REMOTE AGREE: Agreeing into the Workspace and a Case Study in Control

by

Alexandra Motut

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Department of Linguistics
University of Toronto

REMOTE AGREE: Agreeing into the Workspace and a Case Study in Control

Alexandra Motut, Doctor of Philosophy, 2022

Department of Linguistics, University of Toronto

Abstract

I argue for a revision to the available generative operations of the grammar motivated by an expansion of the properties of I-Merge to its counterpart, E-Merge. Following Motut & Oda (2011), I argue that Agree into the syntactic Workspace is possible without subsequent E-Merge. I call this Remote Agree. I expand the concept of the Workspace to include a Workspace element, π, which can be Agreed with by a probe without (E-)Merging into the tree, and articulate rules for how Remote Agree interacts with Spellout in the Workspace to avoid overgeneration.

Using Remote Agree, I argue for a semi-syntactic analysis of arbitrary control and long-distance control that unifies these two non-obligatory control (NOC) phenomena and is consistent with a movement account of obligatory control (OC). The Remote Agree approach to NOC allows for the formulation of a new generalization of the NOC/OC distribution based on Spellout objects and the locality of operations in the Workspace. Arbitrary control is analyzed as Remote Agree with π as a last resort at Spellout of a Spellout object, after unsatisfied features on v and T have percolated to C. Long-distance control is established by Remote Agree with π followed by a coindexation operation between π and a DP in another Spellout object. This account explains the syntactic distribution of NOC (versus OC), its nullness and subject-oriented nature, and its [+HUMAN] requirement. Finally, I suggest possible new empirical applications for Remote Agree.

And a direct note to any future graduate students who spend time poring over the pages of theses acknowledgements wondering if they, too, will someday get there: A doctoral dissertation might give you much, but there is a lot of life out there to be lived beyond its pages. I hope you go seek it.

Table of Contents

Chapter 1: Remote Agree and Puzzles of Construal

This thesis explores a theoretical possibility that is made available by the Minimalist framework (Chomsky 2005), and proposes a solution to an empirical puzzle about construal phenomena—syntactic control and anaphora—that resists easy analysis. The goals are therefore both empirical and theoretical. While it may seem at first that the theoretical proposal presented in this chapter and elaborated in Chapter 3, the Remote Agree Hypothesis, might lack empirical application, I will argue that it provides a fruitful new possible explanation for the difficult puzzles of construal in long-distance and arbitrary control. The full scope of the implications of the theoretical proposal advanced is too broad to be treated here; in the final chapter I thus suggest potential fruitful avenues for future work.

We begin with the theory.

1.1 Minimalism and Generative Grammar

This work is situated in the generative grammatical tradition of linguistics. The "grammar" in this framework is a theoretical construct by which we understand the operations of the Language Faculty, at least parts of which we assume to be innate, and parts of which we assume to be specific to language rather than to general cognition ("we" being practitioners of linguistics in the generative grammar framework). The Grammar is not a model of production or comprehension. It is not expected to have psychological reality, but to be a theoretically useful construct in explaining how and why grammars do what they do (and the Grammar does what it does). Chomsky identifies the three criteria of descriptive, observational, and explanatory adequacy to evaluate our grammatical models, with explanatory adequacy being the highest benchmark. For a theory to achieve explanatory adequacy, its grammatical models must be able to account for how a learner (e.g. a child acquiring a native language) would

acquire that language. Thus, the model of the grammar constructed by the theory must not be so complicated or implausible as to be unlearnable.

Learnability in generative grammar is strongly linked to what we call Plato's Problem, also known as the poverty of the stimulus (POS) argument. Introduced by Chomsky (1980, 1986), it has been a guiding consideration in the generative enterprise. The POS argument states that the input to the child language learner is impoverished: not everything in the grammar can be learned only from the surrounding environment. This foundational assumption necessitates the existence of a Language Faculty, or some innate component of mind that is dedicated to and uniquely purposed for acquiring language. I assume a Language Faculty which is at least partly language-specific (rather than more broadly cognitive in scope) and at least partly innate. This is a very conservative and fairly weak claim about the Language Faculty, and one that I suspect is generally defensible despite recent arguments (Christiansen and Chater 2008, 2016; Hinzen 2012; Dresher 2018, among many others) that more language processes than we initially thought may be able to be handled by domain-general aspects of cognition.

Concerns about learnability inform our models not only at the level of the individual learner but also at the phylogenetic level. Hornstein (2009) claims that as generative grammar moved from the Government-Binding era to the Minimalist era, concerns about POS gave way to what he calls "Darwin's problem." That is, if (and this might be a large 'if') language is assumed to have evolved in a very short evolutionary timespan, it is likely the result of only a very small change in human cognition, such as the emergence of a single core (mental) operation. Because Merge, the operation that takes elements X and Y and combines them to form [$_{XP}$ X Y], is generally taken to be the most essential operation of the grammar,

Merge is usually the primary candidate for the single operation whose emergence triggered the near-instantaneous evolution of an entire Faculty of Language (FL) (Chomsky 2007).

It is sometimes said that Government-Binding theory and the Principles and Parameters framework solved the learnability problem at the ontological level, achieving explanatory adequacy and successfully addressing the Poverty of Stimulus problem. But the Government-Binding (GB) era theories were extremely elaborate, complex, with new hypothetically innate rules being proposed continually. The goals of syntactic theory changed: the theory had to account for the empirical coverage of GB but within the most simple, economical, elegant model possible. Regular parallels were drawn to theoretical physics, and syntacticians began to ask not just, *Does it account for the data? Is it explanatorily adequate? Does it address Plato's problem?* But, also: *Does it do all of the above with style?* In other words, is the theory parsimonious, is it elegant? (Chomsky 1995a, 2004, Hornstein et al. 2005, among many other works).

In the evolution of the theory towards more 'minimal' models, different kinds of parsimony were considered—different ways of meeting the minimalist criterion. For instance, Hornstein et al. (2005) distinguish substantive economy from methodological economy. Substantive economy captures how the mental grammar itself (the object we are aiming to model) is economical in terms of mechanisms or resources used. For instance, a derivation might be sensitive to the number of applications of a core operation or the number of lexical items used. Alternatively, the grammar might show a preference for a derivational path that is shorter than another, or less 'costly' computationally in some other way.

Ultimately, proposing that the grammar itself reflects computational efficiency may be somewhat misguided: (Lewis 2010b) notes that generative grammarians'—and particularly Minimalists'—concerns with computational economy, particularly of the

substantive kind, paradoxically occur within a system that we explicitly claim is *not* a model of production (i.e. how grammatical utterances are produced). Why, then, should our grammars care about computational efficiency, if they are not in fact tracking actual computational resources, and nothing is, in fact, being computed in real time/the real world?[1]

Methodological economy is in many ways a more justifiable foundational tenet of Minimalism, being essentially (as Hornstein points out in numerous works) the *Occam's Razor* principle of "all things being equal, the simpler explanation is likely correct." Applied to our linguistic model-making, methodological economy impels us to design the simplest model that covers the empirical explananda. This is (again in Hornstein's words) nothing more than "good science." The trouble with methodological economy, in practice, as a guiding principle between choosing competing theories, is that rarely is all else equal. The sample derivation used by Hornstein et al. (2005: 335–336) in their argument for the relative economy of Merge over Move cases is carefully constructed and not typical of such arguments in the literature. (In this derivation, the grammar picks a supposedly more economical operation, *Merge,* rather than *Move*, at a crucial derivational step, and Hornstein argues that this economy consideration comes into play precisely because the derivational paths are equal in every other way—i.e. same starting Numerations, same number of derivational steps either way. I will take issue immediately below with the idea that Merge is more economical than Move, but another takeaway from this constructed example is that is it exactly that: carefully constructed.)

---

[1] Lewis (2010a, 2010b, *inter alia*), provide insightful discussion and definitions of the types of computational concerns theoretical linguists worry about in our theoretical frameworks.

[2] It's possible this is also misguided. For instance, in the context of a system where we are not in fact measuring actual resources of production, it might be fair to ask whether any two operations have an

Theoretical appeals to economy almost always involve some simplification of the empirical territory. Furthermore, there is no agreed-upon and established set of "facts" to be accounted for, from the explanation of which we can select the most minimalist of competing candidate theories. It is sometimes suggested, by Hornstein and others, that the Government-Binding era set of facts should be taken as this starting point, this agreed-upon set of explananda. But the GB-era facts are heavily weighted towards English and other European languages, and such a stance ignores other empirical discoveries made since.

Approached from either angle, as a response to Darwin's problem or in the more traditional Minimalist terms of providing the simplest explanation for the richest, most complex set of facts, post-Minimalist generative theory is concerned with stripping down the model and theoretical apparatus so that they are as minimal as possible. In other words, we aim to propose a model no more complicated than necessary, which can still account for all the facts.

1.2 Minimalism and its Core Operations

While there are a variety of ways that Minimalist objectives might be approached and met, much of the discussion has focused narrowly on reducing the number of core operations employed by the grammar in syntactic derivations.[2] The core operations of the grammar, together with the basic building blocks of the lexicon, create an infinite number of sentences and utterances. In syntactic theory, a typical set of core operations consists of Merge, the operation that concatenates two objects (mentioned above), and Move, the operation that

---

[2] It's possible this is also misguided. For instance, in the context of a system where we are not in fact measuring actual resources of production, it might be fair to ask whether any two operations have an equal resource-cost. Having two core operations is not *ipso facto* less economical than having one, unless we know something *a priori* about the computational cost of these operations.

displaces an object within another object. Merge and Move are typically argued to be two types of application of the same operation (External Merge and Internal Merge). Other operations that have been proposed are Copy (Nunes 1995), which makes a Copy of a syntactic object, or Search, which looks for relevant goals. There are also operations which manipulate features (Match, Delete, etc.).

Different researchers argue for different basic operations, but there is a trend towards reducing the number of basic operations in the theory, such as reducing Merge and Move to one operation (Merge) with two flavours. (On the other hand, there are theorists who are not concerned with increasing the inventory of core operations, as long as they account for the facts under discussion, a different but also worthwhile route to linguistic exploration.) In many theories, Agree and Merge are taken to be the two primary operations of the grammar, where Merge includes External Merge and Internal Merge (Chomsky 2004 [2001]). External Merge concatenates syntactic objects, and Internal Merge displaces syntactic objects within a tree (itself a syntactic object). Both are conceptualized as Merge: External Merge (EM) is the concatenation of two syntactic objects neither of which is a subpart of the other, and Internal Merge is the concatenation of two syntactic objects one of which is a subpart of the other (see Ex.1). These are the operations I will assume henceforth.

(1) a. Merge (X,Y) $\rightarrow$ [$_{XP}$ XY]   (External Merge)

  b. [$_{ZP}$ Z [$_{XP}$ XY]]

   Merge (Y, ZP) $\rightarrow$ [$_{YP}$ Y [$_{ZP}$ Z [$_{XP}$ XY]]] (Internal Merge)

The debate about which are the core operations of the grammar and how many core operations the grammar has bears on both the Poverty of Stimulus issue and Darwin's problem, as well as our desire to be doing "good science," methodologically speaking. That is to say, it bears both on learnability and on the question of how language could have

emerged in such a short evolutionary timespan. Core operations of the Language Faculty are taken to be innate, and specific to language, rather than part of a general cognitive mechanism. In order to reduce the innate and language-specific component of the grammar as much as possible, we want to posit as few of these core operations as possible. Minimizing core operations also helps to explain how language emerged so quickly on the evolutionary timescale, to avoid redundancy in the model (Chomsky 2007, Hornstein 2009).

There are arguments for eliminating certain core operations and also arguments for choosing between core operations within a given derivation. For instance, Chomsky (1995a) introduces the principle of Merge-over-Move (MOM), later retracted in Chomsky (2004[2001]), to explain the data in (2):

(2)   a. *There seems a cat to be in the bathtub.
        b. There seems to be a cat in the bathtub.

The argument is that there are two derivational paths to be followed at the point of constructing the embedded infinitival: the grammar can either choose to raise [the cat] from its theta-position to [Spec, IP] of the infinitival, or it can merge expletive *there* from the Numeration. There is nothing inherently ungrammatical about (2a): it does not violate a principle of the grammar, and all features (Case, EPP, etc.) are checked and accounted for. But, Chomsky proposes, the grammar has a preference for merging the expletive at the point of constructing the infinitival, over moving the expletive's associate from lower in the structure, and so the economy condition MOM rules out (2a), which is ungrammatical.

I have argued against this principle in previous work (Motut 2009): not only do alternative accounts exist for contrasts like the one in (2), but the theoretical arguments for MOM do not hold. One argument advanced in favour of MOM was based on derivational economy: Move was two operations (Copy + Remerge) while Merge was just one. This

argument disappears if we abandon the assumption that Copy is an independent operation. If Internal Merge and External Merge are taken to be the same operation, the two derivations are of equal length. And, in the years since this argument was first advanced, essentially no strong independent empirical support for MOM has been found.[3]

The MOM principle, perhaps the most high-profile example of an argument for the economical primacy of one core operation over another, was retracted by Chomsky (2004[2001]), in favour of a Merge-only system. There are several arguments for this change, most of them coming from economy considerations like those discussed above. Chomsky (1995a,b) claims that Merge is the only "virtually conceptually necessary" operation, and therefore comes at no cost. In the Minimalist methodology, any additional operations beyond Merge thus need additional justification: they must be empirically necessary to account for the facts. The idea that Merge is "free" and that Merge is the only operation of the grammar are thus intrinsically linked. From the evolutionary perspective, Merge is often assumed to be the best candidate for the single change in the cognitive systems of pre-human individuals that led to the rapid development of the cognitive linguistic system that was the human Language Faculty. (For an alternative suggestion, see Hornstein 2009). Merge is thus often given theoretical primacy.

The primacy of Merge is also apparent in Epstein, Kitahara, and Seely's (2021, 2015) free-Merge project, a line of work taking Merge to be a necessary prerequisite to any kind of grammatical relation, and deriving a number of consequences from this proposal. Minimalist syntactic theory since Chomsky (2004[2001]) generally assumes that Merge has two varieties: External and Internal. The Epstein, Kitahara, and Seely (EKS) argument goes as

---

[3] Although the MOM principle has been cited in ven Gelderen (2004) in as playing a role in the diachronic process of grammaticalization.

follows: if EM (External Merge) is free, then IM (Internal Merge) must be as well. But IM is constrained in ways EM is not: it obeys locality, it must be triggered by the need to satisfy a featural requirement in the tree, etc. So in order to make this story work, the constraints on IM that do not appear to hold of EM must be explained in some other way. This is where so-called third-factor considerations, external to the language faculty itself (Chomsky 2005), are brought in.

Chomsky's (2013) position, along with that of EKS, is that Merge—in both its Internal and External forms—is "free," but conforms to third-factor considerations like Minimal Search and other constraints. This essentially moves locality and economy effects ("Minimal Search" as defined here) into the domain of general cognition. But this move may be suspect in that it shifts the burden of proof from the syntax to general cognition, but provides no concrete elaboration on how to capture the known empirical effects in a cognitively general formulation of Minimal Search. Let's consider the two perspectives on Internal Merge, given in (3). What, in fact, is the difference between the statements in (3a) and (3b)?

(3)   a.   (Internal) Merge is free but conforms to third-factor considerations like
                Minimal Search (Chomsky 2013, Epstein, Kitahara, and Seely 2021).
         b.   (Internal) Merge is not free and conforms to locality.

In (3a) the kinds of locality effects we typically find in the grammar's derivations (e.g. Relativized Minimality, Rizzi 1990, etc.) are not language-specific but rather part of some larger general cognitive preference for conservation of resources. In (3b), the familiar notion of locality functions within a syntactic object/tree: closer goals block more distant ones, and

this is usually relativized to the type of probe involved.[4] But claiming that the operation(s) of the grammar itself are not constrained by a type of locality specific to language is just a way of shifting the burden from one part of the theoretical machinery to another, in this case reducing the degree of scrutiny the mechanism is subject to. We don't yet have reliable ways of distinguishing which of our principles are cognition-general and which are language-specific.

The free-Merge project follows the deduction in (4). That Internal Merge must be unconstrained follows from the assumption that IM and EM are the same operation, and if one is unconstrained, so too must the other be.

(4) a. EM is free;
  b. EM = IM;
  c. Therefore, IM is free.

The observed constraints on IM (locality effects, etc.) must therefore be the result of "third-factor" economy considerations, and we arrive at (3a). However, the implication also works in the opposite direction: if EM and IM are equivalent, then the properties of IM must hold of EM. Thus, if IM has certain constraints, EM should also have them. Thus we arrive at the opposite implication, given in (5) below.

(5) a. IM is constrained by locality and triggered by features;
  b. EM = IM;
  c. Therefore, EM is also constrained by locality and triggered by features.

The free-Merge project may perhaps be influenced by the historical view that (External) Merge is basic. This may be partially why the implication in (4) takes primacy over that in

---

[4] I say "the familiar notion of locality" when in fact there are many different ways in which linguists can and do understand locality. Here I mean locality defined in terms of c-command or command relations within a tree. I propose a different view of locality later on.

(5). However, there is no principled reason to assume that the properties of External Merge (e.g. its 'freeness') should take priority over those of Internal Merge. In other words, there is no obvious reason why the implication in (4) should be prioritized above that in (5).

In fact, there are so-called triggered-Merge proposals as well. Such proposals argue that Merge is not free but must be initiated by a feature or probe, as with Internal Merge. But what features trigger Internal Merge that might be extended to External Merge? Wurmbrand's (2014a) proposal uses phi features to trigger External Merge of arguments into the clausal spine. And Chomsky (2001) proposes "Edge" Features, which are like EPP (Extended Projection Principle) features for Internal Merge, to trigger External Merge. Collins' (2002) proposes the Subcat Relation in (6), which describes how External Merge is triggered by a featural requirement of one of the syntactic objects being merged. Pesetsky and Torrego's (2006) Vehicle Requirement on Merge makes a very similar proposal (7).

(6)   Subcat Relation (Collins 2002: 54–55):
        Let X (a lexical item) be in SO1, and let Y (a lexical item or feature) be in SO2, where SO1 and SO2 are two independent syntactic objects (lexical items or phrases). If X and Y enter into a Subcat relation, then Merge (SO1, SO2).

(7)   Vehicle Requirement on Merge (Pesetsky & Torrego 2006, cited in Wurmbrand 2014a: 131):
        If α and β merge, some feature F of α must probe F on β.

Rezac (2002:19–20) also proposes what he calls "theta-theoretic Merge," and suggests that Merge and Move can be unified if Merge is also feature-driven:

"[…] There is a feature [Σ-], and classical EPP is an instance, that just demands an XP, however it gets it including initiating its own Match + Merge. If this [Σ-] feature then just seeks an XP, Locality will not let it look very far: not beyond the closest XP, give or take a bit. Match and Agree are how we deal with such feature relations. In the movement case, [Σ-] triggers Match with its Locality, finds the closest XP, and Merge

to [Spec, TP] follows. In the base Merge case, the Match is in the Numeration. […] The way we've just implemented EPP, theta-theoretic Merge can also be implemented […]." Rezac (2002:19)

Rezac's proposal is in many ways similar to Chomsky's Edge features, which are like EPP features for External Merge. Their only purpose is to trigger concatenation of syntactic objects. Wurmbrand's and Pesetsky and Torrego's proposals, in contrast, make use of features that are already doing other work in the tree (e.g., phi features).[5]

In contrast to free-Merge proposals, triggered-Merge proposals prioritize the empirically observable constraints on IM, and therefore adopt the implication in (5), rather than the primacy of EM and the implication in (4). Triggered-Merge proposals also offer a new perspective on the relation between the two kinds of Merge and locality. While free-Merge proposals make applications of IM or EM equal and interchangeable, triggered-Merge proposals create a new possibility: whether a derivation chooses an external application of the operation or an internal one might depend on the same kind of locality or Minimal-Search considerations that determine locality within a given tree/syntactic object. In other words, triggered-Merge proposals suggest the possibility of extending something like Relativized Minimality beyond a single syntactic object. This is a key part of the proposal to be made here. I will argue below that triggered-Merge proposals also shed new light on certain empirical phenomena of unlicensed but referentially dependent syntactic objects.

What about Agree? Hornstein (2009, *inter alia*) has argued that the the coverage of Agree is essentially the domain of long-distance movement, and it can all be reduced to

---

[5] We might also consider category features (O'Neil 1995, Martin 1996, Hornstein 1999, Manzini & Roussou 2000), which can trigger EM and satisfy theta-role requirements. However, Motut & Oda (2011a,b), on which the proposal to be made in Section 1.3 is based, did not explore this possibility.

Move (and therefore Merge, since Move is just Internal Merge). Above, I suggested that we ought to be concerned with how many core operations are in our grammar, as part of our general commitment to methodological minimalism (e.g. "good science"). But I also cast some doubt on the idea that fewer operations necessarily result in a more economical system. The system I propose below relies on a probe-goal Agree operation and Internal/External Merge operation(s). Let's briefly examine some reasons for keeping Agree in the picture.

First of all, Agree covers substantial empirical ground, which would need to be covered by another operation. It has been argued by Hornstein (2009) that covert movement (i.e. Internal Merge) could cover the same empirical territory as the Agree operation. However, this would not account for cases where "moved" elements are not identical—most instances of feature sharing or feature transmission between heads and other non-head constituents require some version of an Agree-dependency rather than covert movement. See Boeckx (2009) for some arguments along these lines. I will propose a system using both Agree and (Internal and External) Merge. The proposal for Remote Agree has its origins in joint work with Kenji Oda. Much of the following section (and especially the examples) follows fairly closely follows the argumentation from Motut & Oda (2011a,b).[6]

1.3 The Remote Agree Hypothesis

1.3.1 From Triggered External Merge to Remote Agree

Although we should not dispense with Agree, External and Internal Merge could be better unified: this is a goal I share with the Free-Merge proponents. If EM and IM are the same operation, it should not be the case that one is "free" and the other is subject to various

---

[6] The initial idea for Remote Agree was Kenji Oda's. I extend and further refine it here, and explore its implications in specific empirical domains. In Section 1.3 the argumentation and examples follow very closely the argumentation in Motut & Oda (2011a,b).

constraints. The unification of Merge and Move which began with Chomsky (2004[2001]) is thus incomplete: recasting these operations as Internal and External variants of Merge, while unifying the two operations in definition, did not unify them in terms of their characteristics. Like the Free-Merge proponents, I aim to unify the properties of these operations, but rather than taking the properties of External Merge to be primary and following the implication in (4) above, I hold the properties of IM to be primary and prioritize the implication in (5).

One reason to prefer this route is that the properties of Internal Merge have been extensively studied and articulated in great theoretical detail, with substantial empirical coverage. Locality effects within a single syntactic object are well understood. For instance, closer goals block more distant ones: if a featural requirement of a probe can be satisifed by two syntactic objects, the closer one (where "closer" is usually defined in c-command terms) will satisfy the probe. We will review this in more detail immediately below. That Merge is parasitic on Agree in internal operations is also well understood and empirically supported. However, the "freeness" of External Merge is argued only on theoretical grounds. While the virtual conceptual necessity of Merge is appealing, it does not follow that all locality effects can or should be attributed to third-factor effects like "Minimal Search."

In the probe-goal model of operations articulated in Chomsky (2000/2001) and elaborated on subsequently by many researchers (e.g., Béjar 2003, Béjar and Rezac 2009), Internal Merge is "parasitic" on Agree, meaning that an Agree relation precedes the application of Internal Merge. Objects may be displaced in the tree if an Agree relation between a probe and goal is established; however, a probe-goal dependency may also be established without subsequent Internal Merge of the goal. In (8), once the probe T with an unvalued phi-feature finds the DP with visible phi features, the two form an Agree dependency such that T's unvalued features will be valued by the features of the DP.

14

Movement of the DP to a [Spec, TP] position may follow. The Agree relation thus precedes, and feeds, the movement of the DP.

(8) *Agree:*

a. $[_{TP}$ $T_{u\Phi}$ … $[_{vP}$ $DP_{\Phi:3sg}$ v … $[_{VP}$ V …$]]]$

b. $[_{TP}$ $DP_{\Phi:3sg}$ $T_{\Phi: 3sg}$ … $[_{vP}$ t_{DP} v … $[_{VP}$ V …$]]]$

Expanding this property of Internal Merge to External Merge opens up an unexplored possibility, represented in Table 1 below. Within the syntactic tree, we have two options: dependency formation (Agree) with parasitic (Internal) Merge, and dependency formation without (Internal) Merge. By the logic in (5) above, which extends the properties of the internal operation to its external counterpart, we should end up with the same two possibilities for External Merge.

Table 1

	Internal (two SOs in one syntactic object/tree)	**External (two separate SOs)**
Dependency and Displacement	Agree + (I-)Merge	Agree + (E-)Merge
Dependency, no displacement	Agree (within the tree)	Agree (beyond the tree) = "**Remote**" **Agree**

The usual view of External Merge is that Agree is not involved at all. We have seen, however, that proposals for Agree-triggered EM (Table 1, top right cell) do exist, and that there may be some (theoretical) reasons to prefer them. The option of such an Agree relation being established without triggering Merge—what I call above "Remote Agree"—is to my knowledge not explored at all in the literature except by Motut & Oda (2011a,b) (referred to in the remainder of this section as Motut & Oda), at least not in such a concrete way. With this option, a probe within the tree with unvalued features (T_{uF}) agrees with a syntactic object (SO) in the Workspace (X_F in [9]). (We assume for the moment that the DP already in the

tree does not intervene, just to illustrate the point.) This Agree relation does not trigger Merge of that SO. This is parallel to the internal case, where displacement of the goal does not necessarily follow Agree with a probe. In this case, a dependency is created between a probe on a syntactic head within a complex SO, and an element in the Workspace, which remains in the Workspace. This constitutes an Agree relation between a probe and an element not included in the syntactic object that contains the probe. The two possible outcomes of an external Agree operation are shown in (9) and (10).

(9) *Remote Agree*: *External Agree without External Merge*

$$[_{TP} T_{uF} \dots [_{vP} DP_{\Phi:3sg} \; v \dots [_{vP} V \dots]]] \dots \{ \dots X_F \dots \}$$

Agree

(10) *Agree-Triggered External Merge*:

    a. $[_{TP} T_{uF+EPP} \dots [_{vP} DP_{\Phi:3sg} \; v \dots [_{vP} V \dots]]] \dots \{ \dots X_F \dots \}$ $\rightarrow$

Agree

    b. $[\; X_{\cancel{F}} \; [_{TP} T_{\cancel{uF+EPP}} \dots [_{vP} DP_{\Phi:3sg} \; v \dots [_{vP} V \dots]]]]$
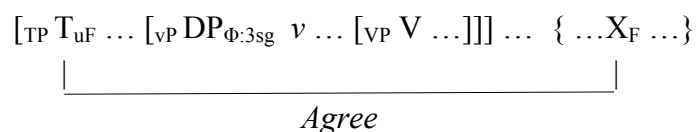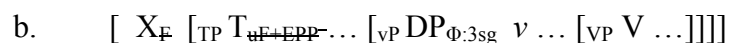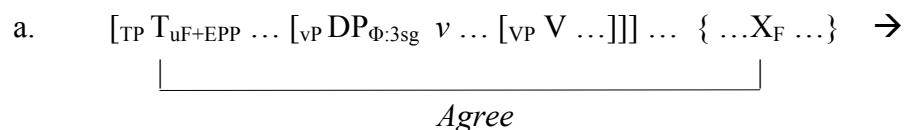
        This proposal of course immediately raises many important questions: How does such an object get Spelled Out? How do we avoid overgeneration? These questions will be addressed in Chapter 3. First, I focus on the theoretical motivation for the proposal rather than its empirical implementation: that is, how does Remote Agree naturally fall out from the current configuration of Minimalist assumptions, and what are its general theoretical benefits to the system? Its possible empirical uses will be explored in later chapters.

1.3.2 External Agree in Motut & Oda (2011): EPP and Edge Features

In Internal operations, the features that drive the application of Merge are operationalized as EPP features: some languages/derivations have EPP in addition to the phi features that

establish the Agree relationship, and so Internal Merge happens. For Remote Agree, Motut &
Oda operationalize the features triggering this external application of the operation as EPP or
Edge features, as in (11) and (12) where P is the probe and G is the goal. The first option in
the figure below, from Motut & Oda (2011a,b) but representing Chomsky (2000), shows the
feature on the probe, P, being satisfied by movement of the goal, G to the specifier of the
probing head. The second option, indicated by a dotted line, demonstrates Agree as the
operation Remerge, which was an option explored by Motut & Oda, following Starke (2001).

Internal Merge (Chomsky 2001) (Figure from Motut and Oda 2011a,b)

(11)

$P_{[uF\text{-EPP}]}$ ... $G_{[F]}$ → $G_{[F]}$ P ... $<G_{[F]}>$ or P ... $P_{[uF\text{-EPP}]}$ ... $G_{[F]}$
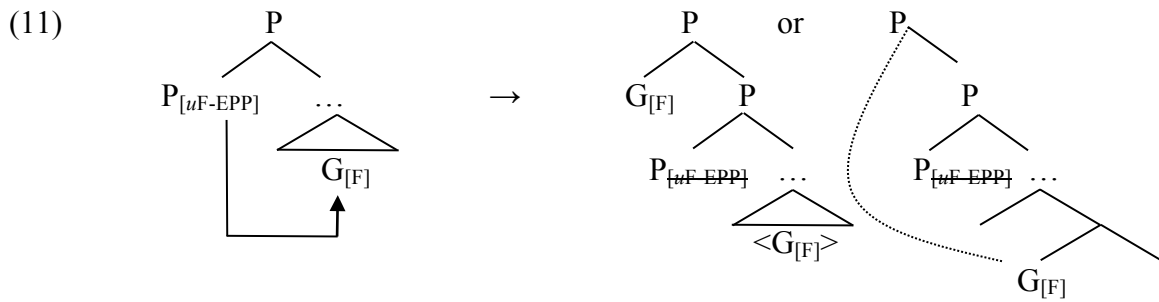
External Merge Triggered by Agree (Motut & Oda 2011a,b)

(12)

$P_{[uF\text{-EPP}]}$ ... ... $G_{[F]}$ → $G_{[F]}$ P $P_{[uF\text{-EPP}]}$ ...
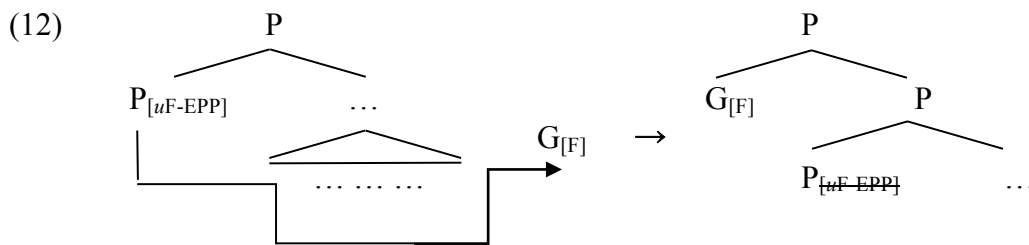
In Internal Merge, the probe is the head containing the EPP feature. The EPP feature piggy-
backs on the agreement between the unvalued feature of the probe and the corresponding
feature of the goal, but the EPP feature does not itself probe: the phi features (or other
substantive unvalued features) probe. Motut & Oda propose that the same mechanism, an
EPP feature on the probe that is parasitic on the probing unvalued (phi) feature, be extended
to External Merge. In (12), Merge is triggered by the unvalued feature *uF* (the probing

feature), and P's ability to trigger an instance of E-Merge is operationalized as an EPP feature on P. (The EPP is a feature on the head itself rather than a feature on a feature.)

This differs somewhat from Chomsky's (2013) use of Edge features, which make a syntactic object available to be Merged. Wurmbrand (2014a), on the other hand, as in the Motut & Oda system, employs some familiar unvalued features, like phi, to trigger the applications of E-Merge. I will assume a feature-driven approach to External Merge here.

The example in (13) shows how a derivation would proceed with a simple intransitive sentence in the Motut & Oda system. The unvalued feature, uF, with parasitic EPP, triggers applications of External Merge in (b) and (d). In (c) an unvalued theta feature triggers the E-Merge of the pronoun into the argument slot of the light verb. Precedent for the use of theta-features is found, for example, in Hornstein et al. (2005), *inter alia*.

Derivation of a simple intransitive sentence (Motut & Oda 2011a,b):

(13)   She laughs:

     a. $\{ \ldots D_{she[\varphi:3sg]}, T_{[u\varphi:\_\_, uD\text{-}EPP, uF3\text{-}EPP]}, v_{[u\theta, uF1\text{-}EPP]}, V_{laugh} \ldots \}$

     b. v probes with $[u\text{F1-EPP}]$ and finds V_{laugh}, and they (External) Merge:

            $\{ \ldots D_{she[\varphi:3sg]}, T_{[u\varphi:\_\_, uD\text{-}EPP, uF2\text{-}EPP]}, [_v v_{[u\theta, \sout{uF1\text{-}EPP}]} V_{laugh}] \ldots \}$

     c. v probes with $[u\theta]$ and finds $D_{she[\varphi:3sg]}$ and they (External) Merge:

            $\{ \ldots T_{[u\varphi:\_\_, uD\text{-}EPP, uF2\text{-}EPP]}, [_v D_{she[\varphi:3sg]} [_v v_{[\sout{u\theta}, \sout{uF1\text{-}EPP}]} V_{laugh}]] \ldots \}$

     d.   T probes with $[u\text{F2-EPP}]$ and finds vP and they (External) Merge:

            $\{ \ldots [T_{[u\varphi:\_\_, uD\text{-}EPP, \sout{uF2\text{-}EPP}]} [_v D_{she[\varphi:3sg]} [_v v_{[\sout{u\theta}, \sout{uF1\text{-}EPP}]} V_{laugh}]]] \ldots \}$

     e.   T probes with $[u\text{D-EPP}]/[u\varphi:\_\_]$, finds $D_{she[\varphi:3sg]}$ and these (Internal) Merge:

            $\{ [_T D_{she[\varphi:3sg]} [_T T_{[\sout{u\varphi:3sg}, \sout{uD\text{-}EPP}, \sout{uF3\text{-}EPP}]} [_v D_{she[\varphi:3sg]} [_v v_{[\sout{u\theta}, \sout{uF1\text{-}EPP}]} V_{laugh}]]] \}$

1.3.3 Remote Agree and Locality Domains

Being probe-triggered and therefore parasitic on Agree is the first property of IM that I propose to expand to EM. The second property that should be expanded from IM to EM, if

the implication discussed above is true, is sensitivity to locality. Locality is pervasive in the

grammar as a well-understood constraint on structural relations internal to the tree. We see

locality effects, for instance, in Superiority effects (14) and in Head Movement (15). In (14)

only the closest wh-word can move to satisfy the probe in the highest clause. In (15), only the

highest auxiliary (head) can move. These examples together illustrate a fundamental

characterization of locality, formalized by Rizzi (1990) as Relativized Minimality, given in

(16). Relativized Minimality crucially highlights the sensitivity of grammatical operations to

intervenors of the same type (e.g. head, DP, Wh-word), or in other words, locality based on

the type and sensitivity of the probe or the required goal. Locality is also sometimes defined

in terms of phases: it is an open problem in the field how to unify these two types of locality.

(14) *Superiority:*

     a. *[$_{CP}$ Who$_1$ does John think [$_{CP}$ who$_2$ would believe [$_{CP}$ that Mary would kiss

       t$_{who-1}$]]]

     b. [$_{CP}$ Who$_2$ does John think [$_{CP}$ t$_{who-2}$ would believe [$_{CP}$ that Mary would kiss who$_1$]]]

(15) *Head Movement Constraint* (Travis 1984)*:*

     a. Could Alex have eaten the pie?

     b. Has Alex eaten the pie?

     c. *Have Alex could eaten the pie?

(16) *Relativized Minimality* (Rizzi 1990): In a configuration [... α ...γ ... β], where α c-

     commands γ and γ c-commands β, γ blocks a relationship between α and β iff γ is of the

     same type as α [...].

Broadening the locality constraints of IM to include EM provides a starting point to

disentangle what Minimalism has to say about applications of IM vs. EM: in other words, we

can say something about the computational cost, or relative locality of a derivation choosing

one or another of these operations at a given stage. In free-Merge accounts, by contrast, both

IM and EM are equally available, since both are cost-free.

In addition to making EM parasitic on Agree, like IM, Motut & Oda propose to extend the Search space of a probe to the syntactic Workspace, making EM constrained by locality in much the same way that IM is. Again, there is some precedent for this idea: van Riemsdijk (2006: 20), in motivating his proposal for "grafting" (see also Citko 2000, 2005 on Parallel Merge), says the following:

> "[The] necessity to constrain grafting [i.e a new case of EM] is not different, in principle, from the necessity of constraining, say, internal merge. The only difference is that we have worked on constraining internal merge (formerly move) for several decades now, while the program to constrain external-internal merge (that is, graft) is only just starting."

His point is that, like Internal Merge, other kinds of Merge must have constraints which await formalization. Chomsky (2013) has also gestured in the direction of Motut & Oda's proposal, remarking that, from a given probe within a given SO, it is "less search"—more local—to apply Internal rather than External Merge:

> "In earlier years it was assumed generally (by me in particular) that displacement is problematic, an 'imperfection' of language, to be explained by external factors, perhaps parsing considerations. A residue of that error is the belief that EM is somehow simpler and preferable to IM. There is no basis for that belief. If anything, IM is simpler, since it requires vastly less search than EM (which must access the workspace of already generated objects and the lexicon)." Chomsky (2013: 41)

However, in this same work, Chomsky immediately pivots and goes on to say, "but the question doesn't arise: both are freely available, given the simplest combinatorial operation, Merge." He thus ultimately opts for the free-Merge hypothesis: both IM and EM are equally available and do not compete with each other, in terms of locality or in any other way computationally. I do not follow him down this route, but find his brief remarks in the

opposite direction tantalizing and informative. Motut & Oda (2011a,b) in fact take the road

not taken by Chomsky and propose that IM requires less Search and is therefore preferred by

locality over an application of EM. The idea that IM might take less "work" computationally

is also picked up by Roeper (2013), who argues based on evidence from child language

acquisition that Internal Merge is less costly computationally than External Merge, for

similar reasons.

In addition to making EM parasitic on Agree, Motut & Oda propose that locality can

be re-defined as in (17) below: probes can Search beyond the syntactic object that contains

them, and goals within the same syntactic object as the probe, are closer than syntactic

objects in the Workspace or in a different syntactic object. Note that this does not address the

case where a probe targets a goal that is a sub-part of a different syntactic object (sideward

movement, Nunes 2004). The term "structure-internal" in (17) means within the same

syntactic object as the probe; "structure-external" means in the syntactic Workspace or not

within the syntactic object containing the probe.

(17)   Redefining Locality for EM (from Motut & Oda 2011a,b):
       Where P is a Probe, and G_1, ..., G_n are potential goals, P Agrees with the closest Goal
       as follows:
       (i)   If P c-commands only 1 structure-internal Goal, G_1, then G_1 is the closest
                goal;
       (ii)   If P c-commands more than 1 structure-internal Goal, G_1, ..., G_n, where G_1
                asymmetrically c-commands G_2, ..., G_n, G_1 is the closest goal;
       (iii)   If P c-commands no potential structure-internal goals, a structure-external
                Goal, G_3, is the closest goal;
       where two elements are "structure-internal" iff there is some node which dominates
       both elements (and one element is "structure-external" to another iff there is no node
       which dominates both elements).

(17) (i) – (ii) reiterate the familiar notion of Minimality as it applies within a single syntactic object, as illustrated in (18)). (17iii) departs from previously formulated notions of locality to describe how locality might work beyond the domain of a single syntactic object. It states that a probe may search for a goal beyond the sub-tree or syntactic object in which it resides, if and only if it does not find a potential goal within that syntactic object.[7] This is represented in (18).

Agree Constrained by Locality (Chomsky 2001) (Figure from Motut & Oda 2011a,b)

(18)　a.
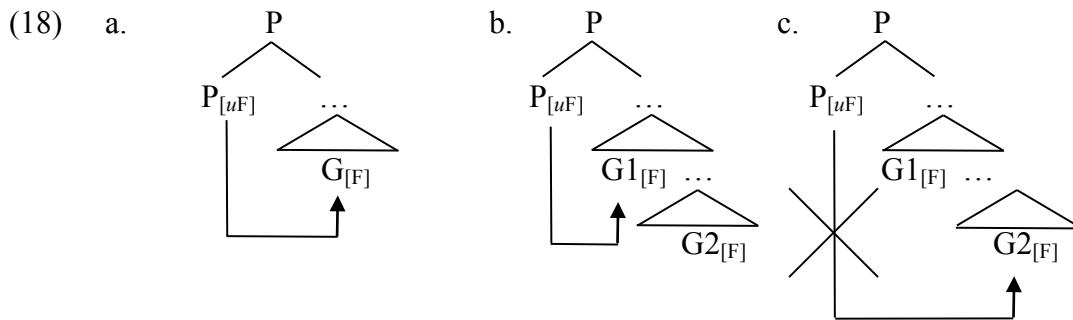
This formulation of locality prioritizes an internal application of Merge over an external one—a kind of Move-over-Merge preference (or IM over EM). However, unlike the original Merge-over-Move principle, this Move-over-Merge preference is not a stipulated property of the grammar. Rather, it is an automatic consequence of applying general principles of locality to the syntactic Workspace. It is not locality itself that is being redefined here; rather, the search space of the probe is being expanded, with consequences for locality.

In the Motut & Oda system, all objects in the workspace are equidistant. Once the probe searches within its SO and finds no potential probe, it then searches beyond that SO.

---

[7] I assume here that 'potential goal' encompasses commonly understood aspects of the mechanics of the probe-goal system and Relativized Minimality, specificity of the probe, etc.

All workspace objects are equidistant from the probe, and any one with suitable features will be a potential goal. This is illustrated in (19).

(19) { [$_{XP}$…**P**…], …, [$_{IP}$ John met a man], [$_{DP}$ the large cat], …, v^0, D^0, L_1, L_2, …, L_n, L_{n+1}, …}

First Merge, that is, Merge of two atomic syntactic objects, is derived as follows: E-Merge is triggered by a probe on a syntactic head (a simplex item such as P in [20]). Such an element has no complement and therefore no potential goal within its SO to block Search into the Workspace: The Workspace contains A, B, and C, but only B has a matching feature [F]. B is therefore the 'closest' potential goal for probe P, a simplex syntactic object.

First Merge as Probe by a Simplex Item (Motut & Oda 2011a,b):

(20)   P$_{[uF\text{-}EPP]}$ … A … B$_{[F]}$ … C …

The system outlined here also eliminates economy-based stipulations like Merge over Move that compare Internal Merge to External Merge and prioritize one based on an assessment of its being somehow computationally "simpler." There is an effective preference for Move over Merge, but as explained just above, it follows from the expanded principles of locality proposed here and to be argued for. This version of locality privileges applications of IM, with less Search, over EM, which requires farther Search, i.e. beyond the SO containing the probe. This is not the same as a trans-derivational comparison and a prioritizing of one core operation over another. In this system, IM and EM really are one core operation, consistently triggered by Agree and constrained by locality. Dispensing with Merge over Move is also a welcome effect, since as argued elsewhere (Motut 2010) there is no empirical basis for this principle. Chomsky (2013) also notes that the principle was incorrect. Shima (2000) has argued for a Move-over-Merge principle rather than Merge over Move, but his logic is

similar to that in Chomsky (1995a) for the original MOM principle. In particular, Shima uses

trans-derivational comparison and argues, based on superraising constructions, that Move is

prioritized over Merge. Neither Shima's proposed Move over Merge nor Chomsky's Merge

over Move unify IM and EM in terms of their properties, and neither has had much empirical

support. This kind of economy condition is eliminated in the approach proposed here, in

favour of a modest expansion of the probe space in the probe-goal system and the consequent

expansion of locality effects beyond the domain of a single syntactic object.

What has been shown so far is that Remote Agree is an available operation given the

assumptions articulated above: that Internal Merge is probe-driven and constrained by

locality and that Internal Merge and External Merge are the same operation. I have also

explored how doing things this way gives us insight into how the derivation chooses between

EM and IM (as opposed to the free-Merge approach, under which both are equally available).

Further details of this system will be developed in later sections, and in some places will

diverge from the proposals made by Motut & Oda. I will argue that Remote Agree functions

as a last-resort mechanism at the edge of certain syntactic objects, and discuss how a

particular view of Spellout prevents the system from overgenerating.

I turn next to several empirical puzzles in the domain of control and anaphor binding.

These will be shown in Chapter 4 to be amenable to analysis with Remote Agree. The first

overview is broad: a more detailed review of key phenomena to be analyzed by Remote

Agree will appear in Chapter 2.

1.4 A Few Empirical Puzzles

We now turn to a sketch of the empirical heart of this thesis. There is a group of related

phenomena in the area of construal, involving syntactic elements that are dependent on other

syntactic elements in order to receive reference. These include controlled PRO as well as anaphors.[8] There are two parallels between control and anaphora that I will focus on. First, long-distance and so-called "unlicensed" versions of control and long-distance-bound anaphors pattern similarly. There is also a parallel across control and anaphora, in that both of these domains exhibit the same patterning between the unlicensed and long-distance cases, which we will group together under the term "non-local" construal phenomena. Both of these parallels are visible in the data below.

(21) *Control phenomena:*

    a. **John** wants [**PRO** to eat all the pastries]. (Obligatory Control)

    b. **John** thinks [that Mary said that [**PRO** helping himself to another piece of pie would be greedy]]. (Long-distance control)

    c. [**PRO_{arb}** to eat the last piece of pie] is a common desire. (Arbitrary Control)

(22) *Anaphor Binding*

    a. **John** told **himself** that eating the last pastry would only anger his colleagues. (Local Anaphor Binding)

    b. **John** expected [that Mary would criticize no one but **himself**] (if he took the last pastry) (Long-Distance bound Anaphor)

    c. Tomo and **myself** are going to share the last piece of pie. (Unlicensed/untriggered anaphor)

        In each set of examples in (21) and (22) there is a licenser and a licensee. The examples in (21a) and (22a) both have **local, obligatory dependencies** between licenser and licensee. In (21a), *John*, the subject of the main clause, controls PRO in the complement clause; in (22a), *John*, the subject of the main clause, binds the anaphor *himself* in object position of the same clause. (21b) and (22b) show **long-distance dependencies** between

---

[8] There may be cases where the referent is not found syntactically but in the discourse context or semantics.

licenser and licensee: in (21b), *John* controls PRO across an intermediate clause containing a potentially intervening DP, *Mary*. In (22b), *John* binds the anaphor *himself* across an intermedicate clause and intervening potential binder, *Mary*. In (21c) and (22c), the dependent elements, PRO and the anaphor respectively, **appear to lack a licenser altogether**. In (21c), PRO in the subject clause is not controlled by another DP; in (22c), *myself* is not bound by *Tomo* or any other DP—there is no other DP to bind this anaphor. The parallels between (21a) and (22a), between (21b) and (22b), and between (21c) and (22c) are the first part of the puzzle: these are local, long-distance and apparently unlicensed cases in both control and binding. There is also overlap in the distributions of the two types of non-local cases (the long-distance and unlicensed ones), in contrast to the local cases in (21,22a). I will discuss this overlapping distribution between (21,22b) and (21,22c). The source of both of these parallels is a puzzle.

Let's take a closer look at the classification and grouping of these phenomena. First, I describe a distinction between the **local** and the **non-local** phenomena, for both control and binding. I will demonstrate briefly that long-distance control patterns with arbitrary control, and that long-distance anaphora pattern with untriggered or unlicensed anaphora. This is the first puzzle of overlapping distributions. Further distinctions are required in order to capture different types of long distance and untriggered/unlicensed phenomena, but we will do so in due course.

The second puzzling parallelism, as mentioned above, is that existing between the control cases and the anaphora cases. The fact that both of these phenomena have long-distance licensed and unlicensed versions also needs to be explained. I will argue that this second parallelism may point to a common mechanism between control and binding and explore whether control and binding phenomena can be unified by a similar mechanism. For

26

control, in particular, I will explain in detail in Chapter 4 how use of the Remote Agree

mechanism can unify long-distance control and arbitrary or unlicensed control, one half of

our parallelism puzzle that can be fruitfully analyzed using Remote Agree.

1.4.1 Lack of Licensing is a Puzzle

The lack of licensing in sentences like (21c) and (22c) itself presents a puzzle.

Anaphors and PRO are defined as dependent on other elements for their reference. The fact

that in both these domains there is a systematic class of cases that appear to entirely lack this

defining property, and, as we will see, show exactly the mirror opposite properties of the

local and obligatory cases, is startling. These uncontrolled PROs and unlicensed anaphors are

only superficially like their obligatorily bound or controlled counterparts. Proposals have

been made, which we will examine in more detail in Section 4.2, for how these types of PRO

and anaphors might be instead licensed by the discourse or the semantic-pragmatic

component of the grammar directly, rather than by the syntax. I will suggest that, in

particular for arbitrary control (but also for unlicensed reflexives), the Remote Agree

hypothesis presents an explanation.

1.4.2 Local versus Non-Local Phenomena

As noted above, in both control and binding, the grammar is confronted with the

problem of assigning reference to a dependent element like an anaphor or a PRO. In the

section below, we'll look at how this is done in the typical, local (and generally obligatory)

cases, and compare these cases to the non-local ones in the next section.

Local phenomena: Local control and anaphora (21a) and (22a) have the properties in (24):

(24)   Properties of the Local Control/Binding Phenomena

(i) The referentially dependent element (PRO or anaphor) must have a suitable antecedent.[9]

(ii) That antecedent must be local.

(iii) The antecedent must c-command the referentially dependent element.

The effect of (24i) is illustrated in (25a) for control and (25b) for binding. *PRO* lacks an antecedent in (25a) since non-referential expletive *it* is not suitable as an antecedent, and *John* is an unsuitable antecedent in (25b) due to mis-matched number features. (For anaphor-binding, generally an antecedent is acceptable if it matches in person, number and gender with the anaphor.)

(25)   a. *It was expected **PRO** to shave. (Cf. John was expected PRO to shave.)
        b. ***The children** helped **herself**. (Cf. The children helped themselves.)

(24ii) is illustrated in (26a) for control and (26b) for binding.

(26)   a. ***John** thinks [that it was likely [**PRO** to help himself]]
        b. ***Alex** thought [that the electrical outlet couldn't hurt **herself**.][10]

(24iii) is illustrated in (27a) for control: John does not c-command PRO (*John's sister* does), and in (27b) *John* does not c-command the anaphor and so is not a potential antecedent.

(27)   a. *[**John**'s sister] tried [**PRO** to help himself to cookies].
        b. *[**John**'s sister] helped **himself** to cookies.

Local, obligatory *anaphor binding* typically occurs in the configuration in (28), where there may be intervening material between the antecedent DP and the anaphor, including whole clauses, as long as the anaphor is still local to and c-commanded by the antecedent:

(28) $[_{CP/TP}$ DP$_i$ ..... [T ... [v/V ... Anaphor$_i$]]]

---

[9] What constitutes a suitable or possible antecedent will differ between PRO and anaphors, and across languages.

[10] There might be two alternative derivations available here, one where the anaphor *herself* is unlicensed or unbound. Some speakers may find this grammatical.

Local obligatory *control* occurs in a similar configuration, with a c-commanding DP

antecedent (29). PRO always occurs as the subject of a non-finite clause.[11]

(29)   [$_{CP/TP}$ DP$_i$ ... [$_{CP/TP}$ PRO$_i$ T$_{inf}$ ......]]

Non-local phenomena:

Strikingly, the properties of both non-local anaphora and non-local control are the mirror

opposite properties of the local, obligatory cases, although the obligatory cases are taken to

be the core phenomena. The properties of the non-local cases are given in (30):

(30)   Properties of Non-Local Control/Binding:
   (i)   The referentially dependent element (PRO or anaphor) does not require an
            antecedent
   (ii)   If there is an antecedent, it need not be local.
   (iii)   If there is an antecedent, it does not necessarily c-command the referentially
            dependent element.

These properties are illustrated in (31) – (33). In (31a), PRO$_{arb}$ has no antecedent and is

understood to refer to persons in general. I discuss the particular properties of arbitrary PRO

in more detail in Section 4.1. In (31b), *myself* refers to the speaker, who is not overtly

represented in the sentence.[12] In (32a), *John* controls PRO across an intervening clause

containing another potential DP antecedent (*Mary*). In (32b), *John* binds *himself* across an

intervening potential antecedent (*Mary*). *John* is not local to (i.e. within the same clause as)

either *himself* or PRO in (32). Property (30iii) is illustrated in (33): In (a) *Alex,* buried in the

complex DP *Alex's friend*, does not c-command PRO although *Alex* is clearly meant to be

---

[11] This is the case in English, but contested in other languages such as Brazilian Portuguese. See
(Boeckx et al. 2010) and references therein for more details.

[12] Some syntactic proposals to account for speaker-linked phenomena like this actually do locate a
"speaker" antecedent in the syntactic tree: we will return to these accounts briefly in Chapter 5.

interpreted as the agent of *keeping*; and in (b), *John* likewise does not c-command the anaphor *himself* with which it is co-referential.

(31)   a. It is illegal PRO$_{arb}$ to park here.

        b. Linguists like myself generally have a good sense of humour.

(32)   a. John$_i$ thought that Mary$_j$ said that [PRO$_i$ shaving himself wasn't important].

        b. John$_i$ expected that Mary$_j$ would thank no one but himself$_i$.

(33)   a. [Alex$_i$'s friends] believe [that PRO$_i$ keeping herself under control at the banquet is important].

        b. John$_i$'s opinion is that [the chair of the department and himself$_i$ should receive accolades].

The puzzle of these construal phenomena is that it is precisely the defining properties of the canonical cases (i.e. those listed in [24]) that do not hold (or must not hold) in the non-canonical—i.e. the non-local cases. How can these be the same phenomena—local and non-local binding, and local and non-local control—when the two subtypes actually exhibit diametrically opposite characteristics? Indeed, in many analyses they are not even given the same treatment, or analyzed as being computed in the same module of the grammar. In the next section we break down the non-local cases into their two distinct sub-types: those where the referentially dependent items lack licensers and those where these elements are licensed by a long-distance antecedent.

The unlicensed type of control is called arbitrary control, and occurs where an infinitival or gerundive predicate's empty external argument position is interpreted as generic and human. For instance in (34) below, (a) is interpreted to mean that '(anyone) smoking is generally bad for health', and in (b) '(a person) having to wake up early is disagreeable', etc. In English examples of arbitrary control, the arbitrary PRO is typically paraphrased as 'one' (e.g. 'one smoking is generally bad for one's health; one's having to wake up early is generally disagreeable,' etc.).

Examples of Arbitrary control:

(34)   a. (PRO$_{arb}$) smoking is generally bad for health.

        b. (PRO$_{arb}$) having to wake up early is disagreeable.

        c. It is advisable (PRO$_{arb}$) to take a multivitamin everyday.

In the domain of binding, we have examples of so-called untriggered anaphors in English and

other languages, such as Icelandic, as illustrated in (35a) from Thráinsson (1991). In (35b),

yourself refers to the addressee of the utterance, which is not (overtly) projected in the tree.[13]

(35) *Unlicensed Anaphors (Icelandic and English)*

    a.   María var alltaf svo andstyggileg.

         María was always so nasty.

         Þegar Ólafurj kæmi segði hún sér$_{i/*j}$ áreiðanlega að fara.

         When Olaf came said she himself certainly to leave

         'Maria was always so nasty. When Olaf would come, she would certainly tell himself

         [the person whose thoughts are being presented-not Olaf] to leave'

         (Thráinsson 1991: 62)

    b. Linguists like yourself tend to be very humorous people.

In the domain of things that are non-locally licensed, we have long-distance control in (36)

and anaphors that are non-locally bound in (37). In (36), *John* controls PRO across the

intervening complement clause containing *Mary*; in (37a) John binds the anaphor *himself* in a

similar configuration. The Icelandic pronoun *sér* is an example of an anaphor-like pronoun

that often appears in configurations where it is bound non-locally (Sigurðsson 1990). In

(37b), for instance, *formaðurinni* ('the chairman') is the binder for *sér*, which occurs several

utterances away in the discourse and thus is not part of the same syntactic structure).

---

[13] Though there have been various proposals since that of Ross (1970), which argue that the speaker
and the addressee are syntactically, but covertly, represented in the syntactic structure of the clause.
See also Miyagawa (2012, 2017), and Ritter and Wiltschko (2018, 2019), among other works. I return
to these issues briefly in Chapter 5.

(36) *Long-Distance Control*

John$_i$ thinks [that Mary said [that PRO$_i$ shaving himself is vital.]]

(Boeckx et al. 2010: 12)

(37) *Long-Distance Anaphors*

a. **John** was hoping [that Mary would support no one but **himself**].

(Reuland 2011:44)

b. **Formaðurinn$_i$** varð   óskaplega reiður.

The-chairman became   furiously angry

Tilligan   væri   svívírðileg.

The-proposal was (subj.)   outrageous

Væri   henni beint gegn   **sér$_i$** persónulega.

Was (subj.) it   aimed against **self**$_i$ personally.

('It was aimed at him personally, he expressed')   (Sigurdsson 1990: 316)

1.4.3 Unlicensed and Long-distance Cases Pattern Together

Landau (2000, 2013, *inter alia*) argues extensively for the classification of both unlicensed

(arbitrary) and long-distance control under the general banner of 'non-obligatory' control: he

makes the case that these two types of control have the same distribution and thus should be

classified (and analyzed) together. In (30) above, we presented the configuration and typical

properties of the non-local cases, which encompass both unlicensed and long-distance control

and binding. Their classification with these very broad criteria (antecedent not necessary,

antecedent not necessarily local, and antecedent not necessarily c-commanding), are more a

description of the phenomena rather than an argument that their distributions overlap. For

more detailed arguments for identifying long-distance anaphora and unlicensed anaphora, the

reader is referred to Reuland (2011, *inter alia*). For control, we will look at the exact

configuration of non-obligatory control in more detail in Chapter 2.

1.4.4 Interim Summary

Figure 1 below summarizes the distributions of the construal constructions outlined above: local versus non-local phenomena, and then the two sub-types of non-local (or non-obligatory, to use Landau's term) phenomena, unlicensed and long-distance varieties of both control and anaphor binding. Both types of non-local phenomena occur in the same syntactic contexts, which we established immediately above. Both non-local control and non-local binding present exactly opposite properties to the local varieties of control and binding, calling into question the status of the local varieties as canonical, and the unity of these phenomena.

The parallels between binding and control in the classification and distribution of both non-local and local varieties should also make us wonder how the two phenomena are related, and whether they might share a similar underlying or causal mechanism in the grammar. Arbitrary control and unlicensed anaphors present another puzzle: how are they licensed, or if they are exempt from licensing, what in the grammar allows this?

In Chapter 4 I will argue that Remote Agree—that is, the forming of a dependency between a probe in the tree and an element in the syntactic Workspace—can be fruitfully applied to these puzzles, and in particular to arbitrary and long-distance control. This will involve some modification of the original Remote Agree proposal by Motut & Oda (2011a,b).

Figure 1: Types of Construal Phenomena

Construal Phenomena

Local, obligatory
- **Obligatory Control**
- **Binding**

Non-local

Unlicensed
- **Arbitrary Control**
- **Unlicensed Anaphors**

Long-distance
- **Long-distance Control**
- **Long-distance Binding**
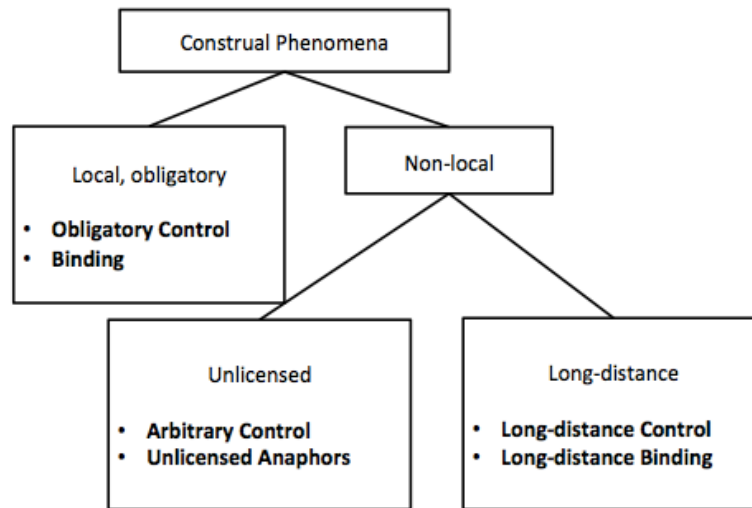
1.5 Remote Agree and Construal Phenomena

On what basis do we look to this Remote Agree operation to solve the empirical puzzles presented here? Why is Remote Agree a promising approach to non-local construal phenomena?

First, both non-local control and non-local binding have been analyzed as the result of Agree dependencies. Landau takes this approach in multiple influential works, diverging from it only with his most recent theory, the Two Tiered Theory of Control (Landau 2015). Reinhart and Reuland (1993) and Reuland (2011) argue for an Agree-based model for anaphor binding in which binding between the antecedent and anaphor is established via a series of successive dependencies formed primarily through Agree-chains (Reinhart and Reuland 1993, Reuland 2011). Hicks (2009), Safir (2004a,b) and others have also proposed Agree-based theories of anaphor binding.

Both of these construal phenomena in their local, obligatory forms can be analyzed as binding via Agree, and both (as we saw in the preceding section) have apparently unlicensed versions (arbitrary control, unlicensed anaphors), theories of which struggle to determine whether and how these referentially dependent elements (PRO, anaphors) are bound. Remote Agree points tantalizingly towards the notion that perhaps these unlicensed construal phenomena *are* bound, just not by an element *in* the syntactic tree.

Both sets of phenomena also have received alternative analyses where the primary mechanism establishing the relation between the antecedent and PRO/anaphor is movement (e.g. Boeckx et al. 2010, Manzini and Roussou 2000, Hornstein 1999, *inter alia*, for control; Boeckx, Hornstein, and Nunes 2010, Kayne 2002, Zwart 2002, *inter alia,* for anaphor binding.) While my proposal is consistent with the movement analysis of control for analyzing obligatory control, I will propose an account of non-obligatory (non-local) control that relies on the Remote Agree mechanism.

In Section 1.3, I discussed some general locality considerations that become relevant if one adopts the Remote Agree hypothesis. In Chapters 3 and 4 (and particularly Section 4.1.1), I will expand on the locality issues that arise and implement a new way of looking at locality in non-obligatory control. I will also touch on anaphor binding, albeit in a more cursory fashion.

Chapter 2: Previous Accounts of Control

2.1 Overview

How does the theoretical proposal of Remote Agree help us approach in a new way the set of puzzles articulated above: the puzzles of projection, apparently unlicensed but licensable elements, and parallelism? Remote Agree provides a new kind of syntactic dependency, and I will demonstrate that it permits a new analysis of arbitrary and long-distance control.

Although my Remote Agree analysis (to be introduced in Chapter 3) will focus on non-obligatory control (NOC), I will first briefly review control in general, including obligatory control (OC), and the main proposals that have been made to account for it. NOC is relatively understudied compared to OC: after briefly reviewing the OC literature, I will review the basic explananda of the control phenomena, focusing on the contrast between OC and NOC, and then review in more depth the configuration of NOC and its previous analyses in the literature.

Control is a phenomenon where a single argument appears to be associated with two theta roles. Configurationally, control in English appears in non-finite clauses that lack an overt subject—that is, infinitivals, gerunds and adjuncts.[14] In the typical cases, the understood subject of the infinitival, which is null, is linked interpretively to another argument, often but not always in a higher clause, from which it receives its reference.

Landau (1999, 2000), in his seminal thesis on the topic, argues for the typology of control given in Figure 2. OC encompasses both exhaustive and partial control, while NOC includes both arbitrary and long-distance control, as mentioned earlier. The difference

---

[14] Finite control, or obligatory control into finite clauses, is cross-linguistically attested, for example in Greek, Romanian, Serbo-Croatian, Persian, Hebrew, among others (Landau 2013: 88). For further details, see Landau (2013), Section 4.1.2.

between exhaustive and partial control is illustrated in (38). In (38a), an example of

exhaustive control, PRO's referent is identical to that of its antecedent. In (38b), an example

of partial control, the referent of PRO includes the antecedent but also whoever else the

director is hoping to meet with. It was Landau (1999, 2000) who first argued that long-

distance control and arbitrary control were both species of non-obligatory control (Figure 2).

Adjunct control, which is not listed in the typology in Figure (2), can be either obligatory or

non-obligatory (Landau 2013, *inter alia*).

Figure (2) (Landau 1999: 12)

```
                          Control
                         /       \
               Obligatory         Non-Obligatory
               /       \           /          \
        Exhaustive   Partial  Long-distance   Arbitrary
```
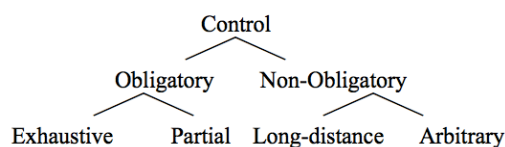
 (38)   a. The director$_i$ hoped [PRO$_i$ to hire four new employees.]

        b. The director hoped [PRO+ to meet virtually.]

        (Adapted from Boeckx et al. 2010: 21)

In the literature, control—where, as noted above, a DP is associated with two theta roles—is

typically contrasted with raising, in which a single argument is associated with two or more

grammatical positions but only one theta role. These two phenomena are syntactically

interesting to contrast because they are superficially similar, as illustrated in (39), but there

are grammatical tests that show them to be different phenomena.[15]

(39)   a. John seemed to enjoy Guatemalan coffee.

        b. John tried to enjoy Guatemalan coffee.

---

[15] See, for example, Perlmutter and Soames (1979: 79–115).

(40) shows the presumed underlying stuctures of each of the sentences in (39), prior to movement.[16]

(40)   a. [seemed [John to enjoy Guatemalan coffee]]

        b. [John tried   [PRO to enjoy Guatemalan coffee]]

The fact that an expletive subject in the matrix clause is grammatical in (41a) but not (41b) is indicative of the fact that in the control construction (41b), there are two theta roles available to be assigned, by both the matrix and embedded verbs, one of which fails to be assigned to the expletive. A host of other tests to distinguish control from raising exist; for a recent overview see Landau (2013: 11–28), Boeckx et al. (2010), among many others.

(41)   a. It seemed John enjoyed Guatemalan coffee.

        b. *It tried John enjoyed Guatemalan coffee.

        Two influential early accounts of control were Equi-NP Deletion and the PRO Theorem. Equi-NP Deletion was a transformation that stipulated that the "controllee" NP would be deleted in exactly the configuration where one finds PRO; later on, a transformation called Super-Equi NP Deletion was introduced to account for non-local control (Rosenbaum 1967, 1970). There was no PRO in this account. These two transformations suffered from empirical shortcomings as well as being largely stipulative/descriptive (Brame 1976; Boeckx et al. 2010: 7-8). In the PRO Theorem, Chomsky (1981) argued that PRO's contradictory status as both pronominal and anaphor made it subject to the conflicting requirements of Principles A and B of the binding theory: it was required to be both bound and free in its binding domain. This contradiction meant the only way to salvage PRO and the binding theory was to stipulate that PRO satisfied these

---

[16] These particular structures are schematic and not entirely consistent with all accounts, but are useful for expository purposes.

conditions vacuously, by not having a binding domain. Simplifying, to have a binding domain meant to have a governor, and so it was stipulated that infinitival Infl was not a governor (while finite Infl was). This approach not only faced empirical disadvantages, but it was also hardly explanatory. It was also not minimalist, in that a separate control module was required specifically for the grammar to interpret PRO (Boeckx et al. 2010: 16). Chomsky and Lasnik (1993) retained PRO and the control module, but abandoned the ambiguous [+anaphor, +pronominal] status of PRO. But their account was hardly less stipulative: non-finite Infl assigned null case, and PRO just happened to be the one and only entity that needed null case. Again, the account lacked explanatory power as it was little more than a technical reformulation of the basic facts of PRO's distribution.

Early as well as recent accounts of control phenomena can also be categorized based on the grammatical mechanism each uses to explain control. Landau (2013: 47–68) provides a summary of these theories organized by grammatical mechanism. I very briefly summarize the main camps here. Predication theories, some but not all of which assume the existence of PRO, take control to be a relation established between a controller and the whole infinitival or gerundive predicate (Williams 1980; Lebeaux 1984; Bach 1979; Chierchia 1984, 1989). Binding approaches to control treat PRO as a null anaphor bound by a higher antecedent within a particular domain, usually the clause containing the infinitival (Manzini 1983, 1986; Koster 1984; Lebeaux 1984, 1985; Borer 1989; Vanden Wyngaerd 1994, Saxon 1989, Sundaresan 2011). As Landau (2013: 55) summarizes, there are binding-control asymmetries that are unaccounted for by these theories, such as implicit arguments' ability to control but not bind (Rizzi 1986); these accounts also have very little to say about the inherent nullness of PRO (although see Sundaresan 2011). Finally, substantial debate has raged between the

Agree and movement camps of analysis for OC. Below, I summarize each of these in somewhat more depth.

In the Agree approach to OC, PRO is a null pronominal element whose relationship to its antecedent is mediated indirectly via Agree relationships with local functional heads. Landau has the most substantial and widely adopted version of an Agree-based analysis of OC, so I focus on his account here. Relying on the Agree operation of Chomsky (2000, 2001), a probing functional head (i.e. *v* or T) establishes a relation with a DP (the controller / antecedent to PRO) in the clause containing the control clause, as well as with PRO (or Agr) in the infinitival (Landau 2000, 2004, 2006, 2007, 2008). This functional head thus acts as the link between controller and PRO/control clause, resulting in a bound variable interpretation. Landau further argues for an elaborate featural system for both I and C heads ([+/− Agr], [+/− T]), and for DPs and lexical elements to implement the cross-linguistic distribution of control. These classifications result in a system where "DP/*pro* environments are [a] natural class [(+T, +Agr) environments], and PRO occurs in the elsewhere case: [+T, -Agr], [-Agr, +T] or [-T, -Agr]," thus "revers[ing] the traditional view, by which PRO-environments form a natural class defined by a single syntactic criterion (e.g. ungoverned, null case, no case, etc.)" (Landau 2013: 67).

Landau's Agree approach to control has the benefit of explaining both partial control and finite control. Partial control was illustrated above in (38b). Finite control occurs where the embedded verb has tense and agreement features, in contrast to English, and is attested in numerous languages, for instance Romanian, Greek, Bulgarian, Serbo-Croatian, Persian, Hebrew, and Dogrib, to name only a few (Landau 2013: 88).

A disadvantage of the Agree theory of OC is that it cannot shed light on the phenomena of backward control, copy control, or split control. Adjunct control is also not

captured by this analysis, despite the fact that it often displays the hallmarks of OC (Landau 2013: 68). Additionally, as Landau (2015) observes regarding his earlier work, the reassignment rule, which specifies that lexical DPs/*pro* appear in just those contexts that are [+T, +Agr] and PRO elsewhere, is stipulative. Furthermore, the features themselves do not always transparently overlap with the expected morphology (Landau 2013: 68). Another problem is that there is seemingly nothing to trigger Agree in OC: as Landau (2015: 13) notes, "[the trigger] cannot be the matrix functional head, for all the uninterpretable features of that head are already checked against the DP controller by the time the complement clause is probed." Wurmbrand (2014b) and Grano (2012) further noted problems with Landau's use of the notion "semantic tense" in the Agree model.

Finally, two large questions remain unanswered in Landau's Agree analysis: the existence of PRO and the lack of connection between NOC and OC. Landau (2015: 14) acknowledges that it would be preferable if PRO could be eliminated in favour of a more general, independently motivated null category. The lack of unity between the analyses of OC and NOC is an issue we will return to. What the Agree model, like many accounts of OC, lacks is a way of unifying the two phenomena. The Agree model assumes that NOC PRO is a logophor determined in the discourse, while OC is a syntactic phenomenon. Landau (2015: 14) muses: "while the empirical differences [between NOC and OC] are undeniable, it is legitimate to ask whether they run as deep as an ontological duality of PRO rather than reflecting higher-level grammatical distinctions that nonetheless hide a unitary structure."

Many, if not all, of the problems with the Agree account that are described above are laid out in detail in Landau (2015: 12-15), which presents a novel approach to control, both OC and NOC, called the Two-Tiered Theory of Control (TTC). The TTC re-draws the landscape of OC and NOC, arguing for a contrast between logophoric control (appearing in

attitude complements) and predicative control (appearing in non-attitude complements) which cuts across the NOC/OC distinction. OC in non-attitude complements is established via predication (predicative control) while OC in attitude complements is established via variable binding (logophoric control). In logophoric control, in contrast to predicative control, a second tier of structure is projected above the predicative layer: logophoric-control complements are propositional while predicative-control complements are property-denoting. In this theory, PRO is "radically impoverished […] a 'reference variable' in the sense of Sigurðsson (2008) or a minimal pronoun in the sense of Kratzer (2009)" (Landau 2015: 23). This account more closely unifies NOC and OC, while also re-analyzing PRO as an element (with lexical entry [D, $u\varphi$]) that is less mystifyingly unique in the grammar.

Finally, there are the movement accounts of control. Hornstein (1999, 2001, 2003, *inter alia*) develops the most widely known and adopted movement account of control. I describe his account, the Movement Theory of Control (MTC), here, although there are other accounts.[17] The MTC argues that control and raising are both types of A-movement, collapsing the distinction between control and raising described as being diagnostic at the beginning of this section. The following description summarizes the MTC based on Hornstein (1999) and Boeckx et al. (2010).

The MTC starts with a theoretical motivation. Early generative grammar made a distinction between Deep Stucture and Surface Structure in a derivation. Deep Structure, prior to movement, was where theta roles were assigned. In Minimalist theory, the distinction

[17] Those who have based analyses on Hornstein's movement account or developed their own movement theories of control include Bowers (1973, 1981, 2008), Manzini and Roussou (2000), Polinsky and Potsdam (2002), Boeckx and Hornstein (2004, 2006), Alboiu (2007), Potsdam (2009), and Boeckx, Hornstein and Nunes (2010).

between Deep Structure (DS) and Surface Structure has been eliminated, freeing up the possibility that theta-roles may be assigned interspersed with movement operations. Hornstein argues that the uniqueness condition on the early Theta Criterion—that each argument can only be assigned one theta role, and each theta role can only be assigned to one argument—is dubious, and argues that arguments can acquire theta roles during the course of a derivation. Control, as stated earlier, is a phenomemon where a single argument appears to be associated with two theta roles, or two argument positions. Inserting a PRO into the (usually) lower position, and linking it with the higher argument DP allows this one-to-one relation between theta-roles and arguments to be maintained.

But if we abandon the Theta Criterion, as well as a distinct level of Deep Structure, and allow movement into theta positions, nothing prevents control from being analyzed as an A-movement chain. (42) shows what such a derivation would look like for a simple OC sentence (example adapted from Boeckx et al. 2010: 44).

(42) Derivation of *John tried to hug Mary*.
   a. Merger of *hug* and *Mary* + θ-assignment → [hug Mary]
   b. Merger of T → [T [hug Mary]]
   c. Merger of *John* + θ-assignment → [John [T [hug Mary]]]
   d. Merger of C → [C [John [T [hug Mary]]]]
   e. Merger of *tried* + θ-assignment → [tried [C [John [T [hug Mary]]]]]
   f. Merger of T → [T [tried [C [John [T [hug Mary]]]]]]
   g. Movement of *John* + θ-assignment → [**John$_i$** [T [tried [C [**t$_i$** [T [hug Mary]]]]]]]

The MTC has some clear advantages. It eliminates PRO and the control module from the grammar in favour of an independently motivated operation—movement—and an an independently motivated element–A-traces. It is also one of the few control theories to adequately handle backward control and so-called copy control, both phenomena attested

cross-linguistically. In backward control, a DP is associated with two theta positions but, unlike in forward control, "in overt syntax, the controller appears in a lower position than the controllee" (Boeckx et al. 2010). (43) illustrates the general schematic of forward vs. backward control, where Δ is a representation meant to indicate neutrality on the question of whether the controller is PRO or a trace.

(43)   a. Forward control: $[DP_1 \ V \ [\Delta_1 \ …]]$
        b. Backward control: $[\Delta_1 \ V \ [DP_1…]]$ (Adapted from Boeckx et al. 2010: 102)

Many properties of OC fall out from the MTC. The nullness of PRO follows from the nullness of traces. Antecedent-controller relations must obey c-command and locality because A-movement obeys c-command and locality. Adjunct-control is handled via Nunes' (2004) "Sideward Movement." However, the MTC also has non-trivial problems.

The MTC does not predict the possibility of split and partial control, given that in the MTC, the head of the A-chain forming the control relation must be identical to its trace. Landau (2003, 2007, 2013: 64) and others also note several cases where the embedded subject's movement to the matrix subject position would be allowed by the MTC but the result is ungrammatical. (44) is one such example.

(44)   *John$_i$ was hated [t$_i$ to live like that].

Landau also points out several cases where the MTC would misclassify an instance of OC as NOC or vice versa, and argues that many robust raising–control contrasts are lost in the MTC analysis (Landau 2003, 2007; Bobaljik and Landau 2009). Hornstein (2001) and Boeckx et al. (2010) also present an analysis of NOC that relies on inserting little *pro* as a repair strategy in contexts where movement is not available.

The analysis I will present for NOC using the Remote Agree mechanism is consistent with the MTC, although it may also be consistent with other proposals for OC in the

literature. Since my focus is primarily on explaining the parallelisms and puzzles of projection for arbitrary and long-distance control, I will have little more to say about OC.

2.2 Defining an Adequate Theory of Control[18]

Boeckx et al. (2010: 5–6) claim that any adequate theory of control must explain the various sub-types of control and their differences; account for the configurational properties, including the available positions for controller and controllee; explain the interpretive properties of control relations, including what determines the antecedent of the controllee, and what relation holds between the two. Finally, Boeckx et al. claim that an adequate theory of control must also account for the nature of the controllee itself—what is it, in the inventory of syntactic categories? In other words, what sort of a thing is PRO, if PRO exists at all? And why?

To this, I would add that an adequate theory must explain why the controllee is always null, whereas the controller is sometimes null or absent altogether (as in implicit and arbitrary control) and sometimes overt. An adequate account must also explain why the controller in non-obligatory control is always interpreted as [+human]. And an adequate theory of control must also explain the complementary distribution of obligatory control and non-obligatory control, and the parallelism between long-distance and arbitrary control, among the "puzzles" referred to earlier. Next I look in more detail at the distributions and properties of OC and NOC.

---

[18] I echo here the beginning subsection of Boeckx et al. (2010: 5), "What any theory of control should account for."

2.3 Explananda: Properties of NOC versus OC

I will shortly turn to applying the Remote Agree mechanism to NOC. First, though, I will describe in more detail the syntactic and semantic properties of the phenomenon. An account of NOC must necessarily address the distributional patterns of NOC with respect to OC, its mirror image. In this section, I thus summarize the distributions of OC compared to NOC, and give a more detailed description of the syntactic and semantic properties of NOC. I then turn to a review specifically of previous accounts of NOC.[19]

One of the most striking things about the OC / NOC syntactic and semantic distributions is their complementarity.[20] The MTC predicts complementarity between NOC and OC: indeed, complementarity is crucial to their account. NOC surfaces only where OC, i.e., movement, is not available. Likewise, the "signatures" that Landau (1999, 2013) describes for NOC and OC present a picture of mirror opposite phenomena. Compared to OC, NOC is the elsewhere case (Boeckx et al. 2010: 195).

---

[19] I draw on the content and organization of Landau (2013: 230–237) throughout this and the following section.

[20] I gloss over a detailed description of the contrast between OC and NC (=non-control) contexts. As Landau (2013: 88) notes, "internal factors (finiteness ingredients) determine whether the subject is PRO or DP/*pro*; second, for the PRO cases, configurational factors […] determine whether PRO exhibits NOC or OC." I am mostly interested in the configurational differences between NOC and OC, and so won't provide a full description of the factors conditioning OC (versus NC). For completeness, (i) shows Landau (2013: 90)'s definition of the OC context and its "finiteness" ingredients.

(i)   In a fully specified complement clause (i.e. the I^0 head carries slots for both [T] and [Agr]),
        a. If I^0 carries both semantic tense and agreement ([+T, +Agr]), NC [no control] obtains.
        b. Elswhere, OC obtains.   (Landau, 2013: 90)

Landau's OC signature is given in (45):

(45)   OC Signature: (Landau 2013: 232)

In a control construction, [X….. [s PRO ….]], where X controls the PRO subject of the clause, S:

    a. The controller X must be a co-dependent of S

    b. PRO (or part of it) must be interpreted as a bound variable.

In OC, controllers of PRO are co-dependents of the controlled clause, and PRO is interpreted as a bound variable. Landau presumably uses "X" rather than "DP" in this definition to allow for implicit control, where the controller might be an implicit argument. A "co-dependent" is either an argument or an adjunct of S. The definition also encompasses both partial and split control ("….part of it…" in [45b]).

The NOC signature is given in (46).

(46)   NOC Signature: (Landau 2013: 232)

In a control construction, [….. [s PRO ….]]

    a. The controller need not be a grammatical element or a co-dependent of S.

    b. PRO need not be interpreted as a bound variable (it may be a free variable).

    c. PRO is [+human].

In contrast to the OC Signature, the NOC Signature seems to dispense with the controller entirely: the controller *need not be a grammatical element*, whereas for OC it must be.[21] Since the controller need not be a grammatical element, this allows for arbitrary control and discourse control. Also contained in the definition is that the controller in NOC need not be a co-argument or co-dependent of S. This allows the definition to encompass long-distance control.

The contrasts between OC and NOC are further demonstrated in the data below, from Boeckx et al. 2010: 196 (these data and main points are originally from Hornstein 1999: 73–

---

[21] This is consistent with Landau's assertion that NOC is not a grammatical phenomenon.

74). (47a), (48b), and (49c) show that OC must have an antecedent (47a), that the antecedent cannot be long-distance (48a), and that the antecedent must c-command the controlled clause (49a). In contrast, we see that NOC has none of these requirements: it does not need an antecedent (47b), can take a long-distance antecedent (48a), or an antecedent that doesn't c-command the controlled clause (49b). In (50) – (52) we see some of the semantic properties of OC versus NOC. (50a) shows that there is only a sloppy reading under ellipsis for OC; whereas for NOC (50b) both strict and sloppy readings are available. (51) shows that OC allows only a bound variable reading while NOC is freer; (52) demonstrates that OC allows only a *de se* reading while NOC allows for both *de se* and *de re* readings.

(47)   a. *It was expected PRO to shave himself
       b. It is illegal to park here.

(48)   a. *John thinks that it was expected PRO to shave himself.
       b John thinks that Mary said that PRO shaving himelf is vital.

(49)   a. *John's campaign expects PRO to shave himself.
       b. John's friends believe that PRO keeping himself under control is vital.

(50)   a. John expects PRO to win and Bill does too
          (Sloppy reading only: Bill expects Bill to win, not John.)
       b. John thinks that getting his resume in order is crucial, and Bill does too.
          (Strict or sloppy reading available: Bill thinks getting John's or Bill's resume in order is crucial.)

(51)   a. [Only Churchill] remembers PRO giving the speech.
       b. [Only Churchill] remembers that PRO giving the speech was momentous.

(52)   a. [The unfortunate] expects PRO to get a medal.
       b [The unfortunate] believes that PRO getting a medal is unlikely.

As we can see, OC and NOC contrast in almost every respect. They also occur in different syntactic environments. The basic syntactic distribution of each can be stated as in (53) below:

(53)   Generalization of the Distribution of OC vs. NOC (Landau 2013: 231)
        a. complement clauses fall under OC;
        b. subject and adjoined (extraposed) clauses fall under NOC.

Landau (2013: 38) gives the following minimal triplet data to illustrate OC in complement clauses (54a) versus NOC in subject and extraposed clauses (54b-c). In (54a), the PRO clause is a complement of the matrix verb, and local control of PRO (by *Mary*) is obligatory (long-distance control by *John* or *we* is disallowed). In (54b), we have a controlled subject clause displaying NOC (i.e. control by either *we* or *Mary*). In (54c), the extraposed clause is controlled either non-locally by *We* or more locally, but still long-distance, by *Mary*.[22]

 (54)   a. We$_i$ thought that John would help Mary$_j$ [PRO$_{j/i*}$ to expose herself/*ourselves]
         b. We$_i$ thought that [[PRO$_{j/i}$ to expose herself/ourselves]] would help Mary$_j$
         c. We$_i$ thought that it would help Mary$_j$ [[PRO$_{j/i}$ to expose herself/ourselves]]

This is the syntactic configuration in which each of NOC and OC occurs, as it pertains to *arguments*. Adjuncts, as Landau (2013: 231) notes, split along similar lines: some act like NOC and some like OC. Landau (2013: 231) notes that event-related adjuncts (those adjoined inside VP) consistently display OC. NOC, on the other hand, is displayed by temporal and absolutive TP-adjoined adjuncts, particularly those that occur sentence-initially.[23] Adjunct clauses like temporal gerunds (55), result clauses (56), and goal clauses

---

[22] (54c) is an extraposition construction, dealt with extensively in the early search-based accounts of control—we examine this construction in more detail in Section 2.4.

[23] Landau (2001: 231) notes that those that occur sentence-finally are more likely to be interpreted as NOC if they are set off with a pause.

(57) display OC (Landau 2013: 221). On the other hand, NOC is sometimes displayed in adjuncts such as initial (fronted) temporal adjuncts (58), final temporal adjuncts (59), and final "without" clause (60). (Examples adapted from Landau 2013: 221, 231–232.)

(55) John tripped [PRO while playing baseball / when running upstairs]

(56) John grew up [PRO to be a firefighter]

(57) John picked up the kids from school [PRO to help his sister]

(58) After [PRO pitching the tents], darkness quickly came.

(59) Lasagna is better [PRO after reheating it]

(60) There is no hope for reconciliation [PRO without accepting the atrocities that have occurred.]

These are the basic syntactic configurations in which we find NOC versus OC. As for the semantics, as the OC Signature in (45) notes, in OC but not in NOC, the controlled clause (or PRO) must be interpreted as a bound variable. In (50a) above, we can see that in ellipsis constructions, for OC only the sloppy reading, where PRO is bound by *John*, is possible . For NOC, as shown in (50b), either the strict or sloppy reading is available. There is also a *de re / de se* distinction between the interpretation of NOC and that of OC (52). With OC, as in (52a), only the *de se* interpretation is available.

One surprising semantic property of NOC that does not appear to be simply the opposite of what is found with OC is that the unexpressed subject must be interpreted as [+human] in NOC. Landau (2013: 234–36) summarizes arguments from Chomsky (1981), Kawasaki (1993), Williams (1992) and others that demonstrate that while OC PRO can sometimes be interpreted as non-human, NOC and arbitrary control are never interpreted as non-human. This contrast is illustrated in the OC examples in (61), versus the NOC examples in (62). In (61a–b), the controller is inanimate and non-human (*your alibi, the car*), and can

control PRO. (61c-d) are OC adjuncts, and PRO can likewise be interpreted as non-human in these examples.

(61)   a. Your alibi$_i$ fails [PRO$_i$ to convince me]

       b. Suddenly, she forced [the car]$_i$ [PRO$_i$ to stop]

       c. This book$_i$ was out of print [before PRO$_i$ becoming a bestseller last summer.]

       d. Granola snacks$_i$ can increase your energy level [without PRO$_i$ increasing your

       blood pressure.]   (Landau 2013: 234–235)

(62) and (63), by contrast, show examples of NOC and arbitrary control, where it is

impossible to get an arbitrary or generic non-human reading for the controlled clause.

 (62)   a. *It is possible [PRO$_{arb}$ to roll down the hill].

             cf. It is possible for the rocks to roll down the hill.

       b. *[PRO$_{arb}$ to snow all day] would be a nuisance.

             cf. For it to snow all day would be a nuisance.

       c.*[PRO to be clear [that we're out of fuel]] would be a nuisance.

             cf. [For it to be clear [that we're out of fuel]] would be a nuisance.

                              (Chomsky 1981: 324–327; cited in Landau 2013: 235)

As Landau (2013: 235) says about the examples in (62a-c), "uncontrolled PRO cannot be

interpreted as [–human], pleonastic *it* of weather predicates or the pure expletive *it* of

extraposition." As the examples in (63) further illustrate, for NOC in (63a) and for arbitrary

control in (63b-d), that *even when* the context or pragmatics of the discourse would favour an

inanimate or non-human referent (i.e. *food* in (63b), or *the window* in (63a), what is instead

preferred is a human referent, even if this makes the sentence odd (Kawasaki 1993).

(63)   a. The open window$_i$ proves that [before PRO$_i$ breaking], it was raining.

                              (Williams 1992, cited in Landau 2013: 236)

       b. [After PRO$_{arb}$ being spoiled in a refrigerator], there is nothing even a

          good cook can do.

       c. The government abolished [PRO$_{arb}$ having to be surrounded by fences].

d. I read stories about [PRO_{arb} falling off a cliff].

(Examples [63b-c] from Kawasaki 1993: 30, cited in Landau 2013: 236)

Based on these examples, we can see that NOC is inherently [+human], whereas OC is not.

NOC, when it is arbitrary control, is also interpreted as generic, and cannot receive a definite

interpretation. The referent of the subject of the infinitival clause in (64) below is generic

(e.g. refers to any person).

(64)   [PRO_{arb} To have a tooth pulled] is no fun.

Finally, note that sometimes we see examples that would appear to challenge the claim made

above, of complementarity between OC and NOC. Some examples appear to be ambiguous

between OC and NOC, or to show both readings. In each of the cases where this appears to

be true, in fact one or both of two situations hold: either (a) there is optionality between an

NOC long-distance reading and an NOC arbitrary reading; or (b), there is structural

ambiguity, or both.

(65)   a.   [PRO_i/PRO_{arb} to eat the last piece of pie] helps John_i
        b.   It would help John_i [PRO_i/PRO_{arb} to eat the last piece of pie]

In (65a), the PRO is contained in a subject clause; therefore, this is an example of NOC

(alternatively, this can be interpreted as arbitrary control). In (65b), there is a reading

available where *John* and PRO are co-indexed, in addition to the arbitrary reading where

anyone is eating the pie. However, this is not obligatory control. The antecedent (*John*) is not

local, and it is not interpreted as binding PRO. This is instead an example of long-distance

control: in fact, the availability of the arbitrary reading is one criterion that establishes this.

I now turn to a review of some of the previous accounts of non-obligatory control.

2.4 Previous Accounts of NOC

Compared to OC, much less has been written about NOC in the literature. A key distinction in the NOC accounts is whether they try to account for both OC and NOC under the same mechanism, or whether they instead account for them using separate mechanisms or even separate modules of the grammar. Landau (2000, 2013), Hornstein (2001, 2003, 2007), and Boeckx et al. (2010) are in the latter camp. Interestingly, it is some of the very early accounts of control (e.g., Grinder 1970) that do try to unify NOC and OC with a syntactic analysis, appealing to locality to explain where local and where long-distance antecedents can control PRO. Landau's (2015) Two-Tiered Theory of Control, which I return to at the end of the thesis, also provides an account unifying NOC and OC, improving on his earlier Agree accounts. I first summarize Hornstein's little *pro* account and its challenges, then turn to Grinder (1970) in more detail. Landau's (2013) and Kuno's (1975) logophoric approaches are discussed in later sections.

2.4.1 Little *pro* "Economy / Last-Resort" Account

Boeckx and Hornstein (2007) and Hornstein (1999, 2001, 2003, 2007) argue that since NOC obtains in island contexts (i.e., precisely those contexts where movement is not possible), NOC can be analyzed as the insertion of a null pronoun, *pro,* by the parser as a kind of last-resort strategy. Boeckx and Hornstein argue that it is the parser rather than the syntax which intervenes to save the doomed derivation, and that "pronominalization is more economical than movement" (Boeckx et al 2010: 197), thus prompting the parser to fill in the gap with *pro* rather than a trace. Landau (2013: 242) refers to their analysis as "a locality account of NOC on the basis of the movement theory of control […] coupled with certain

parsing strategies." (The analysis has parallels to analyses of resumptive pronoun insertion in contexts where movement fails.)

To support this analysis, they draw parallels between NOC and the contexts in which pronouns occur (Boeckx et al 2010: 197). Neither overt pronouns nor NOC require an antecedent; if they do have antecedents, they do not need to be c-commanding or local. Both strict and sloppy readings are available with ellipsis; and there are several other diagnostic similarities.

The contrast between OC in the Movement Theory of Control and NOC in the *pro*-insertion account are illustrated in (66) below. In (66a), obligatory control, the DP *the unfortunate* has moved from the embedded infinitival clause, leaving a trace (this is the MTC). The corresponding NOC sentence in (66b) shows how *pro* appears (inserted as a last resort by the parser) in an island (NOC) context where such movement is not available.

(66)   a. [[The unfortunate]$_i$ expects [**trace$_i$** to get a medal]]
        b. [[The unfortunate]$_i$ believes that [**pro$_i$** getting a medal is unlikely]]

                                                        (Boeckx et al. 2010: 197)

The Movement Theory of Control, combined with Boeckx and Hornstein's account of NOC as little *pro* predicts strict complementarity betweeen OC (movement) and NOC, (insertion of *pro* as last resort). The relation between the two is mediated by an economy condition that applies where pronominalization and movement are in competition.

Yet, perhaps anticipating defeat, Boeckx et al. (2010: 195) do not seem particularly committed to this account of NOC, going so far as to state in the introduction to their analysis in Chapter 6 of *Control as Movement* that "should [our account] of NOC turn out to be partially or totally incorrect, this does not affect the essence of [the account of OC given in]

the previous chapters." In other words, the Movement Theory of Control stands regardless of the success of the *pro*-insertion analysis of NOC.

Their NOC account indeed faces many empirical challenges, summarized by Landau (2013: 242-243). In particular, constraining sideward movement so that it does not overgenerate turns out to be challenging. In addition, Landau points out that "if trace is not an option (i.e., sideward movement fails), the gap (which is *pro*) could take any antecedent that a pronoun could take." This would predict that NOC PRO (*pro* in their account) would have the same distribution as an overt pronoun. Examples like (67) are presented as evidence that NOC PRO does *not* have the same distribution as an overt pronoun: the DP antecedent *Max* can be co-indexed with pronoun *him* but not with PRO, which is argued to be in the same structural position. I will return to this example, and others like it, in Chapter 4, suggesting a different interpretation of the data.

(67)   [The girl who Max_i loved] said that it would be difficult [for him_i/*PRO_i to excuse himself from the party before midnight.]

                                          (Landau 2013: 243, from Grinder 1970: 306)

2.4.2 Apparent Intervention Effects: Early structural accounts of NOC

Many of the early structural or syntactic accounts of control were essentially locality stories: the choice of controller, often including the choice between NOC and OC, was determined by minimal distance or minimal search considerations. Or, it was framed in terms of intervention: available controllers closer to the controlled clause could intervene and prevent control by DPs that were farther away from the controlled clause. The problem was often framed as determining when "local" control would block "long-distance control", or a less local controller (where locality was defined in terms of c-command or command relations in the tree). Accounts that argue for an explanation of NOC and/or OC based on the notion of a

"search domain" include Grinder (1970), Lebeaux (1984), Chierchia and Jacobson (1986), Huang (1989), Sag and Pollard (1991), and Manzini and Roussou (2000) (Landau 2013: 240, fn. 10).

The first structural account of OC and NOC identified in Landau's (2013) historical sketch of the NOC literature is that of Grinder (1970). Since Grinder will be a starting point for discussion in Section 4.2, and is a good example of a purely structural account, I will introduce his intervention constraint here in some detail. The generalization itself was empirically inadequate, but much of the data will be relevant to our later discussion.

Grinder's intervention constraint, updated by Landau (2013) with more recent terminology, is shown in (68). (Grinder's original formulation was made in the terms of Equi-NP deletion and Super-Equi NP Deletion.)

(68)   *Grinder's Intervention Constraint* (as presented in Landau 2013: 238)
       C may control PRO iff C commands PRO and there is no intervener B. B is an
       intervener that blocks control of PRO by C iff B commands PRO, and:
       a.   B precedes PRO and C follows it (… B….PRO…..C)
       b.   OR, both B and C precede / follow PRO, B is closer to PRO than C, and B,C are
            not clausemates (….C [$_S$… B…. PRO …./ ….PRO…. B [$_S$….C…..].

Consider example (67), adapted in (69):

(69)   [[The girl]$_i$ who Max$_j$ loved] said that it would be difficult [PRO$_{i/*j}$ to excuse
       herself$_i$/*himself$_j$ from the party before midnight.]

                                             (Landau 2013: 238, from Grinder 1970: 306)

In this example, the DP *Max* is linearly closer to PRO than *the girl*, but since it does not command into the control clause, it does not block DP *the girl* from controlling PRO, and the reading with *himself* is unavailable. In other cases we observe that intervention seems to occur by a closer DP, where the DPs are on opposite sides of the control clause:

(70) *[That it disturbed Pete [PRO$_j$ to wash herself$_j$]] surprised Eileen$_j$.

     (Landau 2013: 239, originally Grinder 1970: 305)

Pete intervenes here to block *Eileen* from controlling PRO, because although both DPs command the control clause, Pete precedes it. This seems to confim Grinder's original intuition—that long-distance (non-obligatory) NOC could be blocked by a more local controller, forcing OC or local control. (As to why *Pete* is more local here, it has to do with the formulation of Grinder's constraint, which states that where both DPs c-command the control clause and are thus potential controllers, the DP that precedes the control clause will intervene.)

     Grinder (1970: 302) also gives the following example of the intervention constraint applying where both DPs, the actual controller and the potential intervener, are on the same side of the controlled clause:

(71) *John$_i$ said [that it disturbed Sue [PRO$_i$ to make a fool of himself$_i$ in public.]][24]

Grinder claims that the DP *Sue* serves as an intervener, blocking *John* from controlling PRO. Both precede the controlled clause and Sue is closer to PRO, therefore intervening and blocking *John* from controlling.

     Some other examples of Grinder's constraint in action are presented below. In (72), both DPs *Eric* and *Roxanne* c-command and precede PRO; in (73), both DPs *Pete* and *Eileen* c-command and follow PRO. In both cases the DP closest to the controlled clause intervenes and blocks control by the more distant DP. In (74), intervention does not occur since *Tom* and *Harriet* are clausemates. (Examples from Landau 2013: 239 and Grinder 1970.)

(72) *Eric said that Roxanne knew that it would be difficult [PRO to criticize himself].

---

[24] The judgement for this sentence is as reported in Landau (2013) and Grinder (1970); however, this is challenged by Motut & Grant (2017), which will be discussed in Section 2.4.4.

(73)   *[That [PRO washing herself with liquid oxygen disturbed Pete]] surprised Eileen.

(74)   Tom told Harriet that it would be tough [PRO to prevent himself / herself from crying at the funeral]

Grinder's constraint has been challenged on empirical grounds. In particular, his constraint was too restrictive. Judgements about potential intervenors in cases of long-distance control turned out to be much more variable than he predicted. Indeed, Super-Equi extraposition examples like (71), in particular, would come to be analyzed very differently by Landau (2001), and Landau's analysis would in turn be challenged by Motut & Grant (2017).

Landau would in fact predict (71) to be ungrammatical, like Grinder. But he did note that even in these extraposition examples, where one would expect intervention (according to Grinder), there are instances where long-distance control may "ignore" an intervenor, such as (75).

(75)   John$_i$ said [that it helped Sue [PRO$_i$ to make a fool of himself$_i$ in public.]]

The only difference between (71) and (75) is the matrix predicate and thematic status of the argument it introduces: *Sue* is an experiencer in (71) but not in (75). For Landau, this distinction—whether the matrix predicate was psychological or non-psychological, and therefore whether or not it introduces an experiencer argument—was crucial. I turn in the next section to a more detailed review of Landau (1999, 2001) and extraposition, as well as introducing Motut & Grant's (2017) refutation of the key grammatical distinction in that work.

Despite the fact that early structural accounts like Grinder (1970) encountered numerous empirical hurdles, they nevertheless capture a pleasing intuition of the relation between OC and NOC: that the choice of a local versus long-distance controller is in some sense determined by locality, and is at least in part about Search. Landau (2013) summarizes

several of the reasons why these old "structural" accounts failed, reasons which were largely empirical. We will return to these analyses in Sections 2.4.3 and 4.2. In some examples, we will question the judgements themselves; in others, we will see how a new analysis of NOC relying on the Remote Agree mechanism can resurrect, to a certain degree, a kind of "search" flavour to distinguishing NOC from OC, with a revised definition of search space that includes the syntactic Workspace.

First, however, I will briefly take a detour through Landau (2001) and Motut & Grant (2017).

2.4.3 Experiencers in Super-Equi Extraposition: To intervene or not to intervene?

In the NOC literature, particularly the early literature, Super-Equi constructions have received considerable attention. Grinder (1970) and others (Williams 1980, Lebeaux 1984, Manzini and Roussou 2000) predicted that Super-Equi extraposition constructions, as in (71) above and (76), were ungrammatical, but when the control clause was in the subject position of the embedded clause (i.e. "intraposition" in [77]), the construction was predicted to be grammatical. In many early Search-based accounts, the choice of a local versus non-local controller in Super-Equi constructions was hypothesized to be determined by locality: where a local controller was not available, a long-distance one could control, and more local controllers could intervene.[25]

---

[25] The terminology here requires comment. Following many sources in the literature, including Landau (2001), in this section I refer to control by the matrix internal argument, i.e. *John* in (76a-b), as "local control" and control by the matrix external arguemnt, i.e. *Mary* in (76a-b), as "long-distance control." The use of these terms is only meant to indicate relative distance of these respective controllers and not to commit to an analytical stance. In fact, control by both of these arguments is non-obligatory control (and therefore "long-distance"), although Landau (2001) argues that in cases like (76a), local obligatory control is forced by matrix experiencer *John*.

Landau (2001) was the first to thoroughly question the generalization that long-distance control in extraposition constructions like (71) and (76) was *always* ungrammatical, as Grinder and others had predicted. He noted a sensitivity to whether or not the matrix predicate was a pscyhological predicate, and claimed that only experiencers, introduced by a matrix psychological predicate, would intervene and block long-distance control. Non-experiencer arguments were predicted not to be intervenors in this construction. (76) – (77) give Landau's (2001) paradigm, with his judgements.[26]

(76)   Super-Equi Paradigm (Landau 2001) — Extraposition
        a. *Extraposition, psychological predicate ('disturb'):*
           Mary believed that it disturbed John [PRO to speak *her/his mind]
        b. *Extraposition, non-psychological predicate ('help'):*
           Mary believed that it helped John [PRO to speak her/his mind]

The experiencer blocking effect is argued by Landau (2001) to hold only in extraposition constructions and not intraposition, so in (76) but not (77). In (76a), *John,* an experiencer, blocks long-distance control by *Mary*. In (76b), *John* is a theme rather than an experiencer, and does not block long-distance control by *Mary*. Landau argues that this grammaticality contrast is neutralized in the corresponding intraposition constructions, shown in (77).

(77)   Super-Equi Paradigm (Landau 2001) — Intraposition
        a. *Intraposition, psychological predicate ('disturb'):*
           Mary believed that [PRO to speak her/his mind] disturbed John.
        b. *Intraposition, non-psychological predicate ('help'):*
           Mary believed that [PRO to speak her/his mind] helped John.

   Here I give a brief summary of Landau (2001)'s account of the paradigm, assuming for now the judgements as given in that work. Landau makes several structural assumptions. He

---

[26] I will challenge these judgements in the immediate next section based on Motut and Grant (2017).

assumes that extraposition is movement and adjunction to the VP, and that VP-internal

clauses must be peripheral at PF. He assumes that any link in an A-bar chain may be visible

at LF, and that LF can therefore choose to interpret the thematic position of the control/PRO

clause or the VP-peripheral one. He further assumes that experiencers are generated above

causer arguments, which are generated above themes and goals (Belletti and Rizzi 1988). He

argues that when the controlled clause is VP-internal, OC obtains, and when it is extraposed,

NOC obtains.[27]

The contrast between (76a) and (76b) is derived as follows. Since the experiencer

argument is generated higher in the VP structure in its thematic position in (76a), the

controlled clause (causer argument) is already VP-peripheral and does not undergo

extraposition movement. Extraposition movement of the controlled clause is what creates the

context for NOC and therefore the availability of a long-distance controller (i.e. Mary).

Without this movement, only local (i.e. obligatory, in this analysis) control by *John* is

possible. In (76b), on the other hand, *John* is a theme and is generated in a different thematic

position within the VP, relative to the position of the causer control clause argument. Since

John is generated as a sister to the verb (in Landau's analysis, this was the position occupied

by the control clause argument in (76a)), the control clause (causer) is generated in [Spec,

VP], which is not a VP-peripheral position. Thus, this clause undergoes extraposition

movement, adjoining to the right of the VP, where it is in an NOC configuration and is

available to be controlled by *Mary*. This creates a chain where either link is available to be

interpreted at LF: if the base-position copy of the control clause is interpreted, we get local

---

[27] Indeed, these are the configurations we expect.

(obligatory) control by *John*, and if the adjoined / extraposed copy is interpreted, we get long-distance control by *Mary*.

In (77a-b), the intraposition configuration is created when the infinitive moves to [Spec, IP] to check the EPP feature of Infl. This movement occurs regardless of the thematic position of the infinitive (i.e., whether it is generated in [Spec,VP] or as V's complement), therefore removing the context for local, obligatory control, which in these constructions only arises if the infinitive remains within the VP. (No reconstruction is expected here since this is an instance of A-movement, Landau [2001: 125] claims, following Chomsky [1995].) Thus both (77a) and (77b) derive configurations for NOC, where control by either *Mary* or *John* is possible.

Landau (2001) thus revised the Super-Equi paradigm, arguing that obligatory, local control was forced in examples like (76a), for the local experiencer argument (e.g. *John*), but not for non-experiencers. This crucial distinction, dependent on the type of predicate—psychological or non-psychological—in the matrix clause, was an accepted fact about Super-Equi extraposition until Motut and Grant (2017) argued, using experimental evidence, that the judgements in (76) did not in fact hold. I turn to this study next.

2.4.4 Motut and Grant (2017)

Motivated by variability in the reported judgements for this data in the literature as well as by our own intuitions, Motut & Grant (2017) tested the paradigm laid out in (76) with two experiments. The first was a naturalness ratings study conducted with 55 participants (self-reported native English speakers living in the United States) using Amazon Mechanical Turk. Participants were asked to rate on a scale of (1) (completely unnatural to (7) (completely natural) a set of sentences drawn from a 24-item set, where each set consisted of four

conditions corresponding to the four possible combinations of the two independent variables:

type of predicate (psychological or non-psychological) and type of control (long-distance or

local), shown in (78).[28] Psychological predicates introduce experiencer arguments, so in this

way we expected to test for the robustness of the experiencer intervention effect reported by

Landau (2001) and others.

(78) Example item set from Motut and Grant (2017: 390):
     a. *Local control, Non-psych predicate*
       Becky knew that it disgraced Tim (PRO) to grind his teeth during the exam.
     b. *LD control, Non-psych predicate*
       Becky knew that it disgraced Tim (PRO) to grind her teeth during the exam.
     c. *Local control, Psych predicate*
       Becky knew that it angered Tim (PRO) to grind his teeth during the exam.
     d. *LD control, Psych predicate (reported to be ungrammatical)*
       Becky knew that it angered Tim (PRO) to grind her teeth during the exam.

We expected that local controllers might be preferred over more long-distance ones

for parsing reasons. However, if Landau's analysis of experiencer arguments was correct, and

experiencers were in fact categorical intervenors, we would expect to see an interaction in

our results such that this condition—long-distance control with a psychological predicate, or

(78d) above—would be rated significantly worse than the other three conditions.

Instead, we found two main effects for our two main variables, but no interaction.

Sentences with psychological predicates were rated as more natural-sounding than sentences

with non-psychological predicates ($\hat{\beta}$= 1.029 (±0.216), z = 4.766, p < .001), and local

---

[28] Note that I use the same convention here as in the immediately preceding section, of referring to control by the matrix internal argument, whether it be experiencer or theme (i.e. *Tim* in [78]), as "local" control and control by the matrix subject (i.e. *Becky* in [78]) as "long-distance" control. From a configurational perspective, however, both of these are NOC contexts and therefore long-distance control.

controllers were rated as more natural than long-distance controllers ($\hat{\beta}$= -0.633 (±0.289), $z =$ -2.188, $p < .05$). There was no interaction that would support the claim of categorical ungrammaticality for the long-distance/experiencer condition (i.e. [78d]). In fact, the difference between participants' ratings for the local versus long-distance control sentences was greater than for the non-psychological predicates than the psychological ones, indicating that participants were more likely to disprefer long-distance control in the presence of more local *non*-experiencer arguments (rather than experiencers). This is the opposite of what we would have expected if the paradigm in (76), with corresponding judgements, were correct.

These results were then confirmed with an eye-tracking study. 29 native English speakers from the University of Toronto community were recruited to participate in a sentence reading experiment. (Data from 29 additional participants was excluded for a variety of reasons.) Participants read experimental sentences (similar to the sets for the naturalness ratings study), and the movement of the right eye during reading was tracked by an head-mounted Eyelink II eye-tracker (SR Research: Mississauga, Canada). The relevant dependent variables for this kind of study are the duration of fixations on a critical word or region during the reading task (Staub & Rayner 2007). For these experiments we reported on First-Pass Time, which is the sum of all fixations on a region from first entering the region until leaving it, either to the left or right, and Right-Bounded Reading Time, which is the sum of all fixations on a region from first entering it until leaving it to the right.

For First-Pass Time, we found a marginal advantage for local control over long-distance control sentences ($\hat{\beta} = -33.60$ (± 19.69), t = -1.707, p < .1). In line with the results from the naturalness ratings study, we also found an advantage for the psychological predicate condition (experiencers) over the non-psychological predicate condition (non-experiencers), but this difference was not reliable ($\hat{\beta} = 23.15$ (± 21.21), t = 1.092, p = 0.29).

Crucially, the First-Pass Time results do not suggest an interaction such that there was a greater processing penalty incurred for the psychological predicate, local controller condition, which is what we would have expected if experiencers categorically intervene and blocked long-distance control in these constructions.

Table (2) summarizes the judgements as reported (or tested) in the literature. The literature summarized in Table (2) is not exhaustive, but note that those represented in the first two rows presumably reported their judgements based on their own or colleagues' intuitions. They did not test these judgements experimentally, as Motut and Grant (2017) did. Thus we can see that over the course of nearly fifty years of study of NOC in Super-Equi, our assumptions about the grammaticality of the key data itself have undergone significant revision. The judgements as reported in Motut and Grant (2017) are compatible with the claim that in Super-Equi, control by either of the arguments outside the infinitival clause (i.e. *Tim* or *Becky* in [78], *Mary* or *John* in [76]–[77]) is "long-distance" control and falls into the domain of NOC, *contra* Landau's (2001) claim that the experiencer arguments of psychological predicates in the matrix clause force OC.

Table 2

Predicted grammaticality judgements	Intraposition Psych Pred.	Intraposition Non-Psych Pred.	Extraposition Psych Pred.	Extraposition Non-Psych Pred
Grinder (1970), Williams (1980), Lebeaux (1984), Manzini and Roussou (2000)	✓	✓	✶	✶
Landau (2001)	✓	✓	✶	✓
Motut & Grant (2017)	✓	✓	✓	✓

Key: The symbol ✓ means the construction is predicted or reported to be grammatical; the symbol ✶ means it is predicted or reported to be ungrammatical.

Chapter 3: Updating the Remote Agree Hypothesis

In Chapter 2, I reviewed previous accounts of OC and NOC phenomena, including the early syntactic accounts of NOC which Landau (2013) and many others have argued fail on empirical grounds. In Section 2.4.4, I reviewed experimental results from Motut and Grant (2017) that challenged the previously assumed grammaticality judgements for NOC in Super-Equi, a construction that has featured prominently in analyses of NOC. In Chapter 1, a novel hypothesis for Remote Agree—Agree into the syntactic Workspace without subsequent E-Merge of a syntactic object—was presented, based on prior work by Motut & Oda (2011). In the next chapter, I revise and expand the Remote Agree Hypothesis, applying it to the analysis of NOC, the empirical test case for the hypothesis that I explore in this thesis.

In the sections that follow, I articulate assumptions about the Workspace, Spellout of syntactic objects, and the contents of the Workspace that are crucial to implementing the Remote Agree hypothesis. This elaboration will also help to explain how we avoid overgeneration with the Remote Agree Hypothesis. I will also propose that the Workspace contains the element π, a minimal syntactic element which is an "ontological primitive" of the Workspace. The details elaborated in this chapter will allow me to put forward, in Chapter 4, an account of NOC that is partly structural, reviving to a degree the idea that search and locality are relevant to the NOC/OC distinction. I will also modify the Remote Agree Hypothesis, arguing that Agree can target π in the Workspace as a last resort operation at the end of the construction of a syntactic object, which is a different formulation of the proposal than was presented in Motut and Oda (2011a,b).

3.1 Theoretical Assumptions about the Workspace and Operations

Under the Remote Agree Hypothesis (Motut and Oda 2011a,b), a probe can reach beyond the syntactic tree under construction to an object in the Workspace. (This is required in any theory that includes an Agree-triggered External Merge). In addition, there will be reflexes (morphological or formal) of this external Agree operation on the probe due to feature checking between the probe and the goal. Motut and Oda also claimed that Agree operations within the syntactic object of the probe (internal Agree operations) are more local than those that extend beyond the tree into the Workspace. (This is called "Move before Merge" by Motut & Oda, echoing an idea mentioned but not pursued by Chomsky (2013). All of these claims have implications for the Workspace and for Spellout of syntactic objects. My theoretical assumptions about the syntactic Workspace are given in (79). I will articulate each of these in turn.

(79)   *Theoretical Assumptions about the Workspace*
        a. The Workspace is syntactically real.
        b. The Workspace consists of independent subtrees constructed in parallel.
        c. Probes can search into the Workspace.
        d. The Workspace contains, by default, the ontologically primitive syntactic object, π.
        e. Search within the tree is more local than search into the Workspace or into a
        separately constructed syntactic object. The probe is assumed to be the starting point
        of such derivations.

3.2 The Workspace is Syntactically Real

The Remote Agree Hypothesis has theoretical implications for how we conceptualize the Workspace, requiring that the Workspace be syntactically real in a way that many previous proposals do not commit to. In many if not most prior theoretical works, the Workspace is not a real "space": operations do not apply "in it" and it is not part of the

syntactic search space. Frequently, the space in which operations apply is not explicitly discussed at all. In such work, the Workspace is essentially interchangeable with the Numeration; that is, an unordered set of lexical items that will enter into operations and be inserted into the syntactic object under construction. The Numeration in (80a) gives the sentence in (80b).

(80) *The Numeration*[29]

    a. Example Numeration: {the$_2$, dog, cat, bit}

    b. Example Sentence: The dog bit the cat.

The Remote Agree proposal makes the Workspace concrete, since a dependency may form between the syntactic object under construction and an element in the Workspace without the subsequent Merge of that item. Thus, the domain of application of operations is not only the syntactic tree under construction. It is at least plausible that all Agree-triggered conceptions of Merge must involve a similarly concrete view of the Workspace.

Although as syntacticians we make no claims about the real-time processing of the derivations occurring in our mental grammar, we assume for concreteness that our derivations have some properties as if they were occurring in real time (and space). This does not mean that the derivations must proceed in real time/space—they do not—but in our model, terminological correlates of processing (like a "linear" timeline, a "workspace") are necessary to articulate how a derivation proceeds. For instance, subtrees need to be constructed separately, we might imagine in separate subparts of the Workspace. We assume subtrees may be constructed in parallel and that some large trees may require that separate components be constructed independently before they are put together into the larger tree.

---

[29] Note that for simplicity here, I ignore Tense and other functional heads that would also be in the Numeration.

Whether or not one adopts Remote Agree, it must be the case that independent, parallel construction of subtrees occurs in the Workspace. It is required, for instance, for any structure containing a complex specifier. In order to avoid violating the Extension Condition (see [81]), the components of the complex DP must be Merged to form that DP before the DP as a whole is merged as a specifier. These components cannot be Merged one at a time with the rest of the clause, but must first be combined as an independent subtree. Nunes (1995: 82, fn. 48) says: "Assembly of unconnected phrase-markers is independently required for any phrase involving a complex specifier or a complex adjunct […]." This means that for trees, parallel construction of subtrees must happen which contribute to the larger whole.

(81) *Extension Condition*: Applications of Merge can only target root syntactic objects.
    (Hornstein et al. 2005: 63)

Hornstein, Nunes and Grohmann (2005: 63) give the example sentence in (82a). The derivation for this sentence must proceed with first the merger of *the* and *woman*, and *saw* and *George*, independently. If one were to Merge *woman* directly with the structure as shown in (82b), targeting the root object as per the Extension Condition, one would generate the structure in (82c), in which *the* and *woman* do not even form a constituent!

(82)   a. The woman saw George.
        b. [woman [Infl [saw George]]]
        c. [the [woman [Infl [saw George]]]]

The authors note that before the first step in the (correct) derivation combines *the* and *woman* into the DP *the woman*, "there are three root syntactic objects available to the computational system […] In fact, it won't be uncommon that in building a sentence we may have several 'treelets' around prior to their combining into a single big tree" (Hornstein et al. 2005: 63).

These arguments show us is that it is necessary, independently of my proposal, for independent sub-trees to be constructed in parallel in the Workspace. This conclusion also provides an argument in favour of the concreteness of the Workspace. The space of derivations is not just a single syntactic object: our operations must act on many syntactic objects at once and choose between them, and operations must span multiple syntactic objects. This means that Search of possible goals by various probes may span across multiple syntactic objects as well, and the Workspace itself may be the target of Search, which will be a crucial component of my analysis.

The construction of independent subtrees in parallel, and the fact that operations have access to the Workspace, are not the only assumptions required. We also need to articulate how Spellout and locality work among multiple syntactic objects in the Workspace being constructed in parallel. I will also introduce assumptions about what kinds of *syntactic objects* we might find in the Workspace.

3.3 What is in the Workspace? A π Proposal

A consequence of making the Workspace concrete is that we can then populate it with syntactic objects that are inherent to it: the set of objects that are available to any syntactic derivation by default. This of itself may be a novel idea, and I will focus here on just one such type of syntactic object. I will claim that in addition to the lexical items for a given derivation that are to be Merged into the tree, the Workspace contains a kind of primitive feature (or feature bundle) that is special by virtue of being ontologically necessary, and therefore privileged to appear by default in the Workspace and be available to derivations and operations as such. This feature will be instrumental in the mechanics of the analysis I work out below.

By minimally ontologically necessary, I mean it is available to the syntax by virtue of being an ontologically essential component of a derivation, similar to Chomsky's argument that Merge is virtually conceptually necessary, and is thus "free". This feature is available "for free" in any derivation (with some restrictions, to be articulated below).[30]

What is this primitive syntactic object? The simplest hypothesis is that this Workspace element is a very minimally specified feature / syntactic element, so this is the hypothesis I will adopt. I assume that the Workspace, in addition to being a syntactically real space in which operations can apply, is also populated with these very minimal types of syntactic objects, which at their most basic are just bare person nodes whose default interpretation is generic [+HUMAN]. A bare person node, π, without further specification, will be interpreted as [+HUMAN] but with no other identifying feature (gender, participant, etc.). I will keep the make-up of this default element as minimal and underspecified as possible.[31]

Is this a new element? I make the same point as Saab (2014) and others that it is preferable not to expand the ontology of empty categories (e.g. silent/non-overt categories) in our theoretical model unless forced to. Adding a new element to our inventory of empty

---

[30] To be clear, the part of the argument that I'm borrowing from the "Merge is free" story is the part that argues for the necessity of the operation—or in this case, the syntactic element in the Workspace—in the model, not that it must therefore be the only operation, since my model will also include at least Agree.

[31] Note that this element is not quite a "minimal pronoun", although I will compare it to such entities below. It is indeed more minimal than that. It cannot receive independent reference and is not specified for typical phi-features (number, gender, etc.). Furthermore, I am not claiming here that π is the *only* ontologically primitive element which may reside in the Workspace by default. π is the element relevant to the empirical phenomena under discussion in Chapter 4, but it may be that there are other such elements (e.g. perhaps related to the verbal or discourse domains). Identifying all such possible elements would take us much beyond our current scope.

categories is not advisable under basic principles of methodological economy (i.e. "good science" or Occam's razor). Our Workspace element, π, however, may already be used elsewhere in the grammar. Kratzer (2009) proposes a similar (though not identical) very minimal pronoun to encompass several categories of elements that were previously distinct, including anaphors and PRO (see also Landau 2015, and his proposal for PRO as a similarly very underspecified element). Kratzer's minimal pronoun is very underspecified and can be realized as an anaphor, a default morphological shape, or can take on features from other elements in the derivation. My proposed Workspace element, π, is not a pronoun but is similarly very minimally specified. Since π is an element, a feature, used elsewhere in our grammatical derivations, we can say that this Workspace element is not altogether new. I define π in (83).

(83)   **Definition of π**: A minimally specified syntactic element, a bare π node, [+HUMAN] but with no specification for number, gender, etc., which is present in the Workspace by default and available to syntactic operations.

In my system, such an element remains in the Workspace and can Agree with probes in syntactic objects, and supply (minimal) feature specifications to probes within syntactic objects. In Chapter 4, I will show how π, or rather its semantic denotation, can also be co-indexed with a DP elsewhere in the derivation. I'll assume that π cannot be spelled out on its own, nor can it Merge into the syntactic tree under construction: it is simply an inhabitant of the Workspace, a kind of universal primitive available to participate in derivations. Below I describe additional limitations on when and where π is available to a derivation so as to avoid over-generation.

3.4 The Workspace, the Numeration, and π: Types and tokens

A concern may have arisen: What else is in the Workspace, and just how cluttered is it? How does this more concrete conceptualization of the Workspace bear on traditional Minimalist derivational concepts like the Numeration, sub-arrays, and the Lexicon?

First, we need to make a distinction between the Workspace and the Lexicon. The Workspace is not the Lexicon: it does not contain all of the possible syntactic objects and lexical items for any given derivation. If the entire lexicon resided in the Workspace, then external Search operations into the Workspace would have a potentially unbounded space in which to search. Furthermore, as I show in the next chapter, I derive results from appealing to Remote Agree with π as a kind of last resort, available when other constituents/lexical items are *not* available in the Workspace to Merge. Having the entire lexicon in the Workspace would undermine the intended implementation of the mechanism.

Instead, I assume that the Lexicon, composed of lexical types, is distinct from the Workspace, and "feeds" the Workspace tokens as needed (Chomsky 1995a, *inter alia*). I further assume, for reasons that will become apparent later, that there is a distinct Workspace (or sub-Workspace) for each syntactic object under construction. As for the Numeration, if the Numeration (or sub-array) is simply the set of lexical items that will feed a given derivation or sub-derivation, then the Workspace for a given syntactic object may be indistinguishable from the Numeration. I set this possibility aside, and will henceforth refer to the object under discussion as the Workspace.

While π is a universal feature of the Workspace, Agreeing with it in the context of building a syntactic object creates a token for that syntactic object. As we will see, there can be different instances of π corresponding with different Spellout domains. The details will

become clearer in Chapter 4 when we apply the Remote Agree operation and π to the NOC data.

3.5 What is the Minimal Spellout Object?

Another relevant question is what is the minimal unit of Spellout? What kinds of constituents can or must be spelled out? The notion of a "minimal Spellout object" has consequences for how the syntactic Workspace functions and how Spellout proceeds with multiple syntactic / Workspace objects. Below I take a small digression into subsentential fragments to argue that the minimal Spellout unit is neither a sentence nor a clause.

Syntacticians assume that the basic unit of our derivations is a sentence containing one or more clauses. But utterances do not always form complete sentences. In fact, they more frequently don't in natural dialogue. Consider, for instance, the natural-sounding dialogue in (84).

(84) *Dialogue Snippet*

Speaker 1: Which guy did you come to the party with?
Speaker 2: Him.
Speaker 1: Who?
Speaker 2: With the hat.
Speaker 1: How did you get here?
Speaker 2: Drove.

This dialogue contains several full utterances that constitute less than a complete clause or sentence: [$_{DP}$ him], [$_{PP}$ with the hat], [$_{VP/vP}$ drove]. All of these examples appear as responses to questions, but there are other contexts in which sentential fragments may be felicitously uttered. For instance, in the following situations:

(85) a. *Child pointing to a tail disappearing into the shrubbery:*
      Child: [DP A cat!]

   b. *Speaker pointing at an unidentifiable mess on the ground:*
      Speaker: [ADJP Gross!]

Example (85a) is clearly some kind of existential context where the child is communicating 'there is a cat!', or 'I see a cat!'. One plausible analysis of sub-sentential fragments is that there is a lot of covert structure (i.e., it is a full clause or a sentence with elided portions), that what is being spelled out is actually a much larger structure than the sub-sentential constituent that is pronounced. This is, of course, possible, given standard assumptions about the phonological component being processed separately than the semantic component.

For examples like those in (84) where the sub-sentential constituents are uttered in response to questions, it seems plausible that there is elided clausal structure. We can consistently reconstruct what the full response would be. But in cases like (85), it seems less likely that the DP in (85a) and the AP in (85b) contain larger elided or covert structures. In these cases, it is not possible to reconstruct consistently what the intended full-clause utterance would be. In these exclamatory contexts, and possibly elsewhere, I assume that what is being spelled out is just the overt constituent, with no covert structure.

Under this assumption, it is necessary to make a distinction between Spellout objects and phases, which we turn to next.

3.6 Spellout Objects, Syntactic Objects, and Phases: Some definitions

I argued above that subsentential fragments smaller than a clause can be spelled out, that is, both pronounced and assigned a meaning. This will help us determine what is a Spellout object and what isn't, and I will make a technical distinction between, for instance, a

Spellout object and a phase. This distinction will be important to the reformulation of the NOC / OC generalization that I present in the next chapter.

There are many versions of the theory of how and when syntactic objects are spelled out (Uriagereka 1999, Chomsky 2001, Simpson and Wu 2002, Fox and Pesetsky 2005, Kratzer and Selkirk 2007, Fowlie 2013, among many others.) Below I articulate a version of Spellout in which I will assume, as per the discussion in the immediately preceding section, that constituents of any size or category may be spelled out, but not all Spellout objects are phases. Some of these distinctions I make for concreteness; others will be integral to examining the patterns of NOC reviewed in Chapter 2 (to be analyzed in Chapter 4), and to articulating how Remote Agree is to be constrained.

First, some definitions:

(86) a. *Phase:* A *v*P, CP, or DP constituent that, once constructed, is shipped to both LF and PF interfaces, is assigned interpretation and linearized. It is inert to further syntactic operations *once it is spelled out*. We may refer to a constituent as a phase even if it has not yet been spelled out.

b. *Spellout object:* A constituent that can be or has been spelled out. It may or may not be a *v*P, CP, or DP. A Spellout object when spelled out is assigned interpretation and linearized. It is also syntactically inert to further operations. π is not a Spellout object. Individual functional heads are not Spellout objects.

c. *Minimal Spellout object*: The smallest syntactic unit that can be spelled out.

d. *Workspace object*: Any syntactic object in the Workspace. This may include functional heads, features and feature bundles, constituents, sub-trees, etc. π is a Workspace object, along with the lexical items of the Numeration, etc.

Any constituent that can be pronounced and assigned meaning can be spelled out. This means that Spellout of constituents is not limited to *v*P, CP, and DP phases. A PP, for example, as in the dialogue in (84), or a TP could be a Spellout object, but they are not

phases.[32] Not all Spellout objects are phases, but all phases are Spellout objects. Functional heads are *not* Spellout objects, and neither is Workspace π. In a sentence with multiple CP complements, embedded recursively, as in (87), each CP is a phase, spelled out after the completion of the CP above it (Chomsky 2001), and these are all Spellout objects.

(87)   [John said [that Mary thought [that Jane wanted [PRO to eat all the pie.]]]]

Once spelled out, all Spellout objects, including phases, are inert or "frozen" to further operations (cf. Chomsky's 2001 Phase Impenetrability Condition, and earlier proposals by Wexler and Culicover 1977, 1980). Spellout objects encompass the following three categories, the first two of which may not be phases: (i) moved constituents;[33] (ii) complex specifiers or other constituents that must be Merged with another syntactic tree as a whole; (iii) phases within a larger Spellout object or tree.

3.7 Avoiding Overgeneration with Remote Agree: Rules of Merge and Spellout

(88) adds a further assumption about Spellout and E-Merge to our understanding of derivations. Recall that I assume, as is standard, that multiple sub-trees must be constructed in the Workspace in parallel. I also propose that each of these subtrees is a Spellout object, and Spellout objects *must* be spelled out during the course of the derivation of the utterance. They cannot remain in the Workspace and *not* be spelled out at some point in the derivation, since this would clutter the Workspace and increase the amount of Search necessary for a probe.

 (88) *Spellout objects must be spelled out:* A Spellout object must be spelled out during
         the course of a derivation.

---

[32] PP may, in fact, be a phase (Abels 2012), but I assume its status as such is less consistently assumed than vP, CP and DP.

[33] See Wexler and Culicover (1977, 1980) on the frozen status of moved constituents.

Table (3) puts all of these pieces together and summarizes the rules of Spellout and E-Merge in the Workspace, once we incorporate Remote Agree and the availability of Workspace π.

A full DP, because it is a Spellout object, may be spelled out on its own in the Workspace, and pronounced (at the A-P interface) and assigned reference (at the C-I interface), or E-Merged into another object in the Workspace. However, π, as a Workspace object that does not meet the minimum criteria for Spellout, is neither E-Merged nor independently spelled out. In this way it differs from a pronoun. A fully specified pronoun[34] in the Workspace *is* a Spellout object and must be spelled out. It can be E-Merged with another syntactic object, and it cannot remain in the Workspace (not spelled out) the way π can.

---

[34] That is, a DP or other nominal category such as φP. See Déchaine and Wiltschko (2002) and much subsequent work for discussion.

Table 3: Rules of E-Merge and Spellout
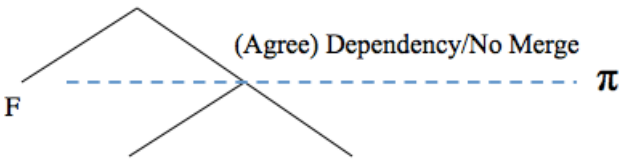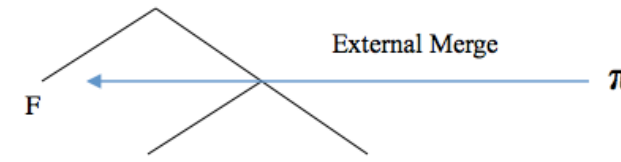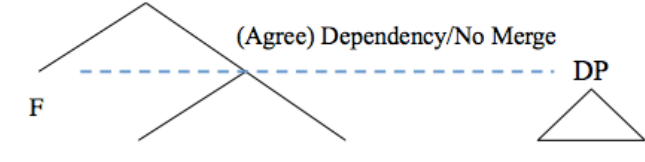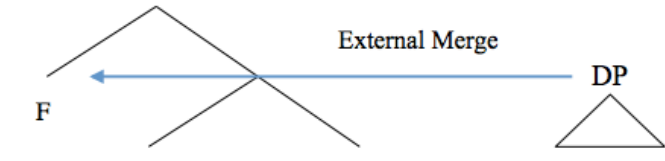
Ex.		Is Spelled Out?
a.	(Agree) Dependency/No Merge — F ... π	Yes
b.	External Merge — F ... π	No
c.	(Agree) Dependency/No Merge — F ... DP	No
d.	External Merge — F ... DP	Yes

Table 3 shows the options available for E-Merging versus forming a Remote Agree dependency without E-Merging, across two different types of syntactic objects in the Workspace: Remote π or a DP. Dotted lines indicate Agree and solid lines indicate E-Merge. The structure in (a) can be spelled out because Workspace π is allowed to remain in the Workspace without being spelled out. A dependency spanning multiple syntactic objects in the Workspace is a Spellout object, since π cannot be spelled out independently. Indeed, π's

lack of phonetic content means the syntactic tree in (a), even with its Remote dependency, can still be linearized.

The structure in (b) is not a Spellout object as π must remain in the Workspace. (c) is not a possible Spellout structure because the DP in the Workspace is independently a Spellout object and as such must be spelled out. However, so must the other sub-tree, and the dependency linking the two sub-trees makes this an un-linearizable structure. (d) is a very standard case: a complex DP is constructed as an independent sub-tree and E-Merged into the other tree. This is a Spellout object (can be spelled out) once the DP has E-Merged with the other sub-tree. The DP itself is a Spellout object as well, and will be spelled out separately. These assumptions about how Spellout works are necessary to constrain the Remote Agree operation and prevent it from over-generating.

In sum, Remote Agree with a DP in the Workspace must be followed by E-Merge of that DP at some point during the derivation. Remote Agree with a DP without subsequent Merge of that DP is disallowed. However, π can form a Remote Agree link with a probe and remain in the Workspace, as it is not a Spellout object.

Chapter 4: Remote Agree with π Applied to NOC in English

In this chapter, I present an analysis of non-obligatory control (NOC) based on the revised Remote Agree Hypothesis outlined in the previous chapter. I will start by formulating a New Generalization for the NOC / OC distinction in English, capturing the distribution of each in terms of Spellout objects / domains rather than the standard terms (adjunct, complement, etc.).[35] I will then analyze arbitrary control using the Remote Agree hypothesis, proposing that Remote Agree is established as a "last resort" mechanism to Agree into the Workspace (with π) at the completion of a Spellout object (or when a Spellout object is spelled out). A feature percolation mechanism will be assumed that "carries" the unsatisfied features of lower argument-introducing heads, up to the final functional head of the Spellout object, in this case, C. C's (Remote) Agreement with π will then satisfy these unsatisfied feature slots on other heads.

After analyzing arbitrary control, I move on to an analysis of long-distance control that relies on the above-mentioned Remote Agree mechanism as the first step in its derivation. I will propose that (non-syntactic) co-indexation between the denotation the arbitrary control clause (which Remote Agrees with π) and a DP antecedent in another Spellout object results in long-distance control. Then I will review the pragmatic factors like logophoricity and topicality that have been argued in the literature to be relevant to the selection of long-distance antecedents for NOC, and argue that my account is consistent with them.

---

[35] This discussion is focused on English. In particular, I make no claims about, e.g., null subject languages, whose control properties may differ substantially.

4.1 Arbitrary Control

As I demonstrated earlier, "arbitrary control" is a phenomenon where we associate a thematic argument slot with a generic or arbitrary reference. Because its syntactic distribution is similar to that of long-distance and obligatory control, i.e. it appears in non-finite clauses lacking a subject, it is referred to as "control." But there is no "controller" or antecedent, at least no apparent one.

Arbitrary control shows the same distribution, moreover, as long-distance (LD) control, which *does* have a controller—a DP from which it gets its reference. Despite the fact that arbitrary control does not, it is traditionally analyzed as PRO_{arb}—a special kind of PRO that does not need an antecedent. One might surmise that, if LD control, which has the same distribution as arbitrary control, needs an antecedent, then so does arbitrary control. But what then is the antecedent for arbitrary control? I will argue that in arbitrary control constructions, there is no PRO to require an antecedent at all. There is no silent, Caseless DP Merged to satisfy EPP, or some other subject-requiring feature.[36] But there is nonetheless a *source* for the "reference" ascribed to the apparently empty argument slot. That source is Remote π in the syntactic Workspace. I claim that, instead of PRO_{arb}, what is standardly referred to as "arbitrary control" in the literature is the interpretation that arises when an empty argument slot and subject position are satisfied by an Agree relation between a functional head or heads in the clause and π in the Workspace.

To give a preview of the proposed analysis, I will argue that in arbitrary control, an Agree relation is established between a functional head and π, satisfying certain features on that functional head. In Section 4.1.1, I will suggest a feature percolation mechanism such

---

[36] I will assume, likewise, for OC, an account that also lacks PRO. My account of NOC is compatible with the movement theory of obligatory control, *inter alia*.

that unsatisifed features can "percolate" up to the C head and Agree with π at Spellout, satisfying all unsatisfied features on that head.

The element π supplies the appropriate interpretation for arbitrary control: it is a very minimal bare π node interpreted by the semantic interface as [+HUMAN] but with no other specification of gender or person, etc. This is precisely the interpretation we want for arbitrary control. We saw earlier that one of the most striking—and perhaps unusual features of arbitrary control (and long-distance control)—is its persistent [+HUMAN] interpretation, in contrast to OC, which lacks this requirement. I will propose in the following sections an account of both arbitrary control and long-distance control that relies on this Agree with π mechanism. For NOC's interpretation to be supplied by [+HUMAN] π solves the mystery of why NOC is always [+HUMAN] and where this generic, arbitrary interpretation comes from. From an interpretation standpoint, π supplies exactly the right meaning.

From a syntactic perspective, we need to answer the question of why Remote Agree with π, forming an arbitrary control interpretation, is only available in the specific syntactic contexts where arbitrary control (or NOC more generally) appears, and not elsewhere. For example, why is it not available in contexts where obligatory control appears?

The distribution of NOC, including arbitrary control, contrasts with that of OC. This complementarity is a key and foundational feature of the distributions of each of these phenomena. Landau, Hornstein, and others, as we saw earlier, have argued that NOC occurs in contexts that are opaque to movement. Landau's (2013: 231) generalization on the distribution of OC vs. NOC is given in (89).

(89)   a. complement clauses fall under OC;

         b. subject and adjoined (extraposed) clauses fall under NOC.[37]

4.1.1 A New Generalization

I make a new claim about the distribution of NOC versus OC: that all of the contexts where

NOC occurs are in clauses that constitute a Spellout object / Spellout domain. OC occurs in

contexts where the controlled clause appears in a larger Spellout domain that encompasses

several clauses.

When control occurs in a complement clause, the controlled clause and its controlling

DP are in the same Spellout domain or Spellout object—essentially, they are in the same

syntactic tree. We may think of this in terms of the transparency of these contexts to

movement (and indeed I will adopt a movement theory of control for OC), but also in terms

of Spellout domains. OC occurs between clauses that have not yet been spelled out. For those

types of clauses in which we observe NOC, according to this generalization, the long-

distance controller and the controlled clause ("PRO") are *not* within the same Spellout object,

but in different Spellout objects.

More concretely, a subject clause is constructed as a separate Spellout object in the

Workspace. Extraposed or moved clauses, and adjuncts, are also constructed as separate

syntactic objects in the Workspace before being E-Merged into the primary tree. They must

be spelled out, and are therefore frozen, in the sense of Wexler and Culicover (1977, 1980)

---

[37] This triplet, repeated from (54) above (Landau 2013: 38), illustrates this:

(i)   a. We$_i$ thought that John would help Mary$_j$ [PRO$_{j/i*}$ to expose herself/*ourselves]

       b. We$_i$ thought that [[PRO$_{j/i}$ to expose herself/ourselves]] would help Mary$_j$

       c. We$_i$ thought that it would help Mary$_j$ [[PRO$_{j/i}$ to expose herself/ourselves]]

prior to their Merge with the main syntactic tree under construction. Some examples below illustrate these differences.[38]

(90)   a. **[**John wants [(PRO) to smoke cigarettes indoors]**]**_{Spellout Domain}

      b. **[**(PRO) To smoke cigarettes**]**_{Spellout Domain} is bad for one's health.

      c. John thought that it annoyed Susan **[**(PRO) to smoke cigarettes**]**_{Spellout Domain}

      d. **[**John wanted [(PRO) to try [(PRO) to smoke cigarettes indoors]]**]** _{Spellout Domain}

(90a) is the context for OC, while (90b,c) are contexts for NOC: the controlled clause is a subject clause in (90b), an example of arbitrary control where there is no antecedent, and in an extraposed (i.e. moved) position in (90c), where there is a long-distance antecedent.[39] In neither (90b) nor (90c) is there a possible antecedent in the *same Spellout domain* as the position marked with PRO. (90a) and (90d) are examples of OC. (90a) is a simple case of OC, with two clauses, and the controlled clause and its obligatory antecedent are in the same Spellout object. In (90d), OC holds across multiple clauses, and although the antecedent of the controlled clause, *John*, is two clauses removed from the controlled clause, all the clauses are within the same Spellout object.[40] Thus we arrive at what I will call the *new generalization* of NOC versus OC, in (91).

---

[38] The bolded brackets, **[ ]**, indicate the boundaries of the Spellout object or domain. I have used "PRO" in these examples for ease of exposition although I will not adopt a PRO analysis.

[39] In fact there are two possible long-distance antecedents: *John* or *Susan*.

[40] The relevant empirical domain here is the Spellout object or domain rather than the phase. In (90d), there may be a phasal boundary between *John* and the most embedded infinitival, but this object as a whole has not been spelled out. Another way to get at the same idea might be to think of the distinction in terms of derivational timing.

(91)   *New Generalization, NOC vs OC:*

    a.   OC occurs in environments where the controlled clause and its antecedent occur within the same Spellout object.[41]

    b.   NOC occurs in environments where the controlled clause is in a different Spellout object from the source of its reference—π or a DP antecedent.

I now digress briefly to discuss the locality implications of this new generalization. In the old "structural" (syntactic or search-based) accounts, the complementarity between NOC and OC was derived by having more local (in the sense of c-command or command) intervenor DPs block long-distance control (NOC) and force OC; the closer antecedent to PRO would take precedence. The old generalization from such accounts (e.g.,Grinder 1970, Lebeaux 1984) was thus that the choice between a local and long-distance antecedent, and thus between local and long-distance control, was determined by locality. This intuition has some sense to it. Dependencies of all sorts tend to favour more local relations over less local ones: why should this not also hold for the antecedent of PRO?[42] This story worked less well with arbitrary control, however (how does one compare the locality of possible antecedents when PRO lacks an antecedent altogether?) In the account proposed here, we still have a locality story, but it is not quite the *same* locality story.

I propose a partial resurrection of the structural accounts of NOC through the use of Remote Agree with π, and by including the concept of "Spellout object" and the syntactic Workspace in the locality calculations for a given derivation. Syntactic objects thus become relevant to locality computations: resolving a dependency or satisfying an unvalued feature

---

[41] This definition remains agnostic as to whether the controlled subject is PRO, trace, copy, or some other element.

[42] Grinder (1970), whose account comes before the introduction of PRO, framed this in terms of local NPs as intervenors blocking control by less local ones.

on a probe within the same syntactic object as the probe is more economical to the derivation (i.e. it requires less Search) than resolving it by going to the Workspace.[43] If π, and therefore by extension arbitrary control (Remote Agree with π) is available only if a probe's features cannot be satisfied by an element within the tree, we end up predicting that arbitrary control is a kind of last-resort strategy for satisfying the requirements of the clause (i.e. features on v and T). This is summarized in (92). Arbitrary control will therefore arise only if these requirements cannot be satisfied otherwise by an element in the probe's c-command domain or specifier. The syntactic Workspace (and π) is the "non-local" option for resolving unsatisfied features on v and T (or an unfilled argument slot in the verbal phrase and subject position), for a given clause and syntactic object.[44]

(92)   Remote Agree with π applies as a last resort at Spellout of a Spellout object.

To say the syntactic Workspace is the "non-local" option brings the Workspace, as a specific theoretical construct, into the locality computations of the derivation. (This is, to my knowledge, a new way of looking at the Workspace.) And it allows us to form a new generalization about the distribution of NOC, including arbitrary control, versus obligatory control: NOC appears when a probe Agrees with π in the Workspace.

---

[43] This is similar to Wurmbrand's (2014a) Agree-triggered Merge account, where E-Merge happens as a last resort when a probe's unvalued features cannot be satsified by an element already in the tree, although for Wurmbrand this is not anything like Remote Agree but instead simply a last-resort insertion / Merge triggered by economy.

[44] Note that this kind of locality is not framed purely in terms of c-command or linear order, but rather is defined over potentially several syntactic objects: the Workspace is not part of the c-command domain of a probe within a given tree. Instead, tree (syntactic object) boundaries are relevant to the amount of Search, and thus the relative locality, of a given probe.

The Workspace also contains DPs in the Numeration, not just π. To a probe, these DPs are also less local than a goal within the same syntactic object as the probe. Like Wurmbrand (2014a), I predict that E-Merge of DP, or Remote Agree with π is a last resort. The restrictions on Spellout that were outlined in Chapter 3 will help to rule out unwanted overgeneration: Remote Agree with a DP will result in E-Merge. It is only π that has the ability to stay in the Workspace in a dependency with a probe in the tree. The element π itself does not go "searching" for an antecedent or to fulfill unvalued features: it has none. It remains available in the Workspace as a last resort to fulfill specifications on probes in the tree, but does not itself need satisfying.

How do we arrive at the complementarity between OC and NOC, based on the generalizations above? Let's walk through some derivational paths to see how this is achieved. If we adopt the movement approach to OC, then the complementarity between OC and NOC will fall out from the presence or absence of a DP in the Numeration / Workspace, as well as some other assumptions.

We assume that the Lexicon is distinct from the Workspace: they cannot be identical because giving the derivation access to the entire Lexicon would introduce too much complexity into the Workspace.[45] Thus we still require some form of sub-array or Numeration in our understanding of how derivations work. We will assume that the lexical items for a given derivation are supplied to the derivational Workspace for the construction of a specific syntactic object. Remote π is always available in the Workspace. We will also introduce the stipulation that overt DPs in the Workspace or other lexical items (functional heads) will take priority over Remote π in Searches of the Workspace. If there is a DP

---

[45] That is, Search of the Workspace, prior to E-Merge or Remote Agree, would have to go through each item in the Workspace. This was discussed in Chapter 3.

available in the Workspace, it will be Merged, taking precedence over Remote Agree with π. This departs from Motut & Oda (2011a,b), who claimed that all objects in the Workspace were equidistant from the probe in the tree.[46] This is how it turns out that Remote Agree (and therefore NOC) is available only as a last resort, in the absence of another DP in the Workspace.

Let's look at how a basic derivation proceeds from the point where we have already Merged V and a DP object, as well as v / *Voice*,[47] and v is searching, as shown in (93). Assume that E-Merge of functional heads and arguments up to this point was triggered by Agree. At the point in (93) where v is looking for something to satisfy the external argument position, we return to the locality relations discussed earlier. According to Motut & Oda (2011a,b), Search of the complement of the probe is more local than Search of the Workspace: hence, "Move" (or I-Merge) over E-Merge. In this case, v probes its complement, but the object DP already in the structure is inactive (its features are already checked). Thus v is free to probe the Workspace, which it does. This is where the derivational paths diverge depending on what the Workspace contains.

(93)   [$_{vP}$ $v_{u\varphi:\_\_}$ [$_{VP}$ V DP]] ------------------------   Workspace: {....}
                        Agree/Search

If there is a DP in the Workspace, it will Merge into [Spec, vP]. From here, the derivation can take two paths. If finite T is then Merged (ignoring intermediate projections like Asp for now), we get the structure in (94a). This clause is licit as it stands and could be spelled out as

---

[46] I continue to assume that all objects besides π are equidistant from each other and from the probe in the tree.

[47] I use v as the functional head that licensed the external argument, for consistency with most of the control literature. It can equally be read as Voice.

a declarative clause: its derivational future might involve embedding within another syntactic object, but the clause could be Spelled out as a simple declarative. If a nonfinite T is Merged above the *v*P, on the other hand, and DP were to be Moved to Spec of this T, that would produce an illicit structure (as a matrix clause). Embedding CP1 under an obligatory control verb like *want* (94c) will result in a structure where DP2 has raised to the Spec of a finite T and the structure is therefore acceptable. (I take no position on the exact derivation of obligatory control via movement in this case, including whether DP2 passes through the Spec $T_{[-fin]}$ position in the lower clause.[48] The exact details are not important: what matters is that our locality considerations, plus the presence of an available DP in the Workspace, and further Merge of lexical items, result in an obligatory control structure.)

(94) a. $[_{CP} [_{TP} DP_2 \ T_{[+fin]} \ [_{vP} \ v_{u\varphi:SG.MASC} \ [_{VP} V \ DP_1]]]]$

John   helps Allie

b. $*[_{CP} [_{TP} DP_2 \ T_{[-fin]} \ [_{vP} \ v_{u\varphi:SG.MASC} \ [_{VP} V \ DP_1]]]]$

*   *John to   help Allie*

c. $[_{CP2} [_{TP} \mathbf{DP_2} \ T_{[+fin]} \ [_{vP} \ ..... \ [_{CP1} C \ [_{TP} \ \mathbf{\textit{t}_{DP2}} \ T_{[-fin]} \ [_{vP}.......]]] \ ]]$

John   tries   t_John to   help Allie

If there is no available DP in the Workspace, the derivational path leads us to Remote Agree with π, as follows. If there is no DP in the Workspace and Numeration for the construction of this clause, then the unsatisfied argument features, in this case uφ,[49] will remain unvalued. The derivation does not immediately crash, however (although it may at a later point).

---

[48] I will discuss ECM constructions in (106).

[49] There is precedent for assuming that uφ features can be the probing features in Agree-triggered (E-Merge), e,g., Wurmbrand (2014a).

Instead, the uφ on *v* "percolates" upward in the tree to the T and C projections. In this scenario, features on T will also percolate as they cannot be satisfied without a DP in the Workspace to E-Merge. We end up with the unsatisfied features from T and *v* on the C head, as in (95).

(95) $[_{CP}$ $C_{EPP:\_\_+u\varphi:\_\_}$ $[_{TP}$ $T_{EPP:\_\_+u\varphi:\_\_}$ $[_{vP}$ $v_{u\varphi:\_\_}$ $[_{VP}$ V DP$]]]]$

At this point, when C Searches the Workspace, it will find Remote π to Agree with, which will satisfy all of the unsatisfied features on the C head, automatically satisfying those in the lower heads as well:

(96) $[_{CP}$ $C_{EPP:\,\pi+u\varphi:\,\pi}$ $[_{TP}$ $T_{EPP:\,\pi+u\varphi:\,\pi}$ $[_{vP}$ $v_{u\varphi:\,\pi}$ $[_{VP}$ V DP$]]]]$

Although π is a permanent syntactic object in the Workspace, it does not Agree with just any head at any point in the derivation. A link with π can only be established at the end of building a Spellout object, or at Spellout of the syntactic object. This makes Remote Agree with π in the Workspace the last resort for satisfying these unsatisfied features. Thus, we derive the "New OC / NOC Generalization" formulated above.[50]

The complementarity of OC vs. NOC, at least as it pertains to OC versus arbitrary control (we return to long-distance control in Section 4.2), comes down to whether there is a DP in the Workspace that can Merge and satisfy argument features on *v*; the rest follows from the locality considerations already described. As discussed above, *v* searching the

---

[50] We might ask why Remote Agree with π does not occur in finite clauses. In English, we know that subjects of finite clauses need to be overt, and so Workspace π is not a suitable element to satisfy the features of finite T as it would leave the clause without an overt subject. We might encode this on T, whether by appealing to Case theory or the more traditional (i.e. Chomskian) view of EPP only occurring on finite clauses. This would require an adjustment to the features of T used in the derivations in this chapter, since I am assuming EPP is one of the features (on [-fin] T) that Remote π satsifies via percolation from C.

Workspace will either find a DP to satisfy $u\varphi$ or not. Agree with a DP in the Workspace will necessarily precede E-Merge of that DP into the structure: recall from the rules listed in Section 3.7 that DP, unlike π, cannot Agree with a probe in a syntactic tree *without* E-Merging. Crucially, once a DP has Merged, it will become the most local goal to any higher probe within the same Spellout object, until that object is spelled out. The rules of locality as we've defined them—that Search within the complement of the probe is more economical than Search into the Workspace and therefore more economical than External Merge—will guarantee that the DP in the tree is a more local option for further operations by higher probes within the tree, and that Remote Agree with π will be a less local and thus dispreferred option for higher probes. This guarantees that Remote Agree with π, and therefore arbitrary control, only occurs within a clausal Spellout object/domain where there is no available DP to satisfy the argument requirements of v.[51]

Returning to several of our earlier examples from (90) in (97) below (replacing PRO with trace to more clearly demonstrate my line of thinking), we see a contrast between the environments for arbitrary control and obligatory control. In (97a) and (97b), both showing obligatory control, the Spellout domain encompasses multiple clauses, each lower clause

---

[51] This pertains only to active goals / DPs in the tree: where there is no *active,* available DP goal in the complement of the probe, Agree into the Workspace will be licit. In the examples in (93) – (96), the object DP is already inactive because it has received Case and is inert. In this way, Case continues to play a crucial role in the Remote Agree story for NOC. Similarly, in (i) where we presumably have arbitrary control via Remote Agree in the subject clause, we note that *who* has moved. But it has moved for the purpose of satisfying a [+wh] feature. Since it has already received Case as the object of *trust,* it is not an available goal for probing v or T, which will instead Agree with Workspace π. Thus this sentence is predicted to be well-formed.

    (i) Who to trust is important.

selected by the verb of the higher clause.[52] DP *John* has Merged into this structure in the

lowest clause, and then it becomes the most local goal for the probe on *v* in the higher clause

(the obligatorily controlling clause). In this derivation, movement of a DP through multiple

theta positions or argument positions is allowed. I am not committed to the specific details of

the Movement Theory of Control (MTC), but adopt a movement approach here for obligatory

control only.

On the other hand, in (97c), we see that the infinitival clause *[to smoke cigarettes]* in

subject position has no available DP within the domain. This Spellout object must be spelled

out because it will be Merged in subject position of the matrix clause. Thus π in the

Workspace is the probe's (C's) only option for that Spellout object / domain.

(97)   a. **[**John wants [t_{John} to [t_{John} *v* smoke cigarettes indoors]]]**]** Spellout object

        b. **[**John wanted [t_{John} to try [t_{John} to [t_{John} *v* smoke cigarettes indoors]]]**]** Spellout object

        c. **[**To smoke cigarettes**]**Spellout object is bad for one's health.

4.1.2 Feature Percolation and Delayed Feature Satisfaction

Feature percolation is a mechanism whereby unsatisfied features on a given head may

appear on higher projections. Originally introduced to explain why pied-piped phrases (98)

can satisfy [Spec, C]'s requirement for a [+wh] constituent (Lasnik and Saito 1984)—the

[+wh] feature of the wh-word, it was speculated, "percolated" up to the entire pied-piped

phrase—and why the base morphemes of words determine the category of the whole word,

feature percolation has been expanded to account for other syntactic phenomena. For

example, Cowper's (1987) definition is given in (99).

(98)   [To whom] did you write the letter? *(Pied-piping)*

---

[52] Recall that Spellout objects are delimited by bolded square brackets: **[…]**Spellout object.

(99)   In a structure [$_α$ β γ] or [$_α$ γ β] …, α a projection of β, features from γ will percolate to

        α iff β is not specified for those features.   (Cowper 1987: 324)

I will employ a version of the feature percolation mechanism which allows unsatisfied

features to percolate from a functional head up the tree to any other functional head not

specified for those features: satisfaction of those features on the highest head will result in

immediate satisfaction of the features on the lower heads. In this case (100), when C is

Merged and the Spellout object is spelled out, C Agrees remotely with π in the Workspace,

and this element supplies a specification for the features of T and *v*, *now appearing on C*,

thus satisfying these feature slots all the way down the tree.

(100)   [$_{CP}$ C$_{EPP: \, π + uφ: \, π}$ [$_{TP}$ T$_{EPP: \, π + uφ: \, π}$ [$_{vP}$ $v_{uφ: \, π}$ [$_{VP}$ V DP]]]]

        This scenario raises a question: If we can delay satisfaction of *v* and T's features via

Remote Agree until the Spellout of the syntactic object, why can't we also delay satisfaction

of those features *via E-Merge* of DP? This would obviously result in the wrong word order:

DP would then overtly show up in [Spec, CP], the highest head in the Spellout object to

which the features had percolated. Overt DPs require Case, and Spec CP is not a Case-

assigning position in English; thus, delaying E-Merge of DP until Spellout of the Spellout

object would result in the DP not receiving Case. The fact that π is a special kind of syntactic

object allowed to remain in the Workspace makes satisfaction of the Spellout Object's

unsatisfied features via percolation possible: π does not have an overt counterpart in the

structure.

        While I employ the feature percolation mechanism here, I am aware that some have

argued (Cable 2010) that feature percolation is an unnecessary addition to the list of

grammatical operations, particularly within a Minimalist framework that attempts to reduce

the list of core operations, and that, indeed, it is only useful in explaining pied-piping. This is in fact not the case; I am employing the mechanism here for arbitrary control. While I share the concern of unnecessarily cluttering the inventory of syntactic operations, the fact that percolation seems to be needed in other areas of the grammar (here, for one), and also for some complex morphological constructions makes me comfortable employing it here. I leave these concerns to future research.

4.1.3 Subjectness of (Arbitrary) Control

The last-resort approach to Remote Agree with Workspace π derives an interesting fact about the distribution of NOC vs OC: that arbitrary control always appears in subject position. This is true of control, generally, and—somewhat surprisingly—is a fact that many established control theories do not account for in a non-stipulative way. (101) below shows the contrast between a licit (obligatory) controlled subject position (101a) and an illicit (obligatory) controlled object position (101b). The controlled position can be the external argument, but not the object, of an active clause.

(101) a. Richard tried [PRO to catch Molly]

  b. *Molly tried [(Richard) to catch PRO]   (Landau 2013: 108)

Why should a property that is so fundamental to the distribution of control not be adequately explained by the major control theories? Landau (2013: 108), in his sketch of "various theoretical mechanisms employed to guarantee the subjecthood of PRO" admits that "these devices are, for the most part, unprincipled". Most early proposals derive this property of control from various stipulations that PRO is primarily associated with [Spec, T/I] or subject position in the clause. In the null case approach (Chomsky and Lasnik 1993, Martin 2001), for example, it is arbitrarily stipulated that only nonfinite T/I can assign null case; conveniently, PRO is just the sort of syntactic object—in fact the only syntactic object—that

needs null Case. The even earlier GB accounts were no better. As discussed in Section 2.1, according to the PRO Theorem, PRO could not appear in governed positions due to its being [+pronominal, +anaphoric] (Chomsky 1981: 191), and nonfinite [Spec, IP] was the only (or one of the only) ungoverned positions. But this is hardly explanatory: PRO was the only element specified as both [+anaphoric] and [+pronominal], and it was so specified in order to make the requisite facts fall out.

Landau also admits that his own early Agree approach to control also fails to derive the subjectness of PRO: In that model, "the relevant feature on the clausal head (I or C) is [-R] (indicating "referential deficiency") and PRO is the only NP that bears this feature, and […] must be the closest one to I—hence, a subject" (Landau 2013: 109). Again, PRO is designated as featurally unique, and its syntactic environment (nonfinite Spec T/I) designated as the perfectly unique counterpart to fit. Like the preceding accounts, this account is *ad hoc*.

The movement theory of control (MTC) is one of the few previous theories of control (the other being the predicational theory of control) that does in fact account for the subjecthood of OC PRO in a less stipulative way (Landau 2013: 109–110). The theory of non-obligatory control that I advance here, involving Remote Agree, is consistent with the MTC. As for my own account of NOC, the reason that arbitrary control cannot appear in object position is also due to locality. (It's locality all the way down!) In the Movement Theory of Control (MTC), Landau (2013: 109) summarizes that it is the locality of movement operations that derives the essential "subjectness" of obligatory control: "In the [MTC], the subjecthood [of] PRO (ultimately a trace) is reduced to whatever explains the subjecthood of raising targets […] no NP lower than the subject would be able to move across it to the matrix clause. Thus the only possible movement out of a clause is that of its subject." Under this view, the fact that the syntactic context for obligatory control (the

controlled position) is always a subject position reduces to locality constraints on movement. Below I derive the subjecthood of arbitrary control from locality considerations, and the last-resort property of the Remote Agree analysis.

Like obligatory control, arbitrary control also can't appear in object position. The examples below are given first with the traditional notation (PRO$_{arb}$), then showing the (partial) derivations in my analysis. Note that for derivations using Remote Agree and feature percolation, as described in Sections 4.1.1-4.1.2, I adopt certain notational abbreviations for convenience. I will use the notation **(π)** to indicate where a head's features are satisfied by Remote Agree with Workspace π (thus abstracting away from the particular feature sets). **(π)** will be shown in the specifier position of the head whose features are satisfied via Remote Agree. Empty square brackets, **[ ]**, will appear on the head whose features percolate upward to be satisfied on the higher head (**[π]** indicating satisfied by π). (See the original notation, full notation, and new notation in (102) below.)

Note the difference between (102) and (103). In (102) there is an extraposed passive controlled clause with arbitrary control. I will put aside the extraposition movement and focus on the derivation of arbitrary control in the Spellout object which is the infinitival clause. In this clause, there is no theta-marked external argument slot, and thus only one set of argument features (uφ) to be satisfied via Remote Agree. The uφ features on the main verb, V, and the EPP features on T, have all percolated up to the C head since they were not satisfied by E-Merge of an overt DP (there being no available DP in the Workspace). This percolation is indicated in full in (102b) and abbreviated with empty subscript square brackets in (102c). Remote π, represented as **(π)** in (102c), then Agrees with C, thus satisfying the features on these heads (shown in [102b]). Once these features have Agreed with Remote π at C, and the Spellout Object (CP) has been spelled out, the result is an

arbitrary, generic interpretation for the "confrontees" in the clause. Thus passive controlled clauses do not pose a problem for this analysis when they receive an arbitrary interpretation.

(102)   a.   It's unfortunate [PRO$_{arb}$ to be confronted t$_{PROarb}$].

Full Notation (Remote Agree analysis):

b.   [It's unfortunate… [C $_{EPP:\ \pi\ +\ u\varphi:\ \pi}$ T $_{EPP:\ \pi}$ v_{pass} V$_{\boldsymbol{u\varphi:\ \pi}}$]]

Alternative Notation (Remote Agree analysis):

c.   [It's unfortunate …. [$_{CP}$ $\boldsymbol{(\pi)}$ C$_{[\boldsymbol{\pi}]}$ [$_{TP}$ T$_{-fin\ [\boldsymbol{\pi}]}$ [$_{vP}$ v_{pass} V$_{[\boldsymbol{\pi}]}$]]]

In (103) below, on the other hand, we have an active transitive infinitival clause without any overt DPs, where arbitrary control appears in both accusative object position and in subject position, Spec/T$_{[-finite]}$. (This is indicated in the traditional notation in (103a) for clarity.) This creates a problem: there are now two linked theta-positions in the same Spellout object. The sentence is repeated in (103b) with the abbreviated Remote Agree notation.

(103)   a.   *It's unfortunate [PRO$_{arb}$ to confront PRO$_{arb}$.]

Alternative Notation (Remote Agree Hypothesis):

b. [It's unfortunate …. [$_{CP}$ $\boldsymbol{(\pi)}$ C [$_{TP}$ T$_{-fin\ [\boldsymbol{\pi}]}$ [$_{vP}$ $v_{[\boldsymbol{\pi}]}$ V$_{[\boldsymbol{\pi}]}$]]]

Since we have adopted the MTC for obligatory control, chains linking multiple theta features are not a problem in and of themselves: indeed, we must allow a DP to receive a theta-role in the infinitival clause as well as the matrix clause in which the infinitival is embedded for such an analysis to work. But multiple theta features within the same Spellout object are another matter. I propose a reflexivization rule for A-chains that link multiple argument positions *within the same Spellout object.* I posit this rule in (104): within the same Spellout

object, a chain linking two theta positions will result in the lower being realized overtly as a reflexive.[53] Such a rule would take a sentence like (103) and realize it as (105).

(104)   Where an A-chain links two theta-positions within the same Spellout object, the lower position is realized overtly as a reflexive.

(105)   It's unfortunate [to confront oneself].

To return to the traditional arbitrary control analysis: if arbitrary control is analyzed as a null PRO$_{arb}$ pronoun, there is no non-stipulative reason why it couldn't appear in object position in an infinitival transitive active clause. The Remote Agree approach to arbitrary control gives us an explanation. Just as the subjecthood of obligatory control is the result of locality constraints on movement under the MTC, under the Remote Agree approach (which is compatible with the MTC), arbitrary control is the result of the locality constraints on Agree: specifically, that Agree into the Workspace is less local than Agree + Internal Merge within the same syntactic object as the probe. Thus arbitrary control will only arise if the probes, v and T, have first searched the syntactic object, found no active DP, then searched the Workspace and found no active goal (DP) there. At the point when the syntactic object is spelled out and C is Merged, and all the features that have percolated up C, then as a last resort C can Agree with Remote π, satisfying all of these features in one fell swoop. Crucially, though, this only happens if there is no active Agree option for the probe(s) within the same syntactic tree as the probe. Any active DP in the structure will be the closest thing to satisfy the probe in the tree, closer than anything in the Workspace, like π (and DPs in the Workspace will be "closer" than π). Thus the subjecthood of *arbitrary* control is derived by the locality restrictions on Search / Agree.

---

[53] Note that this also accounts for example (102), where reflexivization does not obtain, since there is only one theta position in that sentence as the verb is passive.

I've claimed that arbitrary control results from a last resort Agree mechanism at the Spellout of a Spellout object. I've connected this to the availability, or lack, of DPs in the Workspace to satisfy unresolved features on the heads within this Spellout object. This ensures that arbitrary control only arises in Spellout objects where there were fewer DPs available in the Workspace than argument slots, and also predicts that arbitrary control ("PRO$_{arb}$") will always include the highest nominal argument in the clause. These properties hold because arbitrary control results from a last resort strategy, only in the case where there *are* no available DPs to be moved into the subject position of the clause.

4.1.4 Arbitrary Control in Multiple Clauses

That arbitrary control is always the highest argument in the Spellout object, as a result of C Agreeing with Remote π, can be seen particularly in cases of disjoint "arbitrary" reference in multi-clausal structures, where each clause has a different arbitrary "controller" or interpretation. Within a given Spellout object, there will only be one instance of arbitrary control ("PRO$_{arb}$" or Agree with π by C); there cannot be more than one within a given Spellout object because Remote Agree with π is always the last resort option at the end of the Spellout domain. Furthermore, C at the edge of the Spellout object will always Agree with π when this constituent is spelled out (it does not have the option of waiting for a higher C). Remote π is a permanent feature of the Workspace, but there may be one instance of Agree with this universal Workspace entity *per Spellout object*. These individual instances of Remote Agree with π, one per Spellout domain, might not necessarily co-refer.

What we expect, and what we find, is that the reference of "arbitrarily controlled" clauses can vary independently of each other where they are separate Spellout objects, but not within the same Spellout object. In (106), the argument of *seem* and *done the reading* is

the same generic entity, and *[to seem to have done the reading]* is one Spellout object. (I

have used the more familiar notation with PRO_{arb} in [106a] and my new notation in [106b-

c].) Since *seem* is a raising verb, we can assume there is no CP boundary between the two

PRO_{arb} positions, and feature percolation can proceed all the way to the C head in the second

clause (106). Likewise, in raising-to-object (ECM) constructions (106d), percolation can

happen all the way up to the higher clause, assuming this is all one Spellout domain.

(106)   a. It's smart $[_{CP}$ PRO_{arb} to seem t_{PROarb} to have done the reading$]_{Spellout\ object}$

b. It's smart $[_{CP}$ **(π)** C $[to_{[]}$ seem $[to_{[]}$ have done$_{[]}$ the reading$]]]_{Spellout\ object}$

c. It's smart $[_{CP}$ **(π)** C $[to_{[\pi]}$ seem $[to_{[\pi]}$ have done$_{[\pi]}$ the reading$]]]_{Spellout\ object}$

d. It's good $[_{CP}$ **(π)** C $[to_{[\pi]}$ be believed$_{[\pi]}$ $[to_{[\pi]}$ have done$_{[\pi]}$ the reading$]]]_{Spellout\ object}$

In (107), the interpretation is such that those pleading guilty are not necessarily the

same as those doing the thinking.[54] In (107), the two instances of Remote Agree do not occur

in the same Spellout object but in different ones (I have underlined and bolded the most

embedded Spellout object for clarity).

(107)   a. It's naïve $[PRO_{arb}$ to think [that **$\underline{[PRO_{arb}\ to\ plead\ guilty]_{Spellout\ object}}$** is wise$]]_{Spellout\ object}$

b. It's naïve $[(\pi_1)$ C $to_{[]}$ think$_{[]}$ $[_{CP}$ that $\underline{[_{CP}\ (\pi_2)\ C\ to_{[]}\ plead_{[]}\ guilty]_{Spellout\ object}}$ is

wise$]]_{Spellout\ object}$

c. It's naïve $[(\pi_1)$ C $to_{[\pi1]}$ think$_{[\pi1]}$ $[_{CP}$ that $\underline{[_{CP}\ (\pi_2)\ C\ to_{[\pi2]}\ plead_{[\pi2]}\ guilty]_{Spellout\ object}}$ is

wise$]]_{Spellout\ object}$

These examples, with the Spellout objects demarcated, show that arbitrary control always

appears as the highest "argument" in the Spellout object/domain. In (106a–c), the verbs *seem*

and *done* are connected via feature percolation across two clauses; only one instance of

---

[54] For example, in the context where the opinion expressed is that it's naïve *for lawyers* to think that
for defendants to plead guilty is wise.

Remote Agree with Workspace π occurs at C. In (107), there are two instances of Remote Agree in separate Spellout objects, resulting in an interpretation where there can be two distinct sets of individuals interpreted (i.e. $\boldsymbol{\pi_1}$ versus $\boldsymbol{\pi_2}$). However, these two different generic sets of individuals could also be the same (i.e. the thinkers could also be the pleaders, as in *It's naïve for one to think that for one to plead guilty is wise*.) This is akin to long-distance control, and the mechanism we use to explain this co-reference will be the same. We turn to long-distance control next.

4.2 Long-distance Control

So far we have mostly focused on arbitrary control and the hypothesis that it is established via Remote Agree with π, a permanent feature of the Workspace. We now turn to the other sub-type of non-obligatory control (NOC), long-distance control. Recall that long-distance control (LD-control) appears in the same contexts as arbitrary control: adjuncts (extraposed clauses), subject clauses and other syntactic contexts that are opaque to movement.

　　　While NOC and OC appear in complementary syntactic configurations, as we saw above, LD control and arbitrary control (both of which are subtypes of NOC) appear in the same syntactic contexts, suggesting that they should have a unified account.

(108)　a. **[**(PRO$_{arb}$) to perjure oneself$_{arb}$/herself$_i$**]**$_{\text{Spellout Object}}$ would help Mary$_i$

　　　b. We$_j$ thought that **[[**PRO$_{j/i/arb}$ to perjure ourselves$_j$/herself$_i$/oneself$_{arb}$**]**$_{\text{Spellout Object}}$ would help Mary$_i$.**]** *(Intraposition)*

　　　c. We$_j$ thought that [it would help Mary$_i$ **[**PRO$_{j/i/arb}$ to perjure ourselves$_j$/herself$_i$/oneself$_{arb}$**]**$_{\text{Spellout object}}$]. *(Extraposition)*

The subject position of the subject clause in (108a) could either be arbitrary control or LD control—that is, in the traditional nomenclature, it could be a PRO controlled by *Mary*, a long-distance DP in another Spellout Object. In (108b), which is an intraposition

construction, PRO can be uncontrolled (i..e PRO$_{arb}$ interpretation, again in the traditional

nomenclature), or controlled long-distance by *Mary* or *we,* both antecedents that are in

different Spellout domains from the subject being controlled. In (108c), the control clause

appears in an extraposed position, with an expletive in the canonical subject position. As in

the intraposition case, PRO (to use the traditional terminology for clarity) can either be

arbitrary or can be controlled long-distance by either *Mary* or *we*, both in different Spellout

objects from the infinitival clause (i.e. the context for LD-control). The distributional overlap

established in (108) between arbitrary control and long-distance control cases suggests that

they are established by the same mechanism.

Note that the matrix verb in (108), *help*, is a non-psychological predicate. The

distinction between non-psych verbs and psych verbs is argued to be important by Landau

(2001), important enough to affect grammaticality judgements of those sentences. I have

already presented experimental evidence in Section 2.4.4 from Motut and Grant (2017) that

this distinction is not as robust as Landau claims.

I have suggested that in Spellout objects, but not in smaller domains, arbitrary control

arises via a percolation mechanism that carries unsatisfied features up to the highest head in

the Spellout domain (C for clauses), which Agrees with Workspace π. This Agree operation

can happen *at most* once per Spellout object. While in Section 4.1.4 we saw examples of

clauses with multiple arbitrary "PRO" interpretations, arising from Agree with Workspace π

more than once, the two instances of Agree were always in separate Spellout objects.

Arbitrary control is thus a single-clause phenomenon, whereas LD control is a multiple

clause—and multiple Spellout domain—phenomenon. Where LD control is established, the

apparent antecedent / controller is always in a separate Spellout object or Spellout domain.

Essentially, LD control involves associating an overt DP antecedent in the tree with the

Remote π element. If this DP had been available to satisfy features on v or T in the same Spellout object where arbitrary control is expressed (i.e. the infinitival clause), it would have been Agreed with and Merged there, taking precedence over the last-resort operation of Remote Agree with π.

The involvement of Remote Agree with Workspace π in both arbitrary control and LD control would explain the identical distributions of these two related phenomena. I hypothesize that a long-distance control relation is established in two stages: first, an Agree relation is established with Workspace π, and second, a relationship is established between π and the DP antecedent in the other Spellout object. Workspace π, however, does *not* need an antecedent. We can see this from the way it functions in the arbitrary control examples. It saturates features on functional heads in the tree but does not itself require an antecedent; it is not needy. So the question arises, how and why does π establish a relation with the higher DP in the second Spellout object/domain? What does this relationship look like? Is it syntactic, semantic, or perhaps pragmatic?

I had earlier hypothesized that this second-stage relation between a DP and Workspace π might be another Remote Agree operation. But this approach seems unlikely to work. The DP is itself a Spellout object with its own phi features, and like π is not needy. As shown in (109), there is no requirement that the embedded subject be controlled.

(109)   We thought that it would help Mary [if Bill left].

Rather than an Agree or Search-based relation, it makes more sense to propose a less purely syntactic relation for this second step. In fact, it need not be a syntactic relation at all; a semantic / pragmatic one will do. Landau (2013: 237–244) provides an extensive review of the ways in which purely "structural" accounts of NOC have failed—mostly because attempting to come up with a purely *syntactic* analysis for the distribution of NOC

undergenerates the relevant data. I return to some of these arguments in Section 4.2.3. First, let us consider how pronoun reference is established.

Pronouns, unlike anaphors, can be bound or free. When they are free (i.e. not c-commanded by an appropriate antecedent), pronoun reference can be established via coindexation. *Coindexation*, as pointed out by Grodzinsky and Reinhart (1993), "is ambiguous in the standard binding theory, representing both coreference and bound variable interpretation," but they develop, following Reinhart (1991), a definition for coreference as distinct from anaphor binding. This definition of coreference applies to NPs/DPs within the same syntactic tree or, crucially, even between NPs/DPs in distinct sentences:

> "To assign a pronoun a reference mentioned in a previous sentence, it is clearly not necessary, and usually not assumed, that the two are syntactically coindexed. There is no principled reason why things should be different when a pronoun is assigned a reference mentioned in the same sentence." (Grodzinsky and Reinhart 1993: 77)

This definition of coreference is useful here, since long-distance control often involves a DP in one Spellout object, and a separately spelled out control clause which has formed a link with π in the Workspace. These Spellout objects, in cases known as discourse control, need not even be in the same utterance. For instance, in the example of discourse control in (110) below, or the example of LD control in Swedish in (111) from Lyngfelt (2000: 32), the "antecedent" of the controlled clause is not in the same utterance as the controlled clause itself.

(110)   Deciding honesty was in my best interest, I admitted I took the cash from the till. The prosecutor seemed stunned. Clearly, [(PRO) confessing my crime] was not something they anticipated. (Adapted from Landau 2013: 232)

(111) [Tiger Woods]$_i$ var i praktiken borträknad fran segerstriden, efter tva "manskliga" inledningsronder. Men [efter att PRO$_i$ igar ha tangerat banrekordet vagar], ingen räkna bort [golfens nye "Golden Boy"]$_i$

"Tiger Woods was in practice dismissed from the winning competition, after two 'human' starting rounds. But [after PRO$_i$ having touched the record for the course yesterday], no one dares to disregard golf's new 'Golden Boy'."

These examples, where the LD antecedent and the controlled clause "PRO" are not linked within the same syntactic utterance / tree, lends credence to the idea that the coreference that links Remote π and the DP antecedent in LD control is not the result of a purely syntactic operation. I propose that this secondary link is not syntactic in the narrow sense but an operation of coreference that semantically links the LD antecedent with the controlled clause (and Workspace π).

How would long-distance control be established along these lines in some of the examples we've already described? Below I walk through a sample derivation in detail. In (112a) we see an example of arbitrary control in a subject clause, in the standard notation, and in (112b) in the Remote Agree notation. (Recall that **(π)** is indicated in the specifier of a head (C) that has Agreed remotely with π in the Workspace, and that "**[π]**" is the subscript notation indicating satisfaction of feature(s) via percolation from the head that Agreed with π.) In (112c) we have long-distance control, with an antecedent relation established between *Mary* in the matrix clause and the embedded control clause.

(112) a. **[(PRO$_{arb}$) to perjure oneself$_{arb}$]$_{Spellout\ object}$** would help Mary$_i$]
　　　　b. [$_{CP}$ **(π)** C$_{[\pi]}$ to$_{[\pi]}$ $v_{[\pi]}$ perjure$_{[\pi]}$ oneself]$_{Spellout\ object}$ would help Mary]
　　　　c. [$_{CP}$ **(π_i)** C$_{[\pi]}$ to$_{[\pi]}$ $v_{[\pi]}$ perjure$_{[\pi]}$ **herself$_i$**]$_{Spellout\ object}$ would help **Mary$_i$**]

We have seen examples like (112a) earlier. A derivation of (112c), long-distance control, is shown in (113).

(113) Derivation of (112c) above:

Step 1: [$_{CP}$ (π) C **to**$_{[\pi]}$ $v_{[\pi]}$ **perjure**$_{[\pi]}$ (π)]$_{SO\ Object}$ Remote Agree by C with π

Step 2: [$_{CP}$ …..would help Mary]   Construction of Clause 2

Step 3: [[$_{CP}$ (π) C to$_{[\pi]}$ $v_{[\pi]}$ perjure$_{[\pi]}$ (π)] would help Mary]   Merge

Step 4: [[$_{CP}$ (π_i) C to$_{[\pi]}$ $v_{[\pi]}$ perjure$_{[\pi]}$ (π_i)] would help **Mary$_i$**]   Coreference

Step 5: [[$_{CP}$ (π_i) C to$_{[\pi]}$ $v_{[\pi]}$ perjure$_{[\pi]}$ **herself$_i$**] would help **Mary$_i$**] Reflexivization

The first step is that the C of the embedded (control) CP Agrees remotely with Workspace π in the same manner as if this were an arbitrarily controlled clause. π provides the feature specification that percolates down to the heads with unsatisfied features, namely *v* and T, as well as V. Since this is a transitive clause (with no DP object available in the Workspace, presumably), we end up with a chain that links multiple theta positions (within the same Spellout object) via this percolation. This will bring us back to the reflexivization rule in a moment.

The matrix clause is constructed at Step 2 (*... would help Mary*) and Merged with the embedded (control) clause at Step 3. Now "arbitrary control" or Workspace π, and *Mary* are linked via the coreference relation (Step 4). This operation merely designates that Mary and whatever it is coindexed (*i*) with—here, Workspace π —refer to the same entity. By virtue of this coreference between the "long-distance antecedent" (*Mary* in this case) and the control clause, π receives the referential index associated with *Mary*. This leads to the "arbitrarily controlled" embedded "subject" being interpreted not as arbitrary, but as controlled by Mary. The final step (Step 5) is for the lower position in this chain to be realized phonetically as a reflexive pronoun coreferential with *Mary* (as per [104]).

This indexing of π deserves some comment. The Workspace element π must be capable of being indexed but it does not *require* an index. When it is coindexed with *Mary* in a different clause, the controlled clause is then interpreted as referring to *Mary*. But when it

occurs in an instance of arbitrary control, π does not refer to an entity or have a referential index. By default, then, π is not indexed, but it can be coindexed if the operation coreference is applied between it and a DP with more specified features.

In (113), the object position of the controlled clause, whose features are satisfied by percolation of π from C, is realized by a reflexive pronoun according to (104). Since the percolated features form a link between two theta positions within the same clause, the lower position is realized as a reflexive. Remote Agree with Workspace π alone would result in the phonetic form *oneself*, but since the features in this controlled clause are coreferential with *Mary*, identified from the matrix clause, the felicitous phonetic realization will be *herself*.[55]

Since the referential link between Workspace π (and the argument slots it percolates to) and the long-distance antecedent in another Spellout domain is established via coreference, long-distance control is thus not properly *syntactic* control. This is in line with proposals by Landau (2000, 2001, 2013, *inter alia*) and others (Kuno 1975b), which I will discuss more below, which argue that logophoricity and other pragmatic factors are a crucial aspect in determining the distribution of NOC. However, at the same time, since long-distance control overlaps in the initial stage with the mechanism of arbitrary control—Remote Agree with π—my analysis is also *partially* syntactic. The Agree relation between C and π is syntactic, and this relation forms the first step of any long-distance control relation. This might help explain why in the history of non-obligatory control there have been mixed results with purely structural or syntactic approaches.

---

[55] Some interesting questions might be raised at this point about the ordering of reflexivization and coreference as it pertains to Spellout and the semantic and phonetic realization of the reflexive pronoun. But I leave this for another time.

Let's look at another example. See (114) below (108b repeated). (114a) shows arbitrary control: (114b) is long-distance control, where the controlled clause is either controlled by *Mary* or *we*. In (114) this is presented in the old (PRO) notation, and in (115) - (116) below that in the notation I have adopted for my analysis.

(114)   a. We$_j$ thought that [[PRO$_{arb}$ to perjure oneself]$_{Spellout\ object}$ would help Mary$_i$.]

        b. We$_j$ thought that [[PRO to perjure ourselves$_j$/herself$_i$]$_{Spellout\ object}$ would help Mary$_i$.]

In examples (115) and (116) we see that the infinitival clause is the same for both (115), (116), but that the interpretation of this clause is coindexed with a different DP ("antecedent") in a different Spellout object in each example. The (a) examples show the pre-final step of coindexation and the (b) examples show the final step of reflexivization.

(115)   a. We$_j$ thought that [[$_{CP}$ **(π_i)** T$_{[\pi]}$ $v_{[\pi]}$ perjure$_{[\pi]}$ **(π_i)**] would help **Mary$_i$**.]

        b. We$_j$ thought that [[$_{CP}$ **(π_i)** T$_{[\pi]}$ $v_{[\pi]}$ perjure$_{[\pi]}$ **herself$_i$**] would help **Mary$_i$**.]

(116)   a. **We$_j$** thought that [[$_{CP}$ **(π_j)** T$_{[\pi]}$ $v_{[\pi]}$ perjure$_{[\pi]}$ **(π_j)**] would help Mary$_i$.]

        b. **We$_j$** thought that [[$_{CP}$ **(π_j)** T$_{[\pi]}$ $v_{[\pi]}$ perjure$_{[\pi]}$ **ourselves$_j$**] would help Mary$_i$.]

In both of these examples, whether the controller is *Mary* or *we*, the first step is Remote Agree with Workspace π. After the infinitival clause is spelled out and Merged into the larger structure, either *Mary* or *we* may be coindexed with Workspace π / the controlled clause. But while *Mary* and *we* are both (long-distance) options for the controller in this sentence, it's not true that just any DP in the sentence that is in a different Spellout object / domain can be the antecedent. This raises the question, how free *is* the selection of a possible antecedent to the controlled clause in LD control?

This has been a central concern of structural or syntactic accounts of long-distance and non-obligatory control in the literature. Indeed, Landau (2013) reports that many previous accounts failed because they predicted generalizations for the possible controller for

LD control that were too strict—LD control is in fact much more permissive than, say, Grinder's (1970) intervention formulation (seen in Section 2.4.2).

Proposals like Grinder's, which define the available LD antecedents over typical syntactic dependencies occurring within a single syntactic tree, cannot account for LD control in examples like (110) and (111) which we saw earlier—discourse control and intrasentential control. While using coreference rather than a syntactic dependency (Agree, Move, etc.) helps us avoid many of the pitfalls of the purely syntactic accounts, we might wonder whether they are in fact too permissive. Can *any* DP be coindexed with Workspace π / the controlled clause?

4.2.1 The Coreference Rule: What can be a long-distance antecedent?

A related problem is that coindexation is also how coreference is established between pronouns and their antecedents; however, it has been argued that "NOC PRO does not behave like a pronoun" (Landau 2013: 243). I have argued that coindexation is the mechanism that links the DP antecedent with the percolated features in the control clause, where Agree with Workspace π has taken place. Although in my account, there is no PRO in the structure—the LD antecedent DP is coreferential with π, but π is itself not located in the tree but in the Workspace—what I propose is that π, like a pronoun, can be assigned a referential index. (Or rather, it is the representation of π at the semantic level which is assigned the index.) If the mechanism for establishing LD control is the same as the mechanism for coreference in pronouns, then it might be reasonable to expect the distribution of pronouns with respect to a given LD antecedent to overlap with the distribution of non-obligatory control.

Examples like (117), originally from Grinder (1970: 306), have been adduced as evidence that "NOC PRO" does not behave like a pronoun (Landau 2013: 243).

(117)   [The girl who Max$_i$ loved] said it would be difficult for him$_i$/*PRO$_i$ to excuse himself from the party before midnight.

(118)   [His$_i$/*PRO$_i$ having already shaved] shows that Mary arrived more than three minutes after John$_i$ did.

In (117), DP *Max*, which is inside the DP *the girl who Max loved*, does not c-command PRO and yet, unlike pronominal coreference (i.e. *him*) cannot be coreferential with PRO (or the controlled clause). Pronominal coreference between *Max* and *him* obtains, but NOC is impossible. First, a technical quibble: This data might actually be less of a problem for my theory than it is for pronominal coreference theories of control that in fact project an empty category in the embedded subject position. In Landau (2013), these examples are presented as counter-evidence to Hornstein & Boeckx's theory of NOC control as *pro* (see Section 2.4.1). In their analysis, the PRO in (117) and (118) would be replaced by *pro.* However, in my analysis, there is no pronominal element in this position at all. The features on the T head are satisfied by Workspace π after percolation to C. A direct comparison here is moot.

Landau's claim about this data is that it demonstrates that "NOC PRO is more restricted than a pronoun" (Landau 2013: 243). For two reasons, this concern does not apply to my argument for using coindexation to establish NOC. The first I mentioned above: there is no PRO in subject position. The second reason is that PRO and *him* may not be in the same syntactic position at all in (117). For this data to show that NOC PRO does not behave like a pronoun, we must assume that *him* appears in the subject position of the embedded clause. However, Berman (1973, 1974) argues that this may not the case, and that *for him* in fact appears in the higher clause, where it is the experiencer of *difficult* and controls the clause

below. If this is true, then the argument from this data for NOC not overlapping with the distribution of a pronoun falls apart.

For these reasons, the data in (117) – (118) does not pose a problem for my coindexation account of NOC.

4.2.2 What Else Can Be an "Intervenor"?

Let's step back a moment and consider how this account compares to other accounts of NOC. Much of the syntactic literature on approaches to NOC focuses on what can be an "intervenor" for long (or longer) distance control. The goal of these proposals is to use locality to account for the distribution of possible antecedents to the controlled clause. Since my own analysis of NOC is only a partially syntactic analysis rather than fully syntactic (arbitrary control, within a single syntactic object, is established through a syntactic Remote Agree mechanism, while long-distance control is established non-syntactically via coreference with a DP in another syntactic object), determining what kinds of elements can be "intervenors" is irrelevant. However, my account of NOC is not entirely *non*-syntactic, since the crucial first stage of the relation between the antecedent DP and the controlled clause is established by the syntax, with access to the Workspace and Remote π. I will return to a discussion of these larger issues in Chapter 5.

Although my analysis makes the question of "intervenors" moot, it is useful to take note of the data adduced in such accounts. I also discuss below how my own analysis compares to previous syntactic or structural accounts of NOC. The following discussion is organized around the different types of intervention that are identified in previous works. There are three subtypes: DPs within DPs, experiencers in Super-Equi, and the availability of arbitrary control in the presence of an available antecedent.

Type 1: DP within DP

The first type can be found as early as Chomsky (1981). Landau (2013: 244) notes of this type: "While closer arguments need not intervene for more distant ones, they do generally intervene for DPs embedded inside them (Chomsky 1981)". He notes that this type of "intervention" appears to be more robust than the experiencer type of intervenor that I will discuss further below.

(119) *John$_i$'s friends think it is illegal [PRO$_i$ to feed himself]   (Chomsky 1986: 125)

(120) *[PRO$_i$ to have to feed himself] would annoy John$_i$'s friends.   (Chomsky 1986: 128)

In (119) and (120), John cannot be coindexed with PRO: the controlled clause cannot refer to John. (Compare to: *John's friends think it is illegal to feed themselves* and *To have to feed themselves would annoy John's friends*.) This pattern seems very syntactic. While DP *John's friends* is available to be coindexed with the controlled clause, the DP *John* inside this constituent is unavailable. But if, as in my analysis, non-syntactic coindexation is allowed between the controlled clause and a higher DP referent (or a DP referent in another Spellout object), then we need to consider why it is not possible here. Why would structure seem to be playing a role in these cases? One possible explanation is that the DP *John* is simply buried too deeply inside another, already spelled out constituent to be available for coindexation with the control clause. Another reason might have to do with the relative semantic-pragmatic salience of *John* as a DP embedded within a subject DP. I return to semantic-pragmatic factors in Section 4.3.

Type 2: Experiencer intervenors

Experiencers have played an outsized role in the history of accounts of NOC, particularly in structural accounts of NOC, and particularly for Super-Equi constructions. The general observation, going back to Landau (1999), is that experiencers are intervenors for long-

distance control, while non-experiencer DPs are not. This pattern was described for Super-Equi by Landau (1999, 2000, 2001):

(121) a.   Mary thought that it pleased John [PRO to speak his/*her mind.]

     b.   Mary thought that it helped John [PRO to speak his/her mind.]

     c.   Mary thought that [PRO to speak his/her mind] would please John.

     d.   Mary thought that [PRO to speak his/her mind] would help John.

The judgements in (121) are those in Landau (1999, 2000, 2001), but the ungrammaticality of (121a) with coindexation between Mary and PRO has been challenged by Motut and Grant (2017). I will first describe the generalization that Landau argued for and then explain the revised judgements. Landau's claim was that in extraposed clauses (i.e. [121a,b]) but not elsewhere, an experiencer DP argument selected by a matrix psychological predicate blocks long-distance control. Non-experiencer DPs, or those selected by non-psychological predicates, do not have this blocking effect. (121a) – (121b) are extraposition constructions and (121c) – (121d) are not. In (121b) and (121d), the control clause (with unrealized subject position) can be controlled by *John* or *Mary*, since neither of these arguments is an experiencer (the matrix verb *help* does not select an experiencer). In (121a), the controlled clause can be coreferential with experiencer DP *John*, but not with DP *Mary.*

     John is in fact closer to the controlled clause than *Mary* is in (121a), both in terms of linear order and in terms of hierarchical structure. This makes an intervention analysis tempting, and if the judgements were in fact as Landau reports them, we might not have a problem. In Landau's analysis, the difference between the ungrammatical and grammatical interpretations of examples like (121a) is due to the nature of psychological predicates' argument selection. This was discussed in some detail in Section 2.4.3.

Landau (2001) not only provided an analysis of this experiencer / non-experiencer pattern, but he was the first to describe it. Identifying the pattern itself is one of his major contributions. However, the judgements on this type of intervention were never very clear, in contrast to the Type 1 intervention effects discussed above. Landau himself acknowledges some variation in speakers' judgements in LD control in Super-Equi extraposition contructions. For example, he notes that the grammaticality of LD-control across an experiencer intervenor improves when a modal is added to the intermediate clause (e.g. *Mary thought that it would please John to speak her mind),* and that sentences of this type also improve for some speakers with an intonational pause before the controlled clause (Landau 2001: 113, fn.1, 114). Although he cites several examples from languages other than English, these examples are all personal communications and only single examples are provided.

As summarized in Section 2.4.4, Motut and Grant (2017) argue against the judgements provided in Landau (2001). We conducted a naturalness ratings and eye-tracking study that compared the conditions in (121a) and (121b), extraposition constructions with psych predicates versus non-psych predicates, and found no evidence for a categorical ban on long-distance control with a psycholgical matrix predicate, as Landau's description of the data predicts.

These findings by Motut and Grant are consistent with the account of NOC proposed here, where LD control results from coreference between the controlled clause and either DP in the extraposition cases. One crucial difference is that for Landau, the examples in (121) above, where *John* is an intervenor in (121a) and (121b), involve obligatory control rather than non-obligatory control. In my view, control by *John* in these cases is in fact long-distance control as well, since *John* appears in a separate Spellout object from the controlled clause.

My proposal for analyzing NOC with Remote Agree is thus more consistent with the revised judgements reported by Motut and Grant (2017): there is nothing absolute about experiencers blocking long-distance control. In Section 4.3, I will review what semantic-pragmatic factors *do* limit the possible DP antecedent in LD control.

Type 3: Arbitrary control in the presence of a potential controller

In another type of locality effect discussed in the literature, arbitrary control appears despite the availability of a possible "local" antecedent. Landau (2013) points out that this is a problem for many of the structural accounts as well as for optimality-theoretic approaches to NOC (e.g., Lyngfelt 1999, 2000). Structural accounts of the distribution of NOC versus OC, i.e. those that treat their relative distribution as a matter of locality, run into this challenge: "The possibility of arbitrary control in the presence of a (semantically appropriate) potential local controller is problematic to many accounts, which treat control in terms of 'search domains' […]" (Landau 2013: 240, fn. 10).

Examples (122) and (123) demonstrate this point. Again, like the extraposition examples above, the assumptions about locality are similarly misleading. *Bill* in (122) and *babies* in (123) are assumed to be "local" to PRO$_{arb}$, the empty subject position of the control clause. Accounts like that of Lebeaux (1984) in which arbitrary control is established by an operator at the edge of the matrix clause have to explain why *Bill* or *babies* does not intervene in these cases and force local control. In contrast, in my analysis, arbitrary control is the "local" default option for a given Spellout object; anything beyond that Spellout object is "long-distance" control. Therefore, the DPs *Bill* and *babies* in these examples are in fact non-local long-distance antecedents, available for coindexing as described above, and illustrated in (124), but not "blocking" arbitrary control and forcing OC.

(122) It would help Bill [PRO$_{arb}$ to behave (oneself) in public]   Manzini (1983)

(123)   It is dangerous for babies [PRO$_{arb}$ to smoke around them]   Kawasaki (1993)

(124)   a. It would help Bill$_i$ [(PRO$_i$) to behave himself$_i$ in public].

        b. It is dangerous for babies$_i$ [(PRO$_i$) to smoke].

If arbitrary control is Remote Agree with π, occurring as a last resort within a Spellout object if no DP is present to satisfy unvalued argument features, then I predict that arbitrary control is available in (123), along with the possibility of long-distance control via coindexation, as in (124). These arguments, *Bill* and *babies*, are in fact not local when one considers constraints on movement, either. It is not possible to form a movement chain between the position of PRO$_{arb}$ in (122) and *Bill*: PRO$_{arb}$ appears in an extraposed clause and is the highest argument in its Spellout object.

4.2.3 Criticisms of "Structural" Accounts of NOC

The fact that arbitrary control can appear in embedded clauses where a "semantically appropriate" DP antecedent appears in the matrix clause or elsewhere in the tree raises problems for "structural" accounts of NOC, as Landau (2013: 240, fn. 10) points out. Many attempts have been made to account for the distribution of non-obligatory control in purely syntactic terms (e.g., Grinder 1970, Lebeaux 1984, Manzini and Roussou 2000). Landau (2013: 237–244) provides a summary of these cases and argues that they all, ultimately, fail. The reasons are several, but one of the most pertinent criticisms is that they simply do not fully account for the data.

As Landau (2013) neatly summarizes, most previous structural accounts fail by being too restrictive, ruling out grammatical sentences. Indeed, individuals' judgements for NOC appear to be more permissive than originally thought. Grinder (1970), who was concerned with defining what could be an "intervenor" along the "deletion" (control) path, from the antecedent to the deleted element (i.e. PRO), argued that the closer of two DPs on the left of

the control clause would control PRO, and that a DP on the left of the control clause would block a DP on the right from being the antecedent.

In Grinder's formulation, an intervenor would block deletion by another DP if it commanded the deletion site, and one of two conditions held, both of which have to do with precedence: either the intervenor preceded the deletion site and the other DP followed it; or, both the intervenor and the competing DP preceded the deletion site but the intervenor was closer to it.[56] Some of Grinder's predictions are borne out, but counterexamples were subsequently raised in the literature.

Let's look at some data, following Landau's (2013) summary. For example, contrasting (125a) with (126a-b), Landau reports that (125) is ungrammatical (as predicted by Grinder because *Roxanne* is closer to the control clause than *Eric*), while (126a-b) are grammatical, which would not have been predicted by Grinder's intervention constraint. Grinder's constraint wrongly predicts that *Mary* would intervene in (126a) and block *John* from controlling PRO, contrary to the reported judgement. In (126b), *Sue* or *John* would be predicted to intervene and block *Mary* from controlling PRO, again contrary to the reported judgement. Landau (2013: 240) notes that Chierchia and Jacobson (1986: fn. 6) say that LD control is "merely disfavoured (due to 'perceptual' factors'), not strictly forbidden".

(125)   *$Eric_i$ said that Roxanne knew that it would be difficult [PRO_i to criticize himself]

(126)   a. $John_i$ thought that Mary said that [PRO_i shaving $himself_i$] would not be difficult.
               (Lebeaux 1984, cited in Landau 2013: 240)
         b. $Mary_i$ thought that John said that [PRO_i shaving $herself_i$] would bother Sue.
               (Chierchia and Jacobson 1986, cited in Landau 2013: 240)

---

[56] This is a very slight simplification.

Examples like (126) are often found to be grammatical, according to Landau, where the possible intevenor, *Mary*, is "not a co-argument of the infinitive."

There is also data contradicting Grinder's constraint in examples where potential intervenors appear to the right of the control clause. (127) is predicted to be ungrammatical by Landau (2000) (since *Pete* is an experiencer) as well as by Grinder (1970). In contrast, however, examples like (128) show that "intervention on the right side [is not] an absolute constraint" (Landau 2013: 240, who cites Clements 1975, Kuno 1975b, and Chierchia and Jacobson 1986 on this point).

(127)   *[That [PRO$_i$ washing herself with liquid oxygen] disturbed Pete] surprised Eileen$_i$.

    (Landau 2013: 239)

(128)   [[That [PRO$_i$ exiling himself] might grieve the Queen] never occurred to [the

    minister]$_i$. (Clements 1975, cited in Landau 2013: 240)

(129)   [[That PRO$_i$ losing the race] would upset everyone so much] surprised Sam$_i$

    (Landau 2013: 240)

In (128), *the Queen* does *not* intervene and block *the minister* from controlling PRO, as Grinder's structural constraint would predict; in (129), *Sam* can LD-control the controlled clause without being blocked by *everyone*. Grinder's formulation of an intervention constraint that would explain the apparent "distance" effects of NOC or LD-control therefore fails on empirical grounds. It is too restrictive, ruling out data that is grammatical (if somewhat variably so).

Grinder's (1970) syntactic account of the constraints on antecedence of PRO (for Grinder, not PRO but NP) was thus ultimately disproven. Lebeaux (1984) presents a different structural account, which also addresses the distribution of arbitrary control. For Lebeaux, PRO is an A$'$ anaphor that is bound by an operator at the edge of the clause: "Arbitrary PRO

is conceived of as being bound by an operator in an A′ position associated with the dominating S′ ” (Lebeaux 1984: 263). He also suggests that it might be conceivable that instead of an A′ operator binding PRO, an implicit argument might be doing this binding, but argues that the operator analysis is more plausible.

In Lebeaux's analysis, there is no distinction between PRO_{arb} and OC PRO or NOC PRO. This is characteristic of "search" or structural accounts of control, which unify OC and NOC, as local and non-local instances of the same operator. Indeed, as in my own analysis, for Lebeaux (1984) arbitrary control is the resulting "default" option when a local antecedent cannot be found. Arbitrary control arises when an operator is inserted to bind PRO in the absence of a local antecedent:

> "We would like to account for the fact that an arbitrary control reading is in general only available when a lexical antecedent is not present within the local domain. This may be done by formalizing the intuition that PRO searches for an antecedent within its local domain, and if no such antecedent is present, an operator is inserted to locally close off the sentence. This operator, in turn, may be restricted in reference by coindexing with some lexical antecedent (Chomsky, 1981) or may be left antecedentless, in which case its reference is free, and an arbitrary reading results (generally represented with a universal quantifier […]." (Lebeaux 1984: 264)

The analysis proposed here shares some insights with Lebeaux's. In both accounts, arbitrary control is the default interpretation assigned when a subjectless control clause cannot find a local antecedent, although what counts as a local domain is different for me than for Lebeaux. Also like Lebeaux, I argue that coindexation is the proper relation to establish long-distance antecedence. In other ways, however, my account is quite different. Lebeaux assumes that PRO exists; I, on the other hand, adopt a movement approach for OC, and my account of NOC involves satisfaction of features of a probe without an actual pronoun in

subject position. Lebeaux also tries to unify NOC and OC, while I argue that OC arises by movement and NOC by Remote Agree (plus coindexation).

Landau's criticisms of "structural accounts" extend to Lebeaux's account as well. Citing examples like (130), he observes that a c-command relationship is not required between PRO and its antecedent in LD-control. In these examples, there is no command relation between *Mr. Jones* and PRO in (a) or between *John* and PRO in (b).

(130) a. [**PRO** Perjuring **himself** like that in court] proves that **Mr. Jones** is an unreliable witness.
b. [**PRO** Storming out of the room that way after losing the game] convinced everyone that **John** is very immature

(Richardson 1986, cited in Landau 2013: 241)

Landau argues this is a "damaging" blow to Grinder's attempt at a purely structural description of NOC (Landau 2013: 241). If NOC were a syntactic phenomenon, and the connection between NOC PRO and its antecedent a purely syntactic relation, then we would expect examples like (130a–b) to be ungrammatical since we would not expect a syntactic relation to hold without c-command. Landau is right that this points to NOC being established by a non-syntactic relation (like coindexation) or involving non-syntactic processes or factors.

In Lebeaux's analysis, the null operator is inserted at the matrix clause to bind PRO if no local controller is found. This null operator can then be coindexed with an antecedent elsewhere in the sentence, as already summarized above. However, Lebeaux does have to introduce another rule into his analysis to rule out examples like (118), repeated below as (131). He stipulates that "O_i [the null operator] may not c-command its antecedent" (Lebeaux 1984: 267).

(131)   His / *PRO having shaved already shows that Mary arrived more than 5 minutes after

      John did.

But Landau calls this rule "way too strong": he is correct in that it rules out examples like

those in (130) above, where the DP antecedent of NOC "PRO" does not c-command it but

nevertheless the sentence is grammatical. Landau points out that this stipulation would also

rule out sentences like (132) below, which are grammatical.

(132)   [PRO$_i$ shaving himself] helped John$_i$.

So it seems that neither Lebeaux's nor Grinder's structural formulation of NOC really holds.

These are by no means the only two structural accounts, but Landau (2013) argues that none

of the structural accounts of NOC are adequate in accounting for the data. Most are too

restrictive. And although there are tantalizing signs that NOC might be syntactic—the mirror

opposite distributions of NOC and OC, the pseudo-blocking effects, OC's clear restriction to

a local domain—formulating an analysis of NOC that treats this phenomenon as *purely*

syntactic has failed. Landau puts it like this:

> "Although it seems that closer and more prominent NPs are favoured as controllers
> over more distant and less prominent ones, it proves very difficult to formalize a
> precise notion of intervention in NOC. Other than local experiencer co-arguments,
> that force OC, no other NP in the vicinity of a NOC clause can be said to be an
> absolute intervenor." (Landau 2013: 244)

But even Landau's "absolute intervenor", the experiencer argument in extraposition

constructions, does not give rise to, as he says, "absolute intervention." Motut and Grant

(2017) demonstrate that experiencers are not absolute intervenors—in fact, judgements about

these data are more consistent, as I decribed above, with the general availability of LD

antecedents, with the felicity of an individual antecedent likely due to non-syntactic factors

like context and semantics.

So *purely* structural accounts of NOC seem beset by problems from all sides. However, I did say earlier that I would be "resurrecting" the structural or syntactic accounts of control, to some degree. It is now time to clarify what I mean by this.

The structural or syntactic accounts of control by and large seek to unify OC and NOC by explaining the distribution of local versus long-distance antecedence in terms of syntactic considerations, primarily locality. On the face of it this makes sense: for various other phenomena in language, where there is a long-distance and local version, we expect locality of syntactic relations to determine which holds in a given instance. But control, as noted above, does not seem amenable to a purely syntactic account. Previous accounts have been both too permissive and too strict. Some, such as Grinder (1970) have focused primarily on defining "intervention", which leaves aside how arbitrary control factors into the mix. Other accounts (e.g. Hornstein, Landau), however, argued that OC was a fundamentally different phenomenon from NOC. (For Hornstein, and for Boeckx et al. [2010], NOC wasn't even PRO but *pro*.) OC has received a number of syntactic treatments, many of them highly successful. NOC, on the other hand, has resisted analysis along purely syntactic and structural lines. It is these structural accounts of NOC that Landau (2013) argues have ultimately failed.

Earlier when I wrote of "resurrecting" the structural account of NOC, I was setting the stage for a syntactic analysis that began with different assumptions about the fundamental operations of the syntax. These assumptions, particularly the possibility of Remote Agree, are the core of the proposal of this thesis. With Remote Agree as an option in the grammar, which I argued for from a theoretical standpoint, the notion of what is a "syntactic" relation expands beyond a single syntactic tree. For previous syntactic accounts of NOC or arbitrary control, arbitrary control was an element (PRO in many cases) bound by a syntactic element

in the tree (e.g., Lebeaux's null operator). Indeed, given the syntactic options, this was the only way that arbitrary PRO could be bound. (Some accounts, of course, maintained that arbitrary PRO was not bound.) But having the arbitrary subject "bound" by an element in the tree would predict all kinds of syntactic restrictions that we don't see. Establishing (arbitrary) control via a relation established beyond the tree avoids many of these issues.

What the Remote Agree hypothesis opens up is the possibility of a syntactic account where arbitrary control is derived via a syntactic relation, Agree, while at the same time not fully participating in syntactic relations in the way it would if the covert element were in the tree. This explains not only the distributions of LD and arbitrary control, but also the nullness of the "binder" (π in the Workspace).

All of this said, my account *does* resurrect structural accounts to some degree. What I have presented here could be characterized as a *partially* syntactic account: arbitrary control is established via a decidedly syntactic mechanism—Agree with Workspace π—and while I've argued that LD control is established via (non-syntactic) coreference, the argument previously analyzed as PRO in LD control is saturated by the Remote Agree operation. This means that even LD control is partly a syntactic phenomenon. Furthermore, this overlap in the mechanisms that give rise to LD control and arbitrary control predicts the distributional overlap between LD control and arbitrary control.

While this account does not unite NOC and LD control with OC under a single Search operation within the syntactic tree, as Grinder (1970) and others sought to do,[57] my proposal does unify LD control and arbitrary control by using the same operation, Remote

---

[57] That is, long-distance control is not the less-local Search option by a single probe, or by PRO, compared to local or obligatory control.

Agree, to establish both of them, with coindexation relating the arbitrary subject to its antecedent (for LD control, plus coindexation).

There is, however, still a locality story in this analysis. The idea of locality that turns out to be relevant here is not the relative locality (e.g. like Relativized Minimality) but more a kind of absolute locality, in terms of whether a relation spans more than one syntactic object. Rather than comparing the relative distance of long-distance antecedents compared to other types of antecedents, the derivation compares the distances of goals from the probe relative to the boundaries of Spellout objects.

4.2.4 Interim Summary

Now we return to the question posed earlier: if coreference is the mechanism by which long-distance antecedents are linked with a control clause, what are the restrictions on which DPs can be coindexed?

I've argued that NOC can be analyzed by Remote Agree with π at the level of the Spellout object, with the antecedent determined via coindexation between the control clause and the higher DP (or DP in another Spellout object / domain). Since π (or its semantic representation, or NOC PRO in the literature) cannot be coindexed with just any DP antecedent, semantic-pragmatic factors must influence what can be felicitously coindexed with the control clause. In this respect, I follow Landau in arguing that NOC is not entirely syntactic, although crucially for me it is *partly* syntactic: arbitrary control is established via Remote Agree in the syntactic component of the grammar, and this operation feeds long-distance coindexing with the antecedent.

The idea that NOC PRO is sensitive to pragmatic factors seems obvious, but what do we actually *mean* by that? Which factors are the important ones? (I do not have much new to offer here but will summarize the factors identified in prevous literature.)

4.3 Non-structural Factors in NOC

There are two relevant non-structural factors relevant to NOC: logophoricity and topicality (Landau 2013). Logophoricity has been noted by many control researchers over the years as an important factor in the pragmatics of NOC. Kuno (1975b) was one of the first, but it has also been noted by Williams (1992), Landau (2000), and Landau (2013).

As Landau (2013: 245) summarizes, "logophoric pronouns and reflexives pick their antecedent on the basis of some notion of mental perspective […] In linguistic terms, logophoric antecedents are subjects and objects of mental verbs (*think, realize*), psychological predicate (*disturb, angry*), and communication verbs (*tell, hear*)."
The logophoric nature of NOC is illustrated by the following examples. Logophoricity also explains (a) the experiencer "blocking" effect; (b) an interesting overlap with picture anaphors; and, (c) aspects of the DP-within-DP blocking summarized earlier.

Let's consider some data:

(133) a. John said to Mary$_i$ that it would be easy [PRO$_i$ to prepare herself for the exam].
   b. *John said about Mary$_i$ that it would be easy [PRO$_i$ to prepare herself for the exam.] (Kuno 1975, cited in Landau 2013: 245)

(134) a. [PRO$_i$ having just arrived in town], the main hotel seemed to Bill$_i$ to be the best place to stay.
   b. *[PRO$_i$ having just arrived in town], the main hotel collapsed on Bill$_i$.
   c. [Having just arrived in town], the main hotel was a vision indeed.

   (Landau 2013: 246)

In (133), the small difference between "to" versus "about" changes DP *Mary* from the object

of a communication verb, and therefore a participant in the utterance whose mental

perspective is implied, to a theme ("about Mary") whose mental perspective is not involved,

and this makes the sentence ungrammatical where Mary controls PRO. A similar point is

demonstrated in (134), where *Bill* is a logophoric centre in (a) but only a theme in (b), where

Bill cannot control PRO.

Topicality is another factor that has been argued to be relevant (Landau 2013,

Kawasaki 1993, Adler 2006). Landau (2013: 251) cites examples from Kawasaki (1993) in

support of the idea that "the antecedent of NOC PRO must be the sentence topic—an entry in

the common ground of conversation, on which the current sentence makes a comment":

(135)   a. [After PRO_i collecting some money], a bank account was opened by the $landlord_i$.
       b. *[After PRO_i collecting some money], a bank account was opened by a
          $businessman_i$.

As Landau explains, in (135b), the indefinite NP *a businessman* introduces a new

discourse referent but the definite DP, *the landlord*, does not. Topics must be part of the

common ground, and are therefore old discourse information. They cannot introduce a new

discourse referent. The ungrammatical (135b), where *a businessman*, a new discourse

referent, controls PRO, contrasts with the grammatical (135a), where *the landlord* is definite

and can be more plausibly read as the sentence topic. This contrast is adduced as evidence for

the claim that the antecedent of NOC PRO must be a sentence topic.

The broader point is that the acceptability of NOC varies depending on the

accessibility of the potential controllers as a sentence topic (Kawasaki 1993, Landau 2013).

In examples like (136b), where the controller of PRO is plausibly an implicit agent (i.e. the

washer of the stairs), PRO is linked only pragmatically and not syntactically to the implicit agent of the passive. The judgements below are as reported by Landau (2013: 252).

(136) a. [The rain]$_i$ washed the stairs [before PRO$_{i/?arb}$ entering the basement]

b. The stairs were washed [before PRO$_{arb}$ entering the basement].

This analysis would, in fact, be quite in line with my own analysis, where "arbitrary control" is established at the control clause level but can be pragmatically linked with another referent in the sentence, as in (134a) or in the discourse (as in [134c] above). That is, the referent of an arbitrary "PRO" can be overtly represented in the sentence, or only in the common ground / discourse.

(137) is another example cited by Landau (2013: 253) of purported "sentence-topic" control of a PRO clause. There is an implied topic in the sentence—the persons who are pitching the tents—even though this participant is not grammatically represented in the sentence.

(137)   [After PRO pitching the tents], darkness fell quickly.

Topicality is also suggested to be relevant in (138), which comes from an article entirely about (James) Baker. Thus it is suggested that control of PRO in the first clause by Baker is allowed despite the availability of *top Bush advisors* as a potential controller, because Baker is obviously the topic of this sentence (based on the article context). (Control by *top Bush advisors* would also be available, as we would expect.)

(138)   [Since PRO returning to the White House in August], several top Bush advisors have

urged Baker to be the president's chief surrogate in explaining the administration's

economic recovery plan.   (Kawasaki 1993: 205)

On the one hand, these examples help us better define what pragmatic salience factors are involved in narrowing down which possible discourse or grammatical referents can be coindexed with the control clause. On the other hand, is topicality the crucial factor here?

It's not entirely clear that all of these examples demonstrate the importance of topicality as distinct from logophoricity and the [+HUMAN] requirement of NOC PRO discussed earlier. For starters, in some of them it turns out to be very difficult to get the intended contrasts, and the judgements are nuanced. Furthermore, as Landau points out, identifying topicality as a factor is confounded by its close relationship to logophoricity and the [+HUMAN] requirement on NOC, along with the relative lack of availability of relevant data. Landau attempts to disentangle these factors by constructing the example in (140).

(140) As for the boots$_i$, it was obvious [that [for them$_i$ / *PRO$_i$ to be produced in Italy] would increase their appeal.   (Landau 2013: 255)

(140) is constructed so that *the boots*, which has been fronted in the sentence, is the sentence topic, but it is not able to control PRO in the infinitival clause. But this sentence, as Landau points out, also violates the [+HUMAN] requirement on controllers, since *the boots* is inanimate. *The boots* is not a logophor, but neither is it [+HUMAN], so it is not obvious whether the lack of humanness or the lack of logophoricity is causing the problem in this example. The example in (141), from Richardson (1986), is constructed to try to tease apart the logophoricity requirement and the [+HUMAN] requirement.

(141)   All I can say about Mary$_i$ is that most people I have spoken with agree that while [PRO$_i$ removing herself from the race so quickly] may have pleased the party hacks, it will surely distress the people whose interests she represents.

Mary, which is both [+HUMAN] and the sentence topic, but crucially *not* a logophor (her mental perspective is not implicated since she is being spoken *about*), can still control PRO.

This suggests that "logophoricity is not a necessary condition on NOC, although the [+HUMAN] restriction is; which means that the latter is irreducible to the former, and constitutes a primitive feature of NOC PRO, as proposed in some of the earlier studies (Chomsky 1981, Rizzi 1986a)" (Landau 2013: 256–57).

Landau suggests that both logophoricity and topicality are sufficient, but not necessary, conditions, on top of the absolute ("primitive") requirement that NOC be [+HUMAN]. As Landau notes, this is a tentative conclusion based on limited data, and it is clear that more research needs to be done in this empirical terrain.[58] Despite the lack of clarity in the data, Landau tentatively proposes the generalization in (142) to capture the pragmatic properties of NOC (Landau 2013: 256).

(142)   *Pragmatic Properties of NOC:*
        In an NOC configuration [..DP… [PRO…]…] (order irrelevant), DP may control
        PRO iff DP is [+human] and either a logophoric center or topic-oriented.

The pragmatic factors discussed above, logophoricity and topicality, help to explain many of the effects noted earlier that in earlier discussion I left up to "salience" or "pragmatic" factors. Specifically, these pragmatic factors help us explain more clearly the sensitivity of NOC to experiencer "intervenors", the robust DP-within-DP "blocking" effect, and the parallelism between NOC and picture-anaphora, an observation dating back to Grinder (1970).

We had already noted that experiencers, the arguments of psychological predicates, feature prominently as NOC controllers. Landau (2001) argued that experiencers were absolute intervenors for LD control, forcing local obligatory control in extraposition

----

[58] I personally am not sure that I get an ungrammatical reading for control of *to be produced* by *the boots* in (140).

constructions. Motut and Grant (2017) challenged this conclusion with naturalness ratings data and eye-tracking / processing evidence. Their results support the idea that experiencers do not intervene categorically. Instead, there is a preference in naturalness for control by an experiencer, as well as a preference for a more local controller. But we found no evidence for categorical blocking by the experiencer arguments of psychological predicates. In fact, our findings were more consistent with a pragmatic or semantic preference for an NOC controller that is a logophor (of which experiencer arguments are a subtype) than with an experiencer effect due to a strict grammatical restriction. As Landau (2013: 247) says, "Grinder's intervention effects (except for the case of a local experiencer controller […] reflect the hierarchical embedding of mental perspectives in complex sentences. PRO in Super-Equi seeks an anchoring logophoric centre, and favours the one in whose perspective it is immediately embedded over those to which it is only indirectly related. This is just a preference, however […]." This "preference" is visible in (143) (example [126b] from earlier) and (144) below.

(143)   Mary$_i$ thought that John said that [PRO$_i$ shaving herself] would bother Sue.
(Cherchia and Jacobson 1986, cited in Landau 2013: 240)

(144)   *[That [PRO$_i$ washing herself with liquid oxygen] disturbed Pete] surprised Eileen$_i$.
(Landau 2013: 239; originally from Grinder 1970: 303)

In (143), Mary controls PRO "across" *John*. The reading where Sue controls PRO may also be available. Both *Sue* and *Mary* are logophoric centres in this sentence. For (144), the judgement reported in Landau (2013) is that *Eileen*, although an experiencer, cannot control PRO. There is a closer experiencer, *Pete*, which presumably interferes with control by *Eileen*. However, there is variability in these judgements. When set off with an intonational pause after *Pete*, control by *Eileen* in (144) improves. And in (143), we have already noted earlier

how Chierchia and Jacobson (1986: fn. 6) remark that long-distance control here is disfavoured but not forbidden. The variation in judgements reported in the literature, and the robust results from Motut and Grant's naturalness and eye-tracking studies, are consistent with the idea that possible LD controllers in NOC are quite free but ultimately affected by the pragmatic factor of logophoricity. The "experiencer effects" in NOC would follow from this logophoricity factor, and crucially are not indicative of a *grammatical* constraint.

Landau (2001, 2013), however, insists that in Super-Equi extraposition constructions, experiencer co-arguments remain absolute intervenors, forcing OC and blocking LD. If, as I have proposed, NOC is analyzed via Remote Agree, there is no way to achieve an absolute experiencer blocking effect. Fortunately, this more closely coincides with the data, particularly that reported by Motut and Grant (2017). In fact, my proposal is silent about experiencers, leaving this to the pragmatics.

Logophoricity and topicality also help explain the more robust DP-within-DP type of intervention mentioned in Section 4.2.2. These are examples like (119), repeated below as (145).

(145) *John$_i$'s friends think it is illegal [PRO$_i$ to feed himself]
(146) *[PRO$_i$ to have to feed himself] would annoy John$_i$'s friends.

<div align="right">(Chomsky 1986: 125,128; cited in Landau 2013: 244)</div>

In these examples we noted that it was strange that *John* could not be coindexed with PRO, given that coindexation is the mechanism responsible for establishing long-distance control. Why shouldn't *John* be able to control PRO here?

The answer, Landau suggests, is that a DP will be prevented from controlling PRO if it is contained inside another DP that introduces a new logophoric centre, but not if that "containing DP or NP" does *not* introduce a new logophoric centre. He posits that there is a

class of such NPs, which Landau (2000) calls "transparent nouns," related to but not identical with inalienably possessed nouns, which function as a kind of "logophoric extension" (Landau 2013: 249) of an individual (e.g. *career, development, reputation, image;* cf. *John's career, John's development, John's reputation, John's image).* This class of NPs will not block an individual-denoting DP contained inside them from controlling a control clause (or "PRO"), and the logophoricity factor helps us understand why. These NPs/DPs do not introduce a logophoric centre or a new mental perspective and so, pragmatically, the individual-denoting DP contained inside them is still salient and available to be coreferential with the control clause. Conversely, an NP that *does* introduce a new logophoric centre will block any DPs it contains from controlling.

This factor explains a subset of the intervention cases: for example, why *John* can control PRO in (147a) ([145] above) but not in (147b).

(147)   a. *John$_i$'s friends think it is illegal [PRO$_i$ to feed himself]   (Landau 2013: 248)
          b. John$_i$'s gratitude proved that it was important [PRO$_i$ to feed himself]. (Landau 2013: 248)

However, ultimately even this explanation falls short. In Landau's discussion of the extraposition examples, the choice of one logophoric centre over another is described as a "preference"; by the same logic, the difference between (147a) and (147b) should be a preference, not a grammatical contrast. In other words, why is *John* not also an available antecedent for the control clause in (147a), with *John's friends* merely preferred (rather than required) as the antecedent? With the right intonation and context, the acceptability of (147a) with *John* as the controller seems to improve significantly. The judgments for data of these types of constructions is by no means determined, and are variable enough that they should be tested experimentally rather than by simple introspection.

The treatment given experiencer intervenors in Super-Equi by Motut and Grant (2017) could serve as a model and might reveal more flexibility in these judgements than described in the literature. Given that this remains a problem for other researchers in this area, not just for me, I will leave it to further study.

Finally, as Landau (2013) notes (following Kuno [1975b]), the identified pragmatic factors allow us to capture an interesting overlap between NOC and picture anaphors which dates back at least to Grinder (1970), and has been repeatedly observed since then (Clements 1975, Lebeaux 1984, Landau 2013, *inter alia*). As shown in (148), when *Mary* is a logophoric centre, it can control the control clause, but not when it is a theme.

(148) a. John said to Mary$_i$ that there was a picture of herself$_i$ with a Mafia figure in the newspaper.

b. *John said about Mary$_i$ that there was a picture of herself$_i$ with a Mafia figure in the newspaper.

(Kuno 1975b, cited in Landau 2013: 245)

The examples in (148) above mirror those from (133), repeated below as (149). For both (148a) and (149a), where *Mary* is the addressee (and therefore a participant in the illocutionary act whose mental perspective is implicated), it can control PRO or bind the anaphor long-distance and the sentence is grammatical. For (148b) and (149b), on the other hand, *Mary* is the theme (she is being spoken *about*) whose mental perspective is therefore not implicated. *Mary* can neither control nor bind the anaphor in these examples.

(149) a. John said to Mary$_i$ that it would be easy [PRO$_i$ to prepare herself for the exam].

b. *John said about Mary$_i$ that it would be easy [PRO$_i$ to prepare herself for the exam].

These parallels are explained if we assume that both long-distance reflexives and NOC are sensitive to logophoricity.

We see that logophoricity and, to a lesser extent, topicality, help to explain aspects of NOC that are not fully accounted for by the syntax. But unlike Landau (2013), I would argue that NOC is best explained with an analysis that is at least *partly* structural, namely the Remote Agree analysis presented here. I leave the exact details of the pragmatics for further study, since solving the pragmatic and contextual nuances of NOC is beyond the scope of this thesis, which instead aims to illustrate how a promising theoretical mechanism in the grammar, Remote Agree, can be used to analyze this phenomenon.

My account of Remote Agree, to the degree that it explains the syntactic properties of NOC, is compatible with the pragmatic and contextual factors discussed above. We would expect these factors to influence the felicity of coindexing Remote π (or its denotation) with any given long-distance antecedent. The conclusion that Landau (2013: 255) reaches is that "logophoricity is not a necessary condition on NOC, although the [+human] restriction is; which means that the latter is irreducible to the former, and constitutes a primitive feature of NOC PRO, as proposed in some of the earlier studies (Chomsky 1981, Rizzi 1986)." What he means by this is that the [+HUMAN] restriction cannot be reduced to a logophoricity requirement on NOC PRO.

It is difficult to see in a non-stipulative way how the obligatory [+HUMAN] quality of both LD control and arbitrary control could be explained by a purely semantic account of NOC. As Landau notes, the [+HUMAN] requirement on NOC is not reducible to the pragmatics of logophoricity. Its fundamental nature could be explained, though, as an inherent property of Remote π itself. Indeed, [+HUMAN] seems like a very plausible minimal semantic denotation for the syntactic object we have called Workspace π. A bare π node

carries no additional features typical of pronouns (e.g. gender, number); what would the denotation of this most minimal and underspecified element be? Surely the plain quality of humanness might be enough. Indeed, Landau's description of [+HUMAN] as "primitive" to NOC brings to mind my description in Section 3.3 of Workspace π as a "primitive" of the syntactic Workspace. The [+HUMAN] interpretation of NOC seems to be an irreducible core fact about its interpretation: the Remote Agree analysis of NOC provides a syntactic entity in which to locate this property.

As much as logophoricity and topicality appear to be significant factors in determining where long-distance control can obtain, there is still a fundamentally syntactic component in the analysis I have advanced here for NOC. I turn next to reviewing the Remote Agree Hypothesis as a whole and in its application to NOC.

Chapter 5: Conclusion

In this Chapter, I provide a summary of the key findings in this thesis, and evaluate the Remote Agree Hypothesis along several lines.

First, I argued that Remote Agree is theoretically motivated, and then set out to explore its potential empirical contribution to the analysis of non-obligatory control, relative to other accounts in the literature. My goal was not to construct an exhaustive account of the phenomenon, but rather to provide an empirical test case for the Remote Agree Hypothesis. That is the first standard by which we will assess the proposal given here. In Chapter 3, I developed the Remote Agree Hypothesis in detail, adding assumptions and constraints on Spellout, and articulating the nature of the syntactic Workspace.

I also hoped to shed new light on the phenomena identified in Chapter 2, specifically the puzzles of projection and parallelism in construal phenomena like anaphora and control. These included the parallel distribution (in both the anaphor binding and control domains) between apparently unlicensed referentially dependent elements ("PRO" and unlicensed reflexives) and their long-distance counterparts (long-distance control and long-distance anaphor binding). That is, unlicensed and long-distance anaphors have almost identical distributions, as do arbitrary and long-distance control. Finally, the "unlicensed" nature of referentially dependent elements like PRO and reflexive pronouns was itself a puzzle.

I then concentrated on non-obligatory control, and provided an analysis via the Remote Agree Hypothesis of this (relatively understudied, compared with OC) phenomenon. Based on evidence from Motut & Grant (2017) for a new understanding of the data for LD-control, I proposed a partially structural analysis of NOC, succeeding where previous, fully structural accounts had failed. I then reviewed the pragmatic factors that others have argued

are essential to the distribution of NOC and discussed how the Remote Agree account is at least compatible with existing approaches to these factors. Below, I provide a summary of how the Remote Agree analysis fared with respect to other theories of NOC, bearing in mind that control was merely one test case for the Remote Agree Hypothesis.

I will briefly recapitulate each of these in turn, in reverse order from how they were presented above, starting with an evaluation of the empirical success of the Remote Agree approach, and then relative to the theoretical goals it was introduced to address. Finally, I will suggest some avenues for future work.

5.1 Evaluating the Present Proposal for NOC with Remote Agree

First, let's consider whether the proposal accounts for the characteristics of NOC laid out in Chapter 2. While I did not set out to argue for a new analysis of obligatory control, any satisfactory account of NOC, as I argued earlier, should account for the differences between NOC and OC. In Section 2.2, I articulated some of the benchmarks for an adequate theory of control, including that it explain the configurational properties, semantic properties, and the parallels between LD-control and arbitrary control. Furthermore, an adequate account of NOC, I argued, ought to explain the obligatory [+HUMAN] interpretation of arbitrary and LD control receives (which OC does not have).

In my analysis, Remote Agree with Workspace π is the mechanism by which arbitrary control is established. This part of the analysis is syntactic, using the syntactic operation of Agree, which Motut and Oda (2011) argued could be expanded beyond the tree-under-construction into the Workspace (without subseqent E-Merge). My account captures the overlap in distributions between long-distance control and arbitrary control because long-distance control, at the level of the infinitival ("control") clause, also involves Remote Agree.

At the level of the control clause, long-distance control and arbitrary control are identical. This account explains the parallel distribution of arbitrary control and LD-control: they both occur in subject clauses, moved clauses, and extraposed clauses. These are the environments where Remote Agree with π happens. Of course, in the second stage of the derivation for long-distance control, arbitrary and LD control diverge.

This echoes older insights (i.e., Grinder 1970, Lebeaux 1984) that arbitrary and long-distance control were of the same type, except that in the long-distance case, there was an additional step or linking with a more distant antecedent. In more recent accounts, on the other hand, (e.g., Landau 2013), non-obligatory control and arbitrary control, while classified together, are nonetheless analyzed using quite different theoretical primitives, PRO and PRO_{arb}. And Landau's (2013) and others' claim that NOC is a logophor whose interpretation is not handled by the syntax does not amount to an actual analysis of NOC. The Remote Agree account successfully captures the fact that that there is, in fact, a syntactic component to NOC.

The complementarity between NOC and OC is another question that I dwelt on in some detail. These two types of control appear in mirror opposite contexts, a characteristic which has been described in various previous theories; nevertheless, they are both still *control* phenomena, conditioned by the same syntactic factors (infinitival clauses, etc.). Few accounts have managed to capture both the connectedness of these two phenomena and their intrinsic oppositeness. Boeckx et al. (2010) captured the complementarity, but not the connectedness, by arguing that little *pro* is inserted in environments where the movement theory of control (MTC) is not possible. In some ways, my account is reminiscent of theirs: I adopt the idea of a last resort mechanism (in this case, Remote Agree) in island contexts (the context for NOC). In my analysis, however, one of the crucial factors is whether the

Workspace contains a DP available to E-Merge. Remote Agree with π, and therefore arbitrary control, arise in Spellout objects whose derivational workspace does not contain enough overt DPs to satisfy their argument requirements. Landau (2013), on the other hand, argues that NOC and OC are handled in completely separate modules of the grammar, and that NOC is completely non-syntactic. Landau (2015) acknowledges the lack of connection between NOC and OC in his earlier, Agree-based work, and his Two-Tiered Theory of control attempts to correct for some of these deficiencies. In this work, OC and NOC are not divided into distinct modules of the grammar; instead, a sub-type of OC, logophoric control, is argued to be very closely related to NOC. However, it is not clear that this account explains how NOC can arise in exactly those contexts where movement (or Agree) is disallowed (i.e. capturing the complementarity).

In my account, the New NOC/OC Generalization argued for in Section 4.1 captures the broad distributions of the obligatory versus non-obligatory phenomena in a new way, making it amenable to analysis by Remote Agree. Instead of the traditional categories of complements versus adjuncts, extraposed constituents and subject clauses, I characterize this distribution in terms of when the derivation has access to the Workspace, and therefore to π. It turns out that NOC arises in spelled out objects, while the context for OC is syntactic objects that have not been spelled out. This makes the implication of a Workspace element like π intriguing.

At the same time, I have related NOC and OC in that both have a truly syntactic component, as both involve the operation Agree, making use of a previously unexploited option available for the grammar (Remote Agree). Although I have not presented a new analysis of OC, I have assumed a movement account of OC, which consists of Agree-triggered (E-)Merge, followed by Internal Merge, of a single DP. Its counterpart is an *un-*

Merged Workspace element, π, which Agrees with the percolated features on C to satisfy

argument features in the subjectless infinitival clause.[59]

I have also provided an account that gives a syntactic grounding to the very robust

semantic property of NOC, its [+HUMAN] quality, which is lacking in some theories. Theories

that conflate OC and NOC too closely cannot explain why OC does not display this robust

[+HUMAN] requirement, while NOC does (e.g., Lebeaux 1984). On the other hand, theories in

which NOC and OC are treated as two completely distinct phenomena, for instance with OC

in the syntax and NOC in the semantics / pragmatics (e.g. Landau 2000, 2013 and others),

cannot account for the fact that the [+HUMAN] requirement is seemingly immutable, not

subject to the kinds of contextuality or gradience that other semantic or pragmatic factors,

like NOC's logophoricity or topicality, are.[60] The [+HUMAN] requirement on the

interpretation of NOC seems more likely to be a matter of grammaticality than of felicity: my

account predicts this by putting this requirement into the syntax, literally, in the form of π,

the universal primitive of the Workspace proposed in Chapter 3, with a [+HUMAN] feature

specification.

Even in Hornstein's and Boeckx et al.'s (2010) accounts, NOC would not be expected

to be necessarily [+HUMAN]: it is analyzed as *pro*, which is pronominal, but cross-

linguistically *pro* does not have a [+HUMAN] requirement. Their pronominal account also

faces other hurdles: by making NOC fully pronominal and placing it in the syntactic tree,

their account suffers from various empirical (and theoretical) problems. We reviewed some

---

[59] (Note that I have not necessarily committed myself to Hornstein's and and Boeckx et al.'s version
of the movement theory of control, but I believe my approach to NOC is largely compatible with it.)

[60] Landau (2015) does, however, account for the [+HUMAN] specification of NOC as well as
logophoric OC control.

of these in Sections 2.4 and 4.2.1; by and large these issues relate to the fact that overt pronominals (see, for example, [117]) do not have the same distribution as NOC "PRO." My account side-steps these challenges by locating this element beyond the tree-under-construction, in the Workspace, and making it available only at the point of Spellout. Putting this element in the Workspace also helps us explain the nullness of PRO, a point to which I return below.

The semi-syntactic nature of my account distinguishes it in a literature that is frequently divided on the syntactic versus semantic nature of control. One of the more novel aspects of my analysis is that the division of labour between syntax and semantics is parceled out in a new way. Many previous accounts have argued that OC is syntactic but NOC is not. In my analysis, both OC and NOC are syntactic, but NOC is only *partly* syntactic. It has at its core the syntactic operation of Remote Agree, which on its own accounts for so-called arbitrary control. However, in the long-distance cases the coindexation of the "arbitrarily controlled" clause and the antecedent in another Spellout object is accomplished in the semantics. This makes arbitrary control entirely syntactic, and long-distance control *semi*-syntactic.

Landau's (2000, 2001, 2013) accounts separate NOC and OC fairly categorically. He argues that OC is essentially syntactic (achieved via Agree), and NOC is a logophor arising in a particular syntactic context (a context he describes rather than explains), and that the rest of its properties are determined by the semantics or pragmatics. Landau (2015), which I have regrettably not had scope to treat in detail here, acknowledges the shortcomings of his earlier account, and one of those key shortcomings is the lack of relatedness between NOC and OC. His Two-Tiered Theory of Control helps to account for this, but this approach has other drawbacks. One of the advantages of my approach is that I present an analysis that

independently completes the theoretical picture painted in Chapter 3: Remote Agree fills a gap in the paradigm of Agree and Merge, eliminating the otherwise unexplained absence of (External) Agree without (External) Merge. This theoretical merit should be considered in evaluating the overall success of the proposal in its empirical applications.

My account also has the empirical advantage of being compatible with the most up-to-date grammaticality data for Super-Equi, an important sub-class of NOC data. I reported on the naturalness ratings study and eye-tracking studies in Motut and Grant (2017), where no evidence was found for the experiencer-blocking effect reported by Landau (2001). This removes an anomaly in the data: experiencers are no longer an exception to the generalization that there is relatively free availability of long-distance controllers (i.e., non-syntactically determined antecedents) that occur in a separate Spellout domain from the infinitival clause where Remote Agree with π has occurred. My analysis of NOC with Remote Agree is consistent with this pattern.

There are, however, some things that I didn't demonstrate in the domain of control phenomena with Remote Agree. Given that this was a test case for the Remote Agree mechanism, and not an attempt to solve all of the problems of syntactic control, or even non-obligatory control, some questions were set aside or addressed only cursorily. I did not discuss control in adjuncts, which Landau (2013) argues act variably like OC or NOC, though it is reasonable to suppose that this apparent variation could be related to other locality-related properties of those adjuncts. I did not present a new formulation or analysis of the pragmatic factors governing NOC. I also mentioned only cursorily the analysis of the Two-Tiered Theory of Control (Landau 2015). As mentioned earlier, this analysis divides the control landscape differently than the previous literature does, introducing a distinction between predicative and logophoric obligatory control. My account shares with Landau

(2015) the claim that PRO should be analyzed as a bound, minimal pronoun-*like* element, although in his account this element (pronoun) is rather less minimal than π (and I in fact claim π is not a pronoun). A full comparison with his account is beyond the current scope of this thesis.

5.2 Evaluating Remote Agree: Its theoretical merits

I have shown that Remote Agree can be fruitfully applied to an analysis of NOC, though many questions remain. I have thus demonstrated that the Remote Agree Hypothesis has empirical merit, which was one of the goals of this thesis. I have also more fully articulated some of the additional assumptions required to turn the initial idea of Remote Agree into a working theory.

I provided an explicit formulation of Remote Agree, expanding the original proposal from Motut and Oda (2011a,b), and arguing that Remote Agree with π should apply when the syntactic constituent under construction is spelled out. This preserves the original locality flavour of the Motut and Oda (2011a,b) proposal: a probe attempts to be satisfied by an element within the tree, and this will result in I-Merge, or movement, of a DP if it is already in the tree (e.g. control via movement). If there is no DP in the tree, the probe will Search the Workspace.[61] If the probe does not find a Workspace DP to Merge, then Agree with π is a last-resort mechanism at Spellout. The addition of π to the Workspace, and explicit rules about Spellout and the Workspace, were added to the original proposal. The "last resort at Spellout" nature of the Remote Agree operation departs from the original proposal in Motut and Oda (2011a,b).

---

[61] This is the source of the "Move-over-Merge" effect described in Motut & Oda (2011a,b): goals in the tree are more local than goals in the Workspace.

I pointed out in Chapter 1 that among the construal phenomena, there are some interesting parallels and puzzles into which I hoped Remote Agree might provide insight. I accounted for the parallelism between arbitrary control and long-distance control in the analysis I provided for NOC. I gave hints towards how reflexives might be analyzed within the framework I provided here, although I did not address the parallels in that domain in detail. Another of the puzzles identified in Chapter 1 was the "puzzle of projection": why (for anaphora and arbitrary control) can referentially dependent elements that usually have antecedents occur without visible antecedents? I believe I have answered this for arbitrary control, in a way that clearly relates it to its distributional cousin, long-distance control. Workspace π may be taken in some sense to be an antecedent, although as discussed it is not the antecedent of PRO.

This brings us back to two key issues: the existence (or non-existence) of PRO, and its nullness. These two questions bear on architectural and ontological issues, going well beyond the empirical coverage of control phenomena. Theories are divided on whether or not to posit the existence of a phonologically null PRO. Landau (2013) devotes an entire chapter to arguing *for* the existence of PRO. Hornstein (2001) and Boeckx and Hornstein (2004, 2006) argue for the merit of the MTC in large part based on the fact that it allows us to *eliminate* PRO and the control module from the grammar, because it reduces PRO to a trace, a theoretical element independenty needed for analyses of movement. Indeed, Landau (2015), in his detailed critique of his earlier work, claims that "it is no secret that PRO is a sore thumb in the eyes of many linguists," and ultimately argues that PRO is a minimal pronoun, the kind of element we may find elsewhere in the grammar. Manzini and Roussou (2000) eliminate PRO entirely, arguing that (obligatory) control is established where one DP "attracts" more than one predicate (where Attract is an operation of the grammar they posit).

My account also eliminates PRO, as I adopt a movement approach to obligatory

control that is complementary to my account of NOC, and my analysis of NOC allows

unsaturated features / argument positions to be satisfied directly via Agree with π in the

Workspace, without positing that the subject position of the infinitival is actually filled. In

most other theories, the subject position of the infinitival is either empty or not. If it's empty,

no argument is projected there and the position does not exist. If the position is filled, then

the element that fills it must be null (i.e. PRO). But my account allows for a kind of "in-

between" solution to this "puzzle of projection": there is no projected subject, and yet in

some ways π in the Workspace *is* the subject—it just remains in the Workspace rather than

Merging into the syntactic tree. Does this seem like trickery? If it does, it is well-motivated

trickery, as I argued earlier, filling a theoretical gap in the paradigm of Minimalist operations.

The introduction of π, which I argued for in Chapter 3, raises another question. I

earlier argued for the need to reduce the inventory of null categories in the grammar (see also

Saab 2014), and that ontologically it was best not to add empty categories if we did not

absolutely need to. Thus I have done away with PRO in NOC, which fits in nicely with this

goal, but I have on the other hand introduced π into the grammar. It is worth pausing here to

ask, is the elimination of one null category countered by the addition of another? Is the

balance sheet really any better, with respect to our ontology of grammatical empty

categories?

I would argue, yes, we are better off, for several reasons. PRO was a stipulated entity

assumed only for the one specific grammatical context in which it appears—its uniqueness

has been, as Landau said, a sore thumb for many. The element π is little more than a minimal

feature-like element whose existence ontologically is required by the already assumed

features of the grammar (i.e. we use it elsewhere, albeit usually with additional

specifications). There are also other places to explore its use, as I suggest below. The second

reason is that the nullness of π, as compared to the nullness of PRO, can be derived from its

position in the Workspace, rather than being stipulated. PRO is stipulated to be null because

the position we are trying to describe in the relevant constructions is null. π, on the other

hand, is null arguably because it remains in the Workspace and is not linearizable with

respect to any syntactic tree with which it is linked. Since it is unable to Merge into the tree,

or to be spelled out on its own, its lack of phonetic content could be easily reduced to its

being unlinearizable with respect to any other syntactic object. For these reasons, positing π

may be better than positing PRO.

5.3 Other Possible Applications of the Remote Agree Hypothesis

I focused on NOC as my test empirical case, but there are other domains where the

Remote Agree Hypothesis might be successfully applied. The remarks in this section are

tentative, but I hope they might serve as inspiration for the further exploration of this topic.

5.3.1 Control in/into DP

I argued that Remote Agree could apply as a last resort when a syntactic object is

spelled out. This entails that it will occur at the edge of spelled out or frozen domains or

constituents. While we focused in our test case on NOC, which occurs at the level of CP, it is

conceivable given the mechanics of the proposed system that Remote Agree might occur at

the edge of Spellout objects other than CPs. DP control might present one such example.

Control in DPs is relatively less studied than OC in clauses, and there is much

unresolved in the literature, not only about its proper analysis but also the basic facts. The

intuition that control might occur in DPs as well as in clauses is described nicely by Landau

(2013: 201): "First, in line with the old intuition that DP structure mirrors clausal structure,

the question arises whether DP can host a null subject; next, what type of category is it, and finally, how is control expressed inside nominals?"

Evidence that there may be an unexpressed subject in DPs comes from examples like (150) below from Roeper (1993), and (151). What (150) shows is that inserting a determiner into the NP in (150a) blocks the entailment that the falls from the airplane are *John*'s falls from the airplane. (150a) is meant to be an example of control into a nominal, and Roeper (1993) suggests that this shows that control into NPs is blocked by the addition of a definite determiner (150b). The depictive secondary predicate, *stoned*, in (151) additionally shows that there is an understood null subject in the nominal headed by *discussion*.

(150)   a. John enjoyed [falls from the airplane] → John falls

b. John enjoyed [the falls from the airplane] → somebody falls

(151)   [(PRO) Discussion of these issues *stoned*] rarely produces satisfactory results.

(Safir 1987, cited in Landau 2013: 205).

The following examples from Williams (1985) are furthermore argued to present a puzzle for control into nominals. *John* is the picture-taker in (152a) through (152c), but there is no nominal specifier (i.e. [Spec, NP]) available to host a PRO in (152b) or (152c), unless we assume a very expanded DP with multiple specifiers.

(152)   a. John took [a picture of Mary]

b. John took [Mary's picture]

c. John took [his first picture yesterday]

Landau (2013: 204) convincingly argues that the contrast in (150) should not be taken either way as evidence for control in (150a) or against it in (150b). This contrast, he argues, could be due to many things other than the presence of PRO, and I'm inclined to agree. The data from Williams (1985) could likewise be interpreted in a variety of ways. But these data

illustrate one of the key issues in the literature on control into nominals: it is less clear in the nominal domain than it is in the clausal domain that there is a position for a projected null subject. This has led to control into nominals being interpreted not only as control (i.e. PRO) but also as *pro* or as implicit arguments that are not projected in the syntax.

Remote Agree might help us solve this puzzle as well. We could assume that control into DPs functions much the same way as control into CPs does: unsatisfied argument features percolate to the highest head in the Spellout object (here, a DP or NP), and that head Agrees with π in the Workspace, creating an arbitrary interpretation, or to be linked with a DP in another Spellout object. This would make control into DPs like non-obligatory control into infinitival CPs.[62] We would neatly sidestep the issue of which syntactic projection of the nominal should host a null subject, since the Remote Agree (with π) mechanism enables an empty subject position to be interpreted without Merging any element into the tree.

If Remote Agree applies to Spellout objects as a last resort as they are spelled out, we might look beyond CPs to other types of Spellout constituents for more evidence of this mechanism at work, such as DPs. Furthermore, the expansion of this analysis to other kinds of constituents and Spellout domains highlights another possible argument in favour of the Remote Agree Hypothesis itself, beyond expansion of its empirical coverage.

Some analyses of control (Landau 2015, Manzini and Roussou 2000, Lebeaux 1984) propose that there is either additional structure or some kind of operator at the edge of the CP, that can either bind PRO or provide an interpretation to the subjectless clause. It is fair to

---

[62] Indeed, there seems to be evidence that control *inside* a nominal acts like OC while control from outside *into* a nominal looks more like NOC (Landau 2013). This generalization would be easier to capture across clauses and nominals together using a generalization defined over Spellout objects, as I have suggested for the OC/NOC distinction.

ask how the Remote Agree Hypothesis would be an improvement upon such analyses. Why, in other words, must we say π is in the Workspace rather than utilizing some similar element in the tree itself? Or, phrased differently, why not add π as a special kind of operator at the edge of infinitival CP domains, to provide a similar mechanism but without recourse to the Workspace? Is the expansion of syntax to include Remote Agree really necessary?

If constituents other than CP can be analyzed the same way, this might form the beginning of an argument against the CP-operator analysis. To make this concrete, suppose that nominal control could be analyzed similarly to clausal control using Remote Agree with π. Such an analysis would provide motivation for the existence of π, rather than a specialized operator that just happens to be selected by C and D. A concrete Workspace, hosting π, which can interact with probes in any Spellout object constructed there, could explain clausal-nominal parallels better than the stipulation that similar operators appear at these constituents' edge.

5.3.2 Implicit Arguments

Implicit arguments are a second domain in which the questionable status of those arguments' projection in the syntax opens the door to a possibly fruitful analysis using Remote Agree. Implicit arguments, like the unlicensed anaphors and arbitrary "PRO", are unlicensed in the sense that the argument structure of the verb requires there to be an argument, but that argument does not appear to be projected. For example, passive voice in (153a) includes an implicit agent in the matrix clause, which "controls" the PRO in the infinitival clause. The adjunct in (153b) has no overt argument but it is understood that the implicit agent (the seller) is the one (not) doing the reading. In the Italian null-object example

in (154), 'the people' may or may not be an overtly realized argument, but in both cases it is "they" who are doing the concluding.

(153)   a. The ship was sunk [PRO to collect the insurance].

       b. The books can be sold [without PRO reading them].

(Chomsky 1982 / Williams 1985, cited in Bhatt and Pancheva 2006)

(154)   Questo   conduce (a gente)   a [PRO concludere quanto segue]

       This   leads   (the people) to [conclude what follows].

       'This leads the people to conclude what follows'

(Rizzi 1986, cited in Bhatt and Pancheva 2006)

These examples use control as a diagnostic for the existence of an implicit controller. But here we run into some of the same problems which arose in the brief discussion of DP control above: Where are these implicit arguments projected? Are they projected? How can they control into clauses and adjuncts, as in (153) and (154), if they are not projected? Analyses of implicit arguments represent them in a variety of ways: as instances of *pro* or as phi bundles (Landau 2010). Some argue they are not projected in the syntax at all (Williams 1985, 1987; Baker, Johnson, and Roberts 1989) but are represented merely thematically or within a functional head. Many analyses of implicit arguments rely on diagnostics of syntactic activity such as binding and control, which assume that evidence of syntactic activity is synonymous with evidence for syntactic projection.

Remote Agree problematizes these diagnostics for the presence of implicit arguments by challenging the assumptions that control and binding necessarily implicate projected structure only. With Remote Agree as part of the grammar, the appearance of a subjectless constituent in need of a controller does not imply that a controller must be projected in the tree. If π were implicated in the analysis of these structures, we might arrive at syntactic

analyses (via Remote Agree) where clausal projection is not required. While an analysis of implicit argument phenomena is beyond the scope of this thesis, I suggest that a starting point to a Remote Agree analysis of implicit arguments might begin with (155) below.

(155) $[_{vP\,/\,pP\,/\,nP\,/aP} \ldots v/p/n/a \ldots]_{\textbf{Spellout object}}$ π

|_ __ __ ___ __ __ __ __ __ __|

Perhaps when an argument-selecting head can find no argument in its Numeration / Workspace to E-Merge, then at the point of Spellout, Agree with π becomes an option, resulting in an interpretation of an implicit argument for that argument-selecting head. Many questions present themselves, and refinements to the definitions of Spellout domains and other details would be necessary. But we might expect the combination of the features of the selecting head and the denotation of π, with some modification, to produce interesting results.

5.3.3 Speech-Act Phenomena

One final possible domain of application of the Remote Agree Hypothesis lies in the syntax of discourse participants. As Haddican (2018) and others note, there has been a revival of interest in Ross's (1970) proposal for placing the addressee and speaker participants of an utterance into the syntax concretely (e.g., Speas and Tenny 2003; Miyagawa 2012). Most recent accounts do not adopt Ross's (1970) original proposal wholesale, and there are various proposals, but one representation is given in (156), from Haddican (2018: 1):

(156)   *The neo-performative hypothesis:*
        [Speaker … [Addressee … [CP [TP]]]]

Speas and Tenny (2003) refer to this as the speech act layer of the syntactic tree, above CP. Miyagawa (2012) has used this proposal as the basis to analyze Basque allocutive agreement, as in (157), in which the –k/n suffixes on the auxiliary agree with the gender of the addressee.

(157)   Kotxea garestia izan-go d-u-**k/n**.

        car.abs expensive cop-irr expl-root-**2sg.fam.masc/fem**

        'The car is going to be expensive.'   (Haddican, 2018: 1)

Such evident syntactic marking suggests to some that addressee and speaker ought to be

projected in the syntax to license this agreement, despite the fact that speaker and addressee

are not arguments of the verb and are never realized overtly as "arguments" of a higher

speech-act projection. The addition of null categories "speaker" and "hearer" to the grammar

hopefully ignites in the reader the same scepticism that I hope to have elicited for the

existence of PRO. Similar questions arise: should these null entities be projected in the tree at

all?

I will leave a full investigation of these questions to future work, but I suggest that

this empirical territory is ripe for analysis with Remote Agree. Perhaps Workspace π is linked

pragmatically with speaker and hearer; or, perhaps speaker and hearer themselves are

minimal speech-act-linked pronominal elements universally available in the Workspace, to

interact syntactically with probes in the tree (such as the Basque auxiliaries). The Workspace

opens up possibilities for the existence of such universal, ontologically necessary categories

that seem to be both a part of the syntax and yet not projected in any tree.

5.4   The End, and Beyond

I have presented the Remote Agree Hypothesis and its elaboration, along with a proposal for

a Workspace element, π, which is available for a probe to Agree with as a last resort at the

edge of a Spellout object. I have articulated the theoretical motivation for such an operation

at length: it fills a gap in the paradigm of grammatical operations, adding Agree without E-

Merge to our system, in parallel to Agree without Move (I-Merge). This expansion is

justified by the conflation of I-Merge and E-Merge: the properties of the one ought to hold of

the other, and I argued for the expansion of I-Merge's properties to its external equivalent. I then applied Remote Agree to a test case, non-obligatory control (NOC), a puzzle of non-projection that resists a consensus analysis even more than its more worked-over counterpart, obligatory control. I successfully showed that the Remote Agree Hypothesis can account for many of NOC's configurational features, and argued for its advantages over previous accounts. Along the way, I also presented experimental data from Motut & Grant (2017) that simplified the grammatical distributions of a subtype of NOC, Super-Equi, removing an exceptional case of "blocking" from the NOC data in a way consistent with my predictions. I further suggested many prospective avenues for future empirical testing of the hypothesis.

The potential of this proposal far exceeds the empirical cases on which I have tested it here, but this fact in no way diminishes the significance of those first tests as necessary steps towards a fuller understanding of the implications of the hypothesis. Thus, the final word on Remote Agree remains to be written.

References

Abels, Klaus. 2012. *Phases: An Essay on Cyclicity in Syntax*. Berlin/Boston: Walter de Gruyter.

Adler, Alison. 2006. *Syntax and Discourse in the Acquisition of Adjunct Control*. Doctoral Dissertation, Massachusetts Institute of Technology.

Alboiu, Gabriela. 2007. Moving Forward with Romanian Backward Control and Raising. In *New Horizons in the Analysis of Control and Raising*, ed. William D. Davies and Stanley Dubinsky, 187–211. Dordrecht: Springer.

Bach, Emmon. 1979. Control in Montague Grammar. *Linguistic Inquiry* 10(4): 515–531.

Baker, Mark, Kyle Johnson, and Ian Roberts. 1989. Passive Arguments Raised. *Linguistic Inquiry* 20(2): 219–251.

Béjar, Susana. 2003. *Phi-Syntax: A Theory of Agreement*. Doctoral Dissertation, University of Toronto.

Béjar, Susana, and Milan Rezac. 2009. Cylic Agree. *Linguistic Inquiry* 40(1): 35–73.

Berman, Arlene. 1973. A Constraint on *tough* Movement. *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society.*

Berman, Arlene. 1974. *Adjectives and Adjective Complement Constructions in English*. Doctoral Dissertation, Harvard University.

Bhatt, Rajesh, and Roumyana Pancheva. 2006. Implicit Arguments. In *The Blackwell Companion to Syntax*, ed. Martin Everaert and Henk van Riemsdijk, 558–589. Oxford: Blackwell.

Boeckx, Cedric. 2009. On Long-Distance Agree. *Iberia* 1(1): 1–32.

Boeckx, Cedric, and Norbert Hornstein. 2004. Movement under Control. *Linguistic Inquiry* 35(3): 431–452.

Boeckx, Cedric, and Norbert Hornstein. 2006. The Virtues of Control as Movement. *Syntax* 9(2): 118–130.

Boeckx, Cedric, and Norbert Hornstein. 2007. On (Non-)Obligatory Control. in *New Horizons in the Analysis of Control and Raising*, ed. William D. Davies and Stanley Dubinsky, 251–262. Dordrecht: Springer.

Boeckx, Cedric, Norbert Hornstein, and Jairo Nunes. 2010. *Control as Movement*. Cambridge: Cambridge University Press.

Bowers, John. 1973. *Grammatical Relations*. Doctoral Dissertation, Massachusetts Institute of Technology. [Published by Garland, New York, 1986.]

Bowers, John. 1981. *The Theory of Grammatical Relations*. Ithaca, NY: Cornell University Press.

Bowers, John. 2008. On Reducing Control to Movement. *Syntax* 11(2): 125–143.

Borer, Hagit. 1989. Anaphoric AGR. In *The Null Subject Parameter*, ed. Osvaldo Jaeggli and Kenneth J. Safir, 69–109. Dordrecht: Kluwer.

Brame, Michael K. *Conjectures and Refutations in Syntax and Semantics*. Amsterdam: North-Holland.

Cable, Seth. 2010. Against the Existence of Pied-Piping: Evidence from Tlingit. *Linguistic Inquiry* 41(4): 563–594.

Chierchia, Gennaro. 1984. *Topics in the Syntax and Semantics of Infinitives and Gerunds*. Doctoral Dissertation, University of Massachusetts, Amherst.

Chierchia, Gennaro. 1989. Structured Meanings, Thematic Roles and Control. In *Properties, Types and Meanings II,* Ed. Gennaro Chierchia, Barbara Partee and Raymond Turner, 131–166. Dordrecht: Kluwer.

Chierchia, Gennaro, and Pauline Jacobson. 1986. Anaphoric Properties of Infinitives and Gerunds. In *Proceedings of NELS 16*, ed. S. Berman, J.-W. Choe, and J. McDonough, 57–74. Amherst, MA: GLSA.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Chomsky, Noam. 1975[1955]. *The Logical Structure of Linguistic Theory.* New York: Plenum Press.

Chomsky, Noam. 1977. On WH-Movement. In *Formal Syntax*, Ed. Peter W. Culicover, Thomas Wasow, and Adrian Akmajian, 71–132. New York: Academic Press.

Chomsky, Noam. 1980. On Cognitive Structures and their Development: A Reply to Piaget. In *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*. Cambridge, MA: Harvard University Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1982. Some Concepts and Consequences of the Theory of Government and Binding. Cambridge University Press: MIT.

Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use.* New York: Praeger.

Chomsky, Noam. 1991. Some Notes on Economy of Derivation and Representation. In *Principles and Parameters in Comparative Grammar*, ed. Robert Freidin, PAGES. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995a. *The Minimalist Program.* Cambridge, MA: MIT Press.

Chomsky, Noam. 1995b. Bare Phrase Structure. In *Government and Binding Theory and the Minimalist Program*, ed. Gert Webelhuth. Oxford: Blackwell.

Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In *Step by Step*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by Phase. In *Ken Hale: A Life in Language*, ed. Michael Kenstowic, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2004 [2001]. Beyond Explanatory Adequacy. In *Structures and Beyond*, ed. Adriana Belletti, 104–131. Oxford: Oxford University Press. (first published as MIT Occasional Papers in Linguistics 20).

Chomsky, Noam. 2005. Three Factors in Language Design. *Linguistic Inquiry* 36(1): 1–22.

Chomsky, Noam. 2007. Approaching UG from Below. Ms: Massachusetts Institute of Technology. in *Interfaces + Recursion = Language?*, ed. Uli Sauerland and Hans-Martin Gärtner, 1–29. New York: Mouton de Gruyter.

Chomsky, Noam. 2013. Problems of Projection. *Lingua* 130(1): 33–49.

Chomsky, Noam, and Howard Lasnik. 1993. The Theory of Principles and Parameters. In *Syntax: An International Handbook of Contemporary Research*, ed. Joachim Jacobs. Berlin: DeGruyter. [also published as Chapter 1 of Chomsky 1995a]

Christiansen, Morten H., and Nick Chater. Language as Shaped by the Brain. *Behavioral and Brain Sciences* 31(5): 489–509.

Christiansen, Morten H., and Nick Chater. *Creating Language: Integrating Evolution, Acquisition, and Processing.* Cambridge, MA: MIT Press.

Citko, Barbara. 2000. *Parallel Merge and the Syntax of Free Relatives.* Doctoral Dissertation, Stony Brook University.

Citko, Barbara. 2005. On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge. *Linguistic Inquiry* 36(4): 475–496.

Clements, George N. 1975. Super-Equi and the Intervention Constraint. In *Proceedings of the 5th Meeting of the North Eastern Linguistic Society*, 13–28. Amherst, MA: GLSA.

Collins, Chris. 2002. Eliminating Labels. In *Derivation and Explanation in the Mininalist Program*, ed. Samuel David Epstein and T. Daniel Seely, 42–64. Oxford: Blackwell.

Cowper, Elizabeth. 1987. Pied Piping, Feature Percolation, and the Structure of the Noun Phrase. *Canadian Journal of Linguistics* 32(4): 321–338.

Culicover, Peter W., and Ray Jackendoff. Control is not Movement. *Linguistic Inquiry* 32(3): 493–512.

Culicover, Peter W., and Ray Jackendoff. 2006. Turn Over Control to the Semantics! *Syntax* 9(2): 131–152.

Déchaine, Rose-Marie, and Martina Wiltschko. 2002. Decomposing Pronouns. *Linguistic Inquiry* 33(3): 409–442.

Dresher, B. Elan. 2018. Review of *Creating Language: Integrating Evolution, Acquisition, and Processing*, by Morten H. Christiansen and Nick Chater. *Canadian Journal of Linguistics* 63(3): 468–470.

Epstein, Samuel D., Hisatsugu Kitahara, and T. Daniel Seely. 2015. *Explorations in Maximizing Syntactic Minimization.* New York: Routledge.

Epstein, Samuel D., Hisatsugu Kitahara, and T. Daniel Seely. 2021.. *A Minimalist Theory of Simplest Merge*. New York: Routledge.

Fowlie, Meaghan. 2013. Multiple Multiple Spellout. In *Challenges to Linearization*, ed. Theresa Biberauer and Ian Roberts, 129–170. Berlin: Mouton de Gruyter.

Fox, Danny, and David Pesetsky. 2005. Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics* 31(1/2): 1–45.

van Gelderen, Elly. 2004. Economy, Innovation, and Prescriptivism: From Spec to Head and Head to Head. *Journal of Comparative Germanic Linguistics* 7(1): 59–98.

Grano, Thomas. 2012. *Control and Restructuring at the Syntax–Semantics Interface*. Doctoral Dissertation, University of Chicago.

Grinder, John T. Super Equi-NP Deletion. In *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*, 297–317. University of Chicago: Chicago Linguistic Society.

Grodzinsky, Yosef, and Tanya Reinhart. 1993. The Innateness of Binding and Coreference. *Linguistic Inquiry* 24(1): 69–102.

Haddican, Bill. The Syntax of Basque Allocutive Clitics. *Glossa: A Journal of General Linguistics* 3(1), p. 101.

Hicks, Glyn. 2009. *The Derivation of Anaphoric Relations*. Amsterdam: John Benjamins.

Hinzen, Wolfram. 2012. The Philosophical Significance of Universal Grammar. *Language Sciences* 34(5): 635–649.

Hornstein, Norbert. 1999. Movement and Control. *Linguistic Inquiry* 30(1): 69–96.

Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal.* Oxford: Blackwell.

Hornstein, Norbert. 2003. On Control. in *Minimalist Syntax*, ed. Randall Hendrick, 6–81. Oxford: Blackwell.

Hornstein, Norbert. 2007. Pronouns in Minimal Setting. In *The Copy Theory of Movement*, ed Norbert Corver and Jairo Nunes, 351–385. Amsterdam: John Benjamins.

Hornstein, Norbert. 2009. *A Theory of Syntax: Minimal Operations and Universal Grammar.* Cambridge: Cambridge University Press.

Hornstein, Norbert, Jairo Nunes, and Kleanthes K. Grohmann. 2005. *Understanding Minimalism.* Cambridge: Cambridge University Press.

Huang, C.-T. James. 1989. Pro-Drop in Chinese: A Generalized Control Theory. In *The Null Subject Parameter*, ed. Osvaldo Jaeggli and Kenneth J. Safir, 185–215. Dordrecht: Kluwer.

Kawasaki, Noriko. 1993. *Control and Arbitrary Interpretation in English.* Doctoral dissertation, University of Massachusetts, Amherst.

Kayne, Richard S. 2002. Pronouns and their Antecedents. In *Derivation and Explanation in the Minimalist Program*, ed. Samuel David Epstein and T. Daniel Seely, 133–166. Oxford: Blackwell.

Koster, Jan. 1984. On Binding and Control. *Linguistic Inquiry* 15(3): 417–459.

Kratzer, Angelika, and Lisa Selkirk. 2007. Phase Theory and Prosodic Spellout: The Case of Verbs. *The Linguistic Review* 24(1): 93–135.

Kratzer, Angelika. 2009. Making a Pronoun: Fake Indexicals as Windows into the Properties of Pronouns. *Linguistic Inquiry* 40(2): 187–237.

Kuno, Susumu. 1975a. Three Perspectives in the Functional Approach to Syntax. In *Papers from the Functionalism Parassession,* Eleventh Meeting of the Chicago Linguistic Society, 276–336.

Kuno, Susumu. 1975b. Super-Equi-NP Deletion is a Pseudo-Transformation. In *Proceedings of the 5th Meeting of the North Eastern Linguistic Association*, 29–44. Amherst, MA: GLSA.

Landau, Idan. 1999. *Elements of Control.* Doctoral dissertation, MIT.

Landau, Idan. 2000. *Elements of Control: Structure and Meaning in Infinitival Constructions*. Dordrecht: Kluwer Academic Publishers.

Landau, Idan. 2001. Control and Extraposition: The Case of Super-Equi. *Natural Language and Linguistic Theory* 19(1): 109–152.

Landau, Idan. 2003. Movement Out of Control. *Linguistic Inquiry* 34(3): 471–498.

Landau, Idan. 2004. The Scale of Finiteness and the Calculus of Control. *Natural Language and Linguistic Theory* 22(4): 811–877.

Landau, Idan. 2006. Severing the Distribution of PRO from Case. *Syntax* 9(2): 153–170.

Landau, Idan. 2007. Movement-Resistant Aspects of Control. In *New Horizons in the Analysis of Control and Raising*, ed. William D. Davies and Stanley Dubinsky, 203–325. Dordrecht: Springer.

Landau, Idan. 2008. Two Routes of Control: Evidence from Case Transmission in Russian. *Natural Language and Linguistic Theory* 26(4): 866–924.

Landau, Idan. 2010. The Explicit Syntax of Implicit Arguments. *Linguistic Inquiry* 41(3): 357--388.

Landau, Idan. 2013. *Control in Generative Grammar: A Research Companion*. Cambridge: Cambridge University Press.

Landau, Idan. 2015. *A Two-tiered Theory of Control*. Cambridge, MA: MIT Press.

Landau, Idan, and Jonathan David Bobaljik. 2009. Icelandic Control is not A-movement: The Case from Case. *Linguistic Inquiry* 40(1): 113–154.

Lasnik, Howard, and Mamoru Saito. On the Nature of Proper Government. *Linguistic Inquiry* 15(2): 235–289.

Lebeaux, David. 1984. Anaphoric Binding and the Definition of PRO. In *Proceedings of NELS 14*, 253–274.

Lebeaux, David. 1985. Locality and Anaphor Binding. *The Linguistic Review*. 4(4): 343–363.

Lewis, Richard L. 2010a. Efficiency, Economy, Optimality, Simplicity, Minimality, Search, Mapping, Function, State: An Attempt at Clarifying Definitions. Ms., University of Michigan.

Lewis, Richard L. 2010b. The Strong Minimalist Thesis and Bounded Optimality. Ms., University of Michigan.

Lyngfelt, Benjamin. 1999. Optimal Control: An OT Perspective on the Interpretation of PRO in Swedish." *Working Papers in Scandinavian Syntax* 63: 75–104.

Lyngfelt, Benjamin. 2000. OT Semantics and Control. *Rutgers Optimality Archive* #ROA 411.

Manzini, Maria Rita. 1983. On Control and Control Theory. *Linguistic Inquiry* 14(3): 421–446.

Manzini, Maria Rita. 1986. On Binding Theory and Control. In *Proceedings of the Sixteenth Annual Meeting of NELS*. Amherst, MA: GLSA.

Manzini, Maria Rita, and Anna Roussou. 2000. A Minimalist Theory of A-Movement and Control. *Lingua* 110: 409–447.

Martin, Roger. 1996. *A Minimalist Theory of PRO and Control*. Doctoral Dissertation, University of Connecticut, Storrs.

Martin, Roger. 2001. Null Case and the Distribution of PRO. *Linguistic Inquiry* 32(1): 141–166.

Miyagawa, Shigeru. 2012. Agreements that occur mainly in the main clause. in *Main Clause Phenomena: New Horizons*, ed. Lobke Aelbrecht, Liliane Haegeman, and Rachel Nye, 79–111. Amsterdam: John Benjamins.

Miyagawa, Shigeru. 2017. *Agreement Beyond phi*. Cambridge, MA: MIT Press.

Motut, Alexandra. 2010. Merge over Move and the Empirical Force of Economy in Minimalism. MA thesis, University of Toronto. Published as Volume 33, Toronto Working Papers in Linguistics.

Motut, Alexandra, and Margaret Grant. 2017. Experiencers do not Categorically Block Long-distance Control of PRO: Evidence from Naturalness Ratings and Eye Movements during Reading. In *Proceedings of the 34th West Coast Conference on Formal Linguistics*, ed. Aaron Kaplan, Abby Kaplan, Miranda K. McCarvel, and Edward J. Rubin, 387–394. Somerville, MA: Cascadilla Proceedings Project.

Motut, Alexandra, and Kenji Oda. 2011a. 'Move over Move': Agree-driven syntax in an expanded search domain. Presented at *New Perspectives on Language Creativity: Composition and Recursion,* Université du Québec à Montréal, September 25–27.

Motut, Alexandra, and Kenji Oda. 2011b. What do we mean by 'local'? An Expanded Search Domain Unifying Merge and Move. Presented at *The Minimalist Program: Quo Vadis?* Universität Potsdam, October 3–6.

Nunes, Jairo. 1995. *The Copy Theory of Movement and Linearization of Chains in the Minimalist Program.* Doctoral dissertation, University of Maryland.

Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement.* Cambridge, MA: MIT Press.

O'Neil, John. 1995. Out of Control. in Proceedings of NELS 25, ed. Jill Beckman, 361–371. Amherst, MA: GLSA.

Perlmutter, David M., and Scott Soames. *Syntactic Argumentation and the Structure of English.* Berkeley: University of California Press.

Pesetsky, David, and Esther Torrego. 2006. Probes, Goals, and Syntactic Categories. *Proceedings of the 7th annual Tokyo Conference on Psycholinguistics*, ed. Yukio Otsu. Tokyo: Hituzi Syobo Publishing Company.

Polinsky, Maria, and Eric Potsdam. 2002. Backward Control. *Linguistic Inquiry* 33(2): 245–282.

Postal, Paul M. 1970. On Coreferential Complement Subject Deletion. *Linguistic Inquiry* 1(1): 439–500.

Postal, Paul M. 1974. *On Raising*. Cambridge, MA: MIT Press.

Potsdam. Eric. 2009. Malagasy Backward Object Control. *Language* 85(4): 754–784.

Reuland, Eric. 2011. *Anaphora and Language Design*. Cambridge, MA: MIT Press.

Rezac, Milan. 2002. Cyclic Domains and (Multiple) Specifiers Constructions. Ms., University of Toronto.

Reinhart, Tanya. 1991. Binding and Coreference Revisited. Ms., Tel Aviv University

Reinhart, Tanya. and Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24(4): 657–720.

Richardson, John F. 1986. Super-Equi and Anaphoric Control. In *Proceedings of the 22nd Regional Meeting of the Chicago Linguistic Society*, ed. Anne M. Farley, Peter T. Farley, and Karl-Erik McCullough, 248–261. Chicago: CLS.

van Riemsdijk, Henk. 2006. Grafts follow from Merge. In *Phases of Interpretation*, ed. M. Frascarelli, 17–44. Berlin: Mouton de Gruyter.

Ritter, Elizabeth, and Martina Wiltschko. 2018. Distinguishing Speech Act Roles from Grammatical Person Features. In *Proceedings of the 2018 Annual Conference of the Canadian Linguistic Association.*

Ritter, Elizabeth, and Martina Wiltschko. 2019. Nominal Speech Act Structure: Evidence from the Structural Deficiency of Impersonal Pronouns. *Canadian Journal of Linguistics*. 64(4): 709–729.

Rizzi, Luigi. 1986. Null Objects in Italian and the Theory of *pro*. *Linguistic Inquiry* 17(3): 501–557.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.

Roeper, Thomas. 1993. Explicit Syntax in the Lexicon: The Representation of Nominalizations. in *Semantics and the Lexicon*, Ed. James Pustejovsky, 185–220. Dordrecht: Springer.

Roeper, Thomas. 2013. There-Insertion: How Internal Merge Guides the Acquisition Path. In *Generative Linguistics and Acquisition: Studies in Honor of Nina M. Hyams*, ed. Misha Becker, John Grinstead, and Jason Rothman, 249–270. Amsterdam: John Benjamins.

Rosenbaum, Peter S. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: MIT Press.

Rosenbaum, Peter S. 1970. A Principle Governing Deletion in English Sentential Complementation. In *Readings in English Transformational Grammar*, ed. Roderick A. Jacobs and Peter S. Rosenbaum, 20–29. Waltham, MA: Ginn.

Ross, John Robert. 1970. On Declarative Sentences. In *Readings in English Transformational Grammar*, ed. Roderick Jacobs and Peter Rosenbaum, 222–272. Waltham, MA: Ginn & Co.

Saab, Andrés Leandro. 2014. Syntax or Nothing: Some Theoretical and Empirical Arguments on Implicit Arguments. *Borealis: An International Journal of Hispanic Linguistics* 3(2): 125–183.

Safir, Ken. 1987. The Syntactic Projection of Lexical Thematic Structure. *Natural Language and Linguistic Theory* 5(4): 561–601.

Safir, Ken. 2004a. *The Syntax of Anaphora*. Oxford: Oxford University Press.

Sag, Ivan A., and Carl Pollard. 1991. An Integrated theory of Complement Control. *Language* 67(1): 63–113.

Sato, Yosuke. On the Movement Theory of Obligatory Control: Voices from Standard Indonesian. *Canadian Journal of Linguistics* 56(2): 276–275.

Saxon, Leslie. Control and Control Verbs. In *Proceedings of the West Coast Conference on Formal Linguistics*, ed. E. Jane Fee and Katharine Hunt, 347–357.

Sells, Peter. Aspects of Logophoricity. *Linguistic Inquiry* 18(3): 445–479.

Shima, Etsuro. 2000. A Preference for Move over Merge. *Linguistic Inquiry* 31(2): 375–385.

Sigurðsson, Halldór Ármann. 1990. Long Distance Reflexives and Moods in Icelandic. In *Modern Icelandic Syntax*, Syntax and Semantics 24, Ed. Joan Maling and Annie Zaenen, 309–346. San Diego: Academic Press.

Simpson, Andrew, and Zoe Wu. 2002. Understanding Cyclic Spell Out. In *Proceedings of NELS 32*, ed. Masako Hirotani, Article 13. Amherst, MA: Scholarworks.

Speas, Margaret, and Carol Tenny. 2003. Configurational Properties of Point of View Roles. in *Asymmetry in Grammar*, ed. Anna Maria Di Sciullo, 315–344. Amsterdam: John Benjamins.

Starke, Michal. 2001. *Move Dissolves into Merge: A Theory of Locality*. Doctoral dissertation, University of Geneva.

Staub, Adrian, and Keith Rayner. 2007. Eye Movements and On-line Comprehension Processes. In *The Oxford Handbook of Psycholinguistics* (1st ed.), ed. M. Gareth Gaskell and Gerry T. M. Altmann, 29 pp. Oxford: Oxford University Press.

Sundaresan, Sandhya. 2011. A Reductionist Treatment of Control and Anaphora. Semantic Scholar, No. 221081438

Thráinsson, Höskuldur. 1991 Long-distance reflexives and the typology of NPs. In *Long-distance Anaphora*, ed. Jan Koster and Eric Reuland, 49–75. Cambridge: Cambridge University Press.

Uriagereka, Juan. 1999. Multiple Spell Out. In *Working Minimalism*, ed. Samuel David Epstein and Norbert Hornstein, 257–383. Cambridge, MA: MIT Press.

Wexler, Ken, and Peter Culicover. 1977. Some Syntactic Implications of a Theory of Language Learnability. In *Formal Syntax*, ed. Peter Culicover, Thomas Wasow, and Adrian Akmajian, 7–60. New York: Academic Press.

Wexler, Ken, and Peter Culicover. 1980. *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.

Williams, Edwin. 1980. Predication. *Linguistic Inquiry* 11(1): 203–238.

Williams, Edwin. 1985. PRO and the Subject of NP. *Natural Language and Linguistic Theory* 3(3): 297–315.

Williams, Edwin. 1992. Adjunct Control. In *Control and Grammar*, ed. Richard K. Larson, Sabine Iatridou, Utpal Lahiri, and James Higginbotham, 297–322. Dordrecht: Springer.

Williams, Edwin. 1994. *Thematic Structure in Syntax*. Cambridge, MA: MIT Press.

Wood, Jim. 2012. Against the Movement Theory of Control: Another Argument from Icelandic. *Linguistic Inquiry* 43(1): 322–330.

Wurmbrand, Susi. 2013. QR and Selection: Covert Evidence for Phasehood. in *NELS 42: Proceedings of the Forty-Second Annual Meeting of the North East Linguistic Society*, ed. Stefan Keine and Shayne Sloggett, 619–634. Amherst, MA: GLSA.

Wurmbrand, Susi. 2014a. The Merge Condition: A Syntactic Approach to Selection. In *Minimalism and Beyond: Radicalizing the Interfaces*, ed. Peter Kosta, Lili Schürcks, Steven Franks, and Teodora Radev-Bork, 137–177. Amsterdam: John Benjamins.

Wurmbrand, Susi. 2014b. Tense and Aspect in English Infinitives. *Linguistic Inquiry* 45(3): 403–447.

Vanden Wyngaerd, Guido. 1994. *PRO-legomena: Distribution and Reference of Infinitival Subjects*. Berlin: Mouton De Gruyter.

Zwart, Jan-Wouter. 2002. Issues Relating to a Derivational Theory of Binding. In In *Derivation and Explanation in the Minimalist Program*, ed. Samuel David Epstein and T. Daniel Seely, 269–304. Oxford: Blackwell.