption and* wh*-Agreement*. Kluwer Academic Publishers.

Wells, G. (1979). Describing children's linguistic development at home and at school. *British Educational Research Journal*, 5(1):75–98.

Wexler, K. (1992). Some issues in the growth of control. In *Control and Grammar*, pages 253–295. Springer Netherlands.

Wexler, K. (2004). *Theory of phasal development*, volume 48 of *MIT Working Papers in Linguistics*. MIT Working Papers in Linguistics, Cambridge, MA.

Wexler, K. (2012). *Tough* movement developmental delay: Another effect of phasal computation. In Piattelli-Palmarini, M. and Berwick, R. C., editors, *Rich Languages From Poor Inputs*, chapter 10. Oxford Scholarship.

Williams, E. (1983). Semantic vs syntactic categories. *Linguistics and Philosophy*, 6(3):423–446.

Wurmbrand, S. (2017). Stripping and topless complements. *Linguistic Inquiry*.

Wurmbrand, S. and Haddad, Y. A. (2016). Cyclic spell-out derived agreement in Arabic raising constructions. In Haddad, Y. A. and Potsdam, E., editors, *Perspectives on Arabic Linguistics XXVIII*, pages 193–226. John Benjamins, Amsterdam.