Chapter 2
Theta Theory and Obligatory Control

The detail of the pattern is movement.
—T.S. Eliot, “Burnt Norton”

2.1 Introduction

In this chapter an explicit theory of theta role assignment is proposed. This theory, along
with the independently motivated requirements of movement and feature checking in the
Minimalist Program (described in Chomsky 1993 and less completely in chapter 1 of the
present work), is used to derive the properties of PRO by allowing the assignment of
more than one theta role to long Case-chains which span multiple clauses. In this way, we
seek to unify “control” and “raising” structures in English, showing PRO to be unnecessary.
Without PRO, Control Theory as a separate module of grammar can be trivially eliminated.
This would be desirable, since the proper formulation of Control Theory has always
proven to be a problem, as was seen in chapter 1.

Consider the following sentences with respect to the Minimalist Program.

(1) a. I believe [the doctor₁ to have examined Sam₂]
b. I believe [Sam₂ to have been examined t₂ by the doctor₁]
(2) a. I persuaded the doctor₁ [PRO₁ to examine Sam₂]
b. I persuaded Sam₂ [PRO₂ to be examined t₂ by the doctor₁]

(1a) and (1b) are truth conditionally synonymous, whereas (2a) and (2b) are not. The
standard analysis\textsuperscript{1} for (1-2) is that in (1) the embedded subject raises to the matrix SpecAgr\textsubscript{0}P after Spell-Out to check its Case, whereas in (2) the object is in the matrix clause—PRO is the subject of the embedded clause. PRO has a different controller in (2a) and (2b), accounting for the lack of synonymy. The approach to be adopted in this work assumes that there is raising of the embedded subject to SpecAgr\textsubscript{0}P in both (1) and (2); we assert that that chain receives a single theta role in (1), and multiple theta roles in (2). The difference in theta assignment between (2a) and (2b) accounts for the lack of synonymy.

In this paper we will make the following assumptions concerning Minimalist syntax, many of which are uncontroversial.

• I assume that there is some ordering of Theta Roles by a “Thematic Hierarchy.” The particulars of the Thematic Hierarchy—the nature of the theta role labels and their ordering—are beyond the scope of this thesis, although the discussion of these issues by Grimshaw (1990) and Jackendoff (1983, 1990) is influential.

• Also, following Chomsky (1993), I assume that the principle of Full Interpretation requires an LF representation to consist of all and only “legitimate” objects. Moreover, in order to be legitimate, an object must be interpretable; to be interpretable, an argument must be theta-marked. In particular, for an Case-chain to be interpretable, it must be theta-marked (i.e. the argument of a semantic function) at LF.

• Because theta roles are not “morphological” in any sense, I assume that theta role requirements cannot cause morphologically driven movement. For instance, a object cannot move solely to get a theta role.

• Finally, for expository clarity the English functional projections Tense and

\textsuperscript{1}That is, the standard analysis within the Minimalist Program. In earlier Principles and Parameters Theories, the analysis of ECM verbs did not involve movement—hence “exceptional case marking” over the clause boundary. In fact, examples like these sentences have a long history of discussion, ranging at least back to Aspects (Chomsky 1965).
AgrS are fused into Infl, following Thráinsson (1996). This may be done without any loss of generality, since the chapter (unlike others) will be exclusively concerned with data in English, and in English T always moves to AgrS, and SpecTP is not available as a landing site (Bobaljik and Jonas 1996, Jonas and Bobaljik 1993).

These ideas, as well as the description of the Minimalist Program in chapter 1, will motivate the details of the Theta Theory presented here; the reduction of obligatory control to these principles is serendipitous.

To give an overview of this chapter: in §2.2 we look at the relation between Theta assignment and LF objects and structures; in §2.3 we look at “obligatory control” and compare it to raising; in §2.4 consider the nature of subject and object control, and in §2.5 we consider the interaction of scopal properties and the theory we have advanced here, and demonstrate a solution to a long-standing puzzle. In §2.6 the ever-popular wanna-contraction will be examined. We will derive the properties of these data from basic principles of movement and theta-role assignment. We conclude that PRO is unnecessary and hence that obligatory control is trivially eliminable.

2.2 Properties of Theta Theory at LF

In this section, we will consider how theta role assignment (occurring at LF) is constrained by LF structure. Since the structure of LF is largely presupposed in the Minimalist Program (as we saw in chapter one), this section will mostly be concerned with the conditions under which a particular theta role may be assigned. As a simple example, let’s look at the following sentence Jamie left. In this sentence, the verb has but one theta role to assign, and there is only one DP which can possibly receive that theta role — not only because of the simplicity of the sentence, but also because that DP is quite local to the theta-assigning verb at LF. To answer these questions conclusively, we must consider the sentence’s LF structure (as we will below) but we also must define some relevant concepts.
Let us formalize this intuition in the following way. In (3) we define the Theta Domain, within which a head may assign a theta role. The Theta Domain is identical to the Minimal Domain, as defined in (4) and discussed in chapter 1.

(3) **Theta Domain**

The Theta Domain of a head-chain η (within which η’s theta roles are assigned at LF) is η’s Minimal Domain, as defined in Chomsky (1993).

(4) **Minimal Domain** (from Chomsky 1993)

The Minimal Domain, MD(α), of head-chain α is the set of nodes such that:

(i) each member of MD(α) is dominated by the least full-category maximal projection which dominates every position in α;

(ii) each member of MD(α) is distinct from and does not contain α; and

(iii) no member of MD(α) is dominated by any other member of MD(α).

We want to limit the types of syntactic objects which can potentially be assigned theta roles. Trivially we can limit them to legitimate LF objects, but there is another limitation which we will use. In (5), Visibility is defined so that only Case-chains (and other elements whose Case features are satisfied) will be available for theta assignment.

(5) **Visibility** (adapted from Chomsky 1981)

To be Visible for theta marking at LF, a legitimate LF object must have had its Case feature checked.

Therefore, only syntactic objects with checkable Case-features can be assigned theta roles at LF. This includes Case-chains, the complements of prepositions, and CP/IP clauses as
objects which can be “theta-marked.” Now let us consider the LF structure corresponding to the sentence Jamie left, as given in (6).

The LF structure in (6) is a simple example of theta role assignment according to the definitions in (3), (4) and (5). The Case-chain \{[[DP Jamie], t]\} is shown, as is the chain of

There are two minor points here. First, we note that APs can be theta-marked as well but have no Case features. In chapter 3, we will more closely consider the syntactic role of APs, so the question can be set aside here. The other issue is potentially more important: we want a verb to take the DP object of a PP as an argument—but structurally, it seems that the DP is outside the Minimal Domain of the verb. In fact, we can appeal to a fundamental ambiguity of prepositions: under some circumstances prepositions may be considered lexical heads, and under other circumstances functional heads. The argument for prepositions being functional heads is very strong:

(i) prepositions are “closed-class” items — new prepositions are rarely coined; and
(ii) frequently, prepositions are used purely as grammatical markers, without regard to any independent semantic content the preposition might contribute to the meaning of the utterance.

If we accept that prepositions—or at least some prepositions—are functional heads, then the preposition has a relationship to the determiner and the noun of the DP inside the PP which is analogous to the relationship of the complementizer head to the tense and verb heads in the C’’s complement clause. It then becomes acceptable, even expected, that the set of functional heads associated with a lexical head (Agr’, Tns’ and C’ with V’; and Det’, Agr’ and P’ with N’) operate as a single unit with respect to certain operations, as do the different projections of a single head. These shared operations include the passing of feature values and feature checking; it is therefore not exceptional that we allow a theta role nominally assigned to a PP to ‘percolate’ down to the DP where it can be properly interpreted.
the verb *left* from V to Infl. The Theta Domain of *left* is determined according to (3); every node in the Theta domain of *left* in (6) has a square around it. Clearly, one of the positions in the Case-chain \{[\text{DP} Jamie], t\} is in the Theta Domain of *left* (in fact, both are). Therefore, the verb *left* can assign its theta role to the Case-chain \{[\text{DP} Jamie], t\} at LF.

Before we continue, we must address the additional complexities of a head assigning multiple theta roles. In the standard Government and Binding Theory of Chomsky (1981, 1982, 1986a, 1986b) — the canonical version of Principles and Parameters syntax — it is a consequence of the Projection Principle that it is impossible to assign more than one theta role to an argument. However, in the Minimalist Program, the Projection Principle has been eliminated as a principle of grammar. This was largely inevitable. The Projection Principle had the effect of preserving the thematic relations established at D-structure through all the other syntactic levels. Given the elimination of D-structure and S-structure, and the fact that in the Minimalist Program theta-role assignment is not resolved until LF in any case, the Projection Principle has become conceptually superfluous: either an LF object is interpretable, or it is not. Also, the empirical data explained by the Projection Principle in GB Theory has been subsumed (or is derived) by the mechanisms of structure-building and movement outlined in chapter 1. (For more details of this reduction of the Projection Principle to other mechanisms in the Minimalist Program, see Chomsky 1993, p. 20-21.)

However, there is one relic of the Projection Principle which we want explicitly to preserve: that a given function (i.e. theta role assigner) can assign only a single theta role to a particular argument. This in turn implies that a head can assign only a single theta role to a particular Case-chain. We call this principle Thematic Uniqueness and define it below.
(7) **Thematic Uniqueness**

For a theta-role assigning head \( \eta \) and a legitimate LF object \( \alpha \), no more than one of the theta roles of \( \eta \) may be assigned to \( \alpha \).

There are two arguments for Thematic Uniqueness: one conceptual and the other empirical. Conceptually, in the theory of thematic roles presented here, a head’s theta roles are in a correspond directly to the “argument slots” of that head’s semantic interpretation. Each theta role establishes the requirement for a potential “filler” for that theta role, whether or not that filler is realized syntactically. Given a number of potential LF objects which can be assigned theta roles, the maximally simplest way to formulate a theta-role assignment algorithm is to assume that the theta-role assignment function requires that two different theta roles cannot be assigned to the same object, and that two different objects cannot be assigned the same theta role. Then, (7) is merely a corollary to that. Another conceptual argument refers to the lack of “global” theta role labels. In the absence of unique global labels for theta roles it is minimally complicated if the role label consists of the pair \( <\alpha, \beta> \), where \( \alpha \) is a theta-role assigning head and \( \beta \) is a legitimate LF object which is \( \alpha \)’s argument (and as we have seen above, usually a Case-chain). In order for \( <\alpha, \beta> \) to uniquely identify the thematic relationship, there can be only one theta role assigned by \( \alpha \) to \( \beta \). Empirically, consider the LF structure in (8) for the sentence *Jamie saw.*
Thematic Uniqueness eliminates the possibility that the theta-role assigning head *saw* could assign both its theta roles to the same Case-chain \{[DP Jamie], t\}. If this were permitted, it would allow (8) to be interpreted to mean ‘Jamie saw herself/himself.’ There are languages with null reflexive pronouns and arguably languages with null reflexive morphemes (for instance the verb *wash* in English, as used in the sentence *Chris washed*), but both null reflexive pronouns and null reflexive morphemes are best viewed as theta-grid-changing lexical operations and therefore do not provide counterexamples to Thematic Uniqueness. This topic, and the general problem of reflexivity in morphology and semantics, is dealt with extensively in Reinhart and Reuland (1991).

If there is more than one legitimate LF Case-chain and more than one theta role to be assigned by a head, how do we determine which Case-chain is assigned which theta role? The Theta Hierarchy Condition in (9) ensures that the two orderings relevant to the assignment of theta roles simply “match up”: one ordering is given by the Thematic Hierarchy (which orders theta roles themselves according to a number of criteria), and the other is determined by the definition of Superordinate in (10), which defines a structural
ordering among Case-chains at LF. We are guaranteed that, for any two Case-chains passing through the same Theta Domain, one will be Superordinate to the other. Therefore there is no ambiguity about which Case-chain will receive which theta role LF—based on the derivational history of the corresponding Case-chain, taking into account movement both before and after Spell-Out.

(9) **Theta Hierarchy Condition (THC)**

Assume function $\eta$ assigns theta role $\theta_1$ to LF object $\alpha$, and assigns $\theta_2$ to LF object $\beta$. Then, $\theta_1$ is higher on the thematic hierarchy (of $\eta$) than $\theta_2$ if and only if $\alpha$ is Superordinate to $\beta$.

(10) **Superordinate**

Given two chains $\alpha$ and $\beta$, $\alpha$ is Superordinate to $\beta$ iff some position in $\alpha$ c-commands every position in $\beta$.

For example, consider the sentence *Jamie saw Andy*, and its LF structure (11).
The Case-chains of \{[\text{DP Jamie}], t\} and \{[\text{DP Andy}], t\} are shown, the former connected with a solid line to indicate that it is formed before Spell-Out, the latter with a dashed line to indicate its formation after Spell-Out. By (10), \{[\text{DP Jamie}], t\} is Superordinate to \{[\text{DP Andy}], t\}. Both Case-chains are in the Theta Domain of the V-chain of \text{saw}, as shown in (11). Thus the THC requires that \{[\text{DP Jamie}], t\} be interpreted as having the theta role higher on the Theta Hierarchy of \text{saw}—the “agent” role. Consequently, the chain \{[\text{DP Andy}], t\} is associated with the remaining “patient” theta role, which is lower on the Thematic Hierarchy. This yields the correct interpretation of (11), and any different assignment of theta roles is excluded by (9).

\[\text{(11)}\]
2.3 Control versus Raising Constructions

In this section we compare “control” and “raising” sentences (and their LF structures) with respect to the version of Theta Theory presented in §2.2. We will determine the possible theta assignments for a minimal but canonical set of sentences and their LF structures, deriving their properties using the theory of theta assignment presented here and, critically, without making use of PRO. In particular, the sentences to be compared involve want with a raising/ECM object and with a controlled object, and believe with a raising/ECM object—in addition to the ungrammatical status of believe as a control verb.

First, we consider the sentence Jamie wants Andy to have left, along with its LF structure in (12) below. There are two Case-chains in the LF in (12): \{[DP Jamie], t\} and \{[DP Andy], t′, t\}, and by (10) it is clear that \{[DP Jamie], t\} is Superordinate to \{[DP Andy], t′, t\}. Also, there are two theta-role assigners: the verb left which assigns only one theta role within its theta domain (which is limited to the embedded SpecVP), and the verb wants, which assigns a agent theta role and a theme theta role, the latter of which is semantically constrained to be assigned to a clause (and is assigned obligatorily to the IP). The theta domain of wants is the minimal domain of the verb chain from V˚ to AgrO˚, which contains the embedded IP, and the matrix SpecVP and SpecAgrOP. Only the Case-chain \{[DP Andy], t′, t\} is in the Theta Domain of the verb left. Hence, the THC requires that the Case-chain \{[DP Andy], t′, t\} be assigned the sole theta role of left. Finally, the Case-chain \{[DP Jamie], t\}, must receive the agent theta role of wants; otherwise it would be assigned no theta role at all and hence be an illegitimate LF object (or be uninterpretable at LF) and crash the derivation.
In (13) we see the LF structure of the sentence *Andy wants to have left*. Carefully note the theta domain of *wants* in (13), and how that theta domain differs from the theta domain of *wants* in (12). In (13), the chain of the verb *wants* goes from $V^\circ$ to $I^\prime$, and hence the theta domain of *wants* in (13) is the minimal domain for that chain: SpecIP, SpecVP and the embedded IP. In contrast, the verb chain in (12) goes to $AgrO^\prime$, and therefore we would expect its theta domain to be different.
In (13) there is only a single Case-chain, \{\text{\text{DP}} \ Andy, \ t''', \ t', \ t\}, which moves in a successive cyclic fashion from the embedded to the matrix clause in order to be Case-checked. Being a long Case-chain, \{\text{DP} \ Andy, \ t''', \ t', \ t\} is in the Theta Domain of both \textit{left} —the position \(t_{\text{Andy}}\) in (13)—and in the Theta Domain of \textit{wants} —both \text{\text{DP}} \ Andy\ and \(t_{\text{Andy}}''\) in (13). There are two theta roles assigned at LF, one by each verb (excluding the theme theta role of \textit{wants}, which is constrained to be assigned to the complement IP clause and therefore does not enter into these calculations). However, since there is only one Case-chain to which both theta roles can be assigned, \{\text{\text{DP}} \ Andy, \ t''', \ t', \ t\} is assigned the theta role of \textit{left} as well as the agent theta role of \textit{wants}. Moreover, note that the only
difference between (13) and the LF structure of a raising-to-subject construction like *Jamie seems to have left* is that *seems* does not assign a theta role to its subject, only to its IP complement clause—otherwise, the two are structurally identical at LF.

Next we consider the sentence *Jamie believes Andy to have left*, with the following LF structure.

At LF in (14) there are two Case-chains, {[DP Jamie], t} and {[DP Andy], t’, t}, and the Case-chain {[DP Jamie], t} is Superordinate to {[DP Andy], t’, t}. In (14) as in (12), the only Case-chain in the Theta Domain of *left* is {[DP Andy], t’, t}, so it is assigned the sole
theta role of \emph{left}. Also, as argued earlier concerning (12), the Case-chain \{[\emph{DP Jamie}], t\} must receive the agent theta role of \emph{believes}, because otherwise it would be assigned no theta role at all and hence be an illegitimate object at LF. (\emph{Believes} also assigns a theme theta role to its IP complement, which as before has no effect on these calculations.)

To complete this paradigm, we should consider the unacceptable sentence *\emph{Andy believes to have left} and its LF.

As in (13), there is a single Case-chain, \{[\emph{DP Andy}], t''', t', t\}, which is in the Theta Domain of both \emph{left} —because the node \(t_{\text{Andy}}\) in (14) is—and in the Theta Domain of
wants —since both \([\text{DP Andy}]\) and \(t_{\text{Andy}}\) are. Therefore, in analogy to the LF in (13) of the grammatical sentence *Jamie wants to have left, it should be acceptable to assign to the Case-chain \([\text{DP Andy}], t^{'}, t^{'}, t\) both the theta role of left and the theta role of wants. Therefore, in analogy to (13), *Andy believes to have left is predicted to be a grammatical sentence. However, this is an incorrect prediction. In fact, every verb which enters into raising but not into control constructions would face similar problems. Obviously there is some parametric difference between verbs like want and verbs like believe, and in chapter 3 we will propose a explanation of the difference which explain the ungrammaticality when applied to sentences like (13).\(^3\) To preview that future explanation, we will propose that there are differences between the Tense features of the T’ head in control and raising clauses—effectively, T’ in raising clauses has no LF interpretation, while in control clauses it does. This difference make it impossible in raising sentences for both the matrix and the embedded verb to assign theta roles to the same Case-chain. (Although otherwise unmotivated here, this condition on theta role assignment—Amalgamation—has a number of desirable consequences, as we will see in chapter 3.)

A similar question (but fortunately simpler to answer) concerns control constructions which do not allow corresponding object raising constructions. That is, given the previous discussion, there is no reason why (16a) should be ungrammatical if (16b) is grammatical, since we know that (16c&d) are both grammatical:

\(^3\)In addition to the explanation in chapter 3, the reader might consider the “low-tech” possibility that the sentence *John believes to have left is ungrammatical simply because the ECM verb believe cannot check its uninterpretable Case features against an object, and therefore the uninterpretable Case features make the verb itself uninterpretable at LF. This simple explanation actually has a quite powerful explanatory capacity, since it solves many of the problems posed in chapter 3. Unfortunately, it also poses a number of new problems of its own. For example, a large number of verbs—maybe most “plain” transitive verbs—in English allow a significant degree of optionality in whether they have objects. Consider the following:

(i) John ate.

Why is it that Case-checking is optional with eat but not with ECM verbs like believe or small-clause verbs like consider? Maybe this question has a plausible answer, but most of the obvious answers seen quite stipulative. Why, for instance, would there be a link between Case-checking by verbs and the tense-features of a nonfinite embedded T? Perhaps the LF-interpreted head adjoining to the verb absorbs Case-features obligatorily? Clearly, it is true that selecting an LF-interpreted head and having checkable Case-features are somehow in complementary distribution. More on this in chapter 3.
(16)  a. *Andy\textsubscript{1} tried [Jamie\textsubscript{2} to leave]
    b. Andy\textsubscript{1} wanted [Jamie\textsubscript{2} to leave]
    c. Jamie\textsubscript{1} tried [ t\textsubscript{1} to leave]
    d. Jamie\textsubscript{1} wanted [ t\textsubscript{1} to leave]

However, there is a avenue for expressing these differences in the Minimalist Program. Let us suppose that certain (transitive) verbs have Case features which can be checked after the verb moves covertly to SpecAgrOP, while other (intransitive) verbs do not have such features. This dimension of variation must be allowed in English in any case, to account for normal transitivity and intransitivity. However, if we posit that \textit{wanted} is transitive and \textit{tried} is intransitive, then if would be absolutely clear why (16a) is ungrammatical: the DP \textit{Jamie} cannot have its Case-features checked, due to the intransitivity of \textit{tried}. In fact, it cannot even form a grammatical Case-chain, and as such is not a legitimate object at LF, eligible to be assigned a theta role by a verb.

A possible objection to this is the observation that both \textit{want} and \textit{try} can also take plain DP complements: sentences like \textit{Jamie wanted the wine} and \textit{Jamie tried the wine} are both quite grammatical. However, we can respond that there is no reason to expect that \textit{try} selecting a IP complement clause should have identical Case features to \textit{try} selecting a DP complement, and the same would go for \textit{want}. It is therefore (more or less) a coincidence that both \textit{want}-with-an-IP-complement and \textit{want}-with-a-DP-complement check Case. As we see in (16), \textit{try} has Case features when it selects a DP complement, but not when it selects an IP complement.

2.4 Subject Control versus Object Control

In this section we compare subject and object control structures, and using the Theta Theory defined in §2.2 and §2.3 we will derive their thematic properties without making
use of PRO. Once again a minimal pair of sentences is contrasted: Jamie persuaded Andy to leave and Jamie promised Andy to leave, with their LF in (17) and (18), respectively.

In the LF in (17), corresponding to Jamie persuaded Andy to leave, there are two Case-chains: \{\{DP Jamie\}, t\} and \{\{DP Andy\}, t', t\}. Only the latter is in the Theta Domain of left, so \{\{DP Andy\}, t', t\} is assigned the sole theta role of left. Both \{\{DP Jamie\}, t\} and \{\{DP Andy\}, t', t\} are in the Theta Domain of persuade, and since \{\{DP Jamie\}, t\} is Superordinate to \{\{DP Andy\}, t', t\}, the THC requires that \{\{DP Jamie\}, t\} be assigned the theta role highest in the Thematic Hierarchy of persuade. The “patient” theta role of persuade is assigned to \{\{DP Jamie\}, t\}.
Andy, t', t}, and the embedded IP receives the theme theta role obligatorily.

The LF of “subject control” verbs like promise is given in (18), and differs structurally from the LF of persuade in (17), in that there is a extra “VP shell” — a VP embedded directly inside another VP — which permits there to be an object position in the lower VP shell.

The underlying observation is that promise is a ditransitive verb. It takes two complements, a DP and an IP. This contrasts with persuade, which takes only a single complement IP,
and *promise* can take two complements because of the extra structural positions created by the VP shell. Notice, though, that *promise* and *persuade* assign the same number of theta roles, and both check the same number and type of Case features. The difference between *promise* and *persuade* is their selection of complements, which therefore cannot be derived from the Case and theta properties of these verbs and must be another primitive of their lexical entries. However, the selectional properties of these verbs may be related to their semantics, a possibility which Larson (1991) discusses and which will be considered closely in chapter 4 in relation to control in Romance languages. For now, we will take the selection differences to be a primitive, although possibly derived from other lexical (syntactic or semantic) features of these verbs.\(^4\)

In order to implement these verb’s different selectional properties, the VP-shell analysis of Larson (1988) has been adopted. Within the Minimalist Program, this is the usual approach to ditransitives and other VP structures which contain more than one complement. Otherwise having more than one complement would be impossible in a theory like the Minimalist Program which allows only binary-branching trees. Also the VP-shell approach gives a completely structural explanation for c-command asymmetry: the first VP complement c-commands the second, but the second does not c-command the first.\(^5\) So, the LF in (18) corresponding to *Jamie promised Andy to leave* contains an additional VP shell structure; this was proposed by Larson (1991) specifically for verbs like *promise*.

\(^4\)The extent to which selectional features can be derived from theta and Case features is debatable. (For several points of view in this debate, see Grimshaw 1979, Pesetsky 1982, and Boškovic 1996.) However, as long as we suppose that verbs with similar Case and theta features can differ with respect to their VP shell structure, then some mechanism like selection is needed—of the complement category, of the complement category’s head features, or something similar.

\(^5\)To give one example, this c-command asymmetry can be seen in the occurrence of reflexives, which depends on c-command. Consider these sentences, as discussed by Barss and Lasnik (1986):

(i) a. John showed Bill, himself.
   b. *John showed himself, Bill.*

Condition A of the binding theory requires that a reflexive be c-commanded by its antecedent. This minimal pair demonstrates that condition A is satisfied in (ia) but not in (ib). If a “flat” VP structure—where the verb and both objects are sisters—is assumed for these sentences, we must change the definition of c-command to take into account precedence relations between tree sisters. (Barss and Lasnik 1986 argue precisely for this.) On the other hand, if we accept Larson’s VP shell model, the definition of c-command can remain the same.
The extra VP-shell allows the Case-chain \{[DP Jamie], t', t', t\} to “leapfrog” \{[DP Andy], t\}, as seen in (19). It is allowed because, at the point of the derivation shown in (18), movement to the potential landing site (the empty SpecVP) is equally the “Shortest Move” for both Case-chains. Recall the definition of Equidistance in chapter 1: if \(\alpha\) and \(\beta\) are in the same minimal domain, they are equidistant from \(\gamma\). In (19), the DP Andy does not block the movement of Jamie to SpecVP, because Andy and SpecVP are both in the minimal domain of the chain \{promised, \(t_{\text{promised}}\}\}. Therefore, Andy is no closer to Jamie than SpecVP is, and Andy thus does not cause a Shortest Move violation.

(19)

Two moves are possible at the stage of the derivation shown in (19), one in which the DP
Jamie is moved to the higher SpecVP, and the other in which Andy is moved there. However, the “wrong” derivation (where Andy is moved) never leads to a convergent derivation, because in that derivation, Jamie is “trapped” in the lower VP and unable to check its Case features; the verb chain will not extend any further, and therefore Equidistance will not prevent a higher DP from blocking the movement of Jamie to check its Case feature. On the other hand, the other possibility (where Jamie is moved) leads to the convergent LF shown in (18).

The LF structure in (18) contains two Case-chains: \{[\text{DP Jamie}], t′′, t′, t\} and \{[\text{DP Andy}], t\}. Since only one of the Case-chains, \{[\text{DP Jamie}], t′′, t′, t\}, contains a position in the Theta Domain of left, it receives the theta role of left. In addition, since \{[\text{DP Jamie}], t′′, t′, t\} is Superordinate to \{[\text{DP Andy}], t\}, the THC requires that \{[\text{DP Jamie}], t′′, t′, t\} be assigned the theta role highest in the Thematic Hierarchy of promise. The remaining theta role of promise is assigned to the Case-chain \{[\text{DP Andy}], t\}.

In addition, we can also account for Visser’s generalization (Visser 1973), that passives of subject control verbs are ungrammatical—an observations which is difficult to capture under a standard Control Theory. Consider the examples in (20-21).

(20)  a. Jamie persuaded Andy to leave
     b. Andy was persuaded to leave

(21)  a. Jamie promised Andy to leave
     b. *Andy was promised to leave

Under the analysis of promise we gave previously, it assigns two theta roles to DP’s, an agent and a goal (and also a third: a theme theta role to the IP), it checks Case features, and it selects a DP complement and an IP complement. Although there are many analyses of passivization, they have in common the property that the external theta role of the verb is suppressed and its Case-features are neutralized. Hence, if passivized, promise will
assign only its goal and theme theta roles, it will not check Case features, but it still selects the same DP and IP complements. In the case of promise, passivization does not clause promise to lose its subject because the subject’s theta role is gone—the subject is present anyway, because the embedded clause is still assigning it a theta role. Therefore a sentence like (21b) is ruled out: Andy, raised from the embedded clause, will not be able by itself to satisfy the selectional criteria of promise. The DP Andy cannot move to a complement position in the matrix clause because that would be a violation of Greed; there are no potential features to be checked there.

Thus, there are necessarily two DP’s which need to check their Case features in the matrix clause, despite the passivization of promise. However, a sentence like *Jamie was promised Andy to leave cannot be grammatical, since promise, once passivized, no longer has Case features to be checked against. The DP [DP Andy] cannot check its Case features, and the derivation crashes at LF because an uninterpretable feature (the Case feature of Andy) remains.

Bach’s generalization is another observation which is difficult to derive from the standard Control Theory. It was observed by Bach (1979) that object control verbs cannot appear with a missing object:

\[(22) \quad \begin{align*} 
a. & \quad \text{John promised to leave} \\
b. & \quad *\text{John persuaded to leave} 
\end{align*}\]

It is assumed that verbs allowing such “missing objects” have undergone a process of Detransitivization, and that Detransitivization is a operation (lexical or otherwise) which removes a complement DP from the selection features of the verb, suppresses that DP’s theta role and neutralizes the verb’s Case features. Given that formulation of Detransitivization, it is clear why persuade cannot undergo it. As we saw above, persuade lacks a DP complement; therefore Detransitivization cannot be applied to persuade, since
there’s no complement DP to unselect. In contrast, *promise* selects a DP complement and therefore can be Detransitivized, producing the grammatical (22a).

### 2.5 An Argument from the Scope Properties of Control

We have argued that the distribution and interpretation of control constructions is best explained by feature-checking and movement, rather than by Binding Theory or a separate Theory of Control. In fact, there are a few cases where control sentences behave markedly differently from predictions made by the Binding Theory.

Note that the following raising sentence, (23a), is ambiguous, while the analogous control sentence, (23b), is not.

\[(23)\]
\[
a. \text{Everyone is likely to have left the party.} \\
b. \text{Everyone wants to have left the party.}
\]

There are two interpretations for (23a):

\[(24)\]
\[
a. \text{Every } \{x; \text{likely´ (left´ (x, party)))}\} \\
b. \text{likely´ (Every } \{x; \text{(left´ (x, party)))}\}
\]

In (24a), for every person it is likely that he has left the party, although there still may be people at the party. In (24b), it is asserted that it is likely that no one remains at the party. The control sentence (23b) only has one interpretation, corresponding to the first interpretation of the raising sentence in (24a).

\[(25)\]
\[
\text{Every } \{x; \text{want´ (left´ (x, party)))}\}
\]

However, we demonstrated that these two sentences have the same structure (aside from
lexical choice) and that the only difference between the sentences was in theta assignment at LF. So, how can there differences in interpretation be explained?

First, let us make a commonsense assumption about syntactic variables. In most P&P theories, each Case-chain has exactly one variable, considered to be the head of the Case-chain (or the A-chain). This corresponds to the stipulation that each Case-chain has exactly one theta-assigned position. In the theory presented here, Case-chains may be assigned more than one theta-role by different theta-assigners. Therefore, let us propose a change to the definition of syntactic variable, so that a Case-chain may have more than one variable position in it, corresponding to the position in which each theta role is assigned to that Case-chain.

We also assume a standard “lowering” analysis of the narrow-scope raising interpretation (May 1985: p. 97ff). After a quantifier phrase moves to a position at Spell-Out, the QP may be “reconstructed”—that is, interpreted as if it were lower in the tree than its Spell-Out position. Any position is possible, with the one constraint that the quantifier phrase must still bind its variable positions at LF.

Therefore, it then automatically follows why the control sentence does not have the narrow-scope interpretation: the lowering of the quantifier phrase everyone to embedded-clause scope in (23b) would require the variable corresponding to the agent theta role of want to be unbound at LF, because—once lowered—the quantifier does not c-command that variable. Therefore the sentence (23b) would have a free variable at LF and not be interpretable. On the other hand, there is no theta role assigned by the raising verb in (23a), and hence no variable in the matrix clause. Therefore the quantifier phrase can optionally lower after Spell-Out to take scope only over the embedded clause without violating the condition that variables be bound at LF. Thus this prediction is in agreement with the data presented above.

Another argument for the analysis presented here is also from scope data. Consider the following sentences:
(26)  a. Only Reagan voted for himself
      b. Only Reagan wanted to leave

The first sentence is ambiguous between the following readings:

(27)  a. Only \( \{x, \text{Reagan}(x), x \text{ voted for Reagan}\} \)
      b. Only \( \{x, \text{Reagan}(x), x \text{ voted for } x\} \)

In (27a), Reagan voted for Reagan, and everyone else voted for someone other than Reagan. In (27b), Reagan is the only “self-voter”; Reagan cast a vote for Reagan and no one else voted for himself, but other people might have voted for Reagan. On the other hand, the control sentence lacks the first, so-called “strict” reading; only the second, “sloppy” reading is observed:

(28)  Only \( \{x, \text{Reagan}(x), x \text{ wants } x \text{ to leave}\} \)

In a standard Theory of Control, this cannot be easily explained. Consider a structure like the following:

(29)  \([dp, \text{Only Reagan}], \text{wants [PRO] to leave}]\)

Control Theory (or the Binding Theory) supplies the index to PRO, but we already know that this is not sufficient to prevent the “strict” reading from being generated, as seen with the pronoun in (26b). Therefore, to account for the fact that PRO does not act like a pronoun with respect to quantification, other stipulations must be made in a Control Theory based on PRO.
However, in the theory presented here, these judgements are correctly predicted. Consider the LF structure we propose for (26b):

\[(30) \quad [\text{IP } [\text{DP Only Reagan}]_1 \text{ wants}^+ [\text{VP } t_1 \text{ to } [\text{VP } t_1 \text{ leave}]]]\]

Each position in a Case-chain assigned a theta-role is a variable: in (30), the Case-chain \{[\text{DP Only Reagan}], t_1 \text{ wants}, t_1 \text{ leave}\} is assigned two theta roles by wants and by leave, and therefore the variables in the chain are the head of the chain \text{Only Reagan} and \text{t}_i. Therefore, instead of a coindexed PRO in the lower theta position, we have a variable which is part of the chain containing \[\text{DP Only Reagan}\]. Hence, there is no possibility for that position to be interpreted in a “strict” manner. Variable positions in a chain must be interpreted as variables, and cannot be given some other interpretation (the way a pronoun can be, as in (26a)). In (26b), only the “sloppy” reading is available with the variable, and we need no further stipulations to account for the data.

\subsection{2.6 Wanna-Contraction}

One argument for the reality of PRO has been the so-called “wanna-contraction,” although it occurs with a number of verbs in English. The basic data are well known:

\[(31) \quad \begin{align*}
\text{a.} & \quad \text{They want to see Alex.} \\
\text{b.} & \quad \text{They wanna see Alex} \\
\text{c.} & \quad \text{Who do they want to see } t_1 ? \\
\text{d.} & \quad \text{Who do they wanna see?} \\
\text{e.} & \quad \text{Who do they want } t_1 \text{ to see Alex?} \\
\text{f.} & \quad *\text{Who do they wanna see Alex?}
\end{align*}\]

For some reason, the wh-trace in (31f) blocks the wanna-contraction, possibly by preventing
want and to from being “adjacent” at PF. Note, though, that PRO does not block wanna-contraction, although it too is expected to intervene between the two words. The most commonly accepted analysis of wanna-contraction within GB Theory was advanced by Pesetsky (1982). He claimed that PRO’s position could be postposed. Its position was constrained by Case assignment, and if PRO was not assigned Case then its linear position to the left of the I head to was no longer obligatory. So, along with [IP PRO to VP], it was equally possible to get [IP to VP PRO]. In the latter case, PRO did not intervene between want and to, and therefore contraction could happen. On the other hand, a wh-trace was assigned Case in the position to the left of to, and therefore could not be postposed. Thus a wh-trace would always intervene between want and to.

In Minimalism, this analysis is questionable on several grounds. First, the LCA of Kayne (1994) independently requires that PRO, if it exists, intervene between want and to, because the LCA—beyond requiring binary branching—also forces all branching to be to the left. However, Pesetsky’s analysis assumes that the branch with PRO can go to the right in order to “save” the structure from ungrammaticality. Moreover, feature checking (replacing Case assignment) is no longer dependent on the linear order of the elements. If the analysis of Pesetsky (1982) is rejected, though, the only answer left is to stipulate that PRO and wh-traces have different PF-interpretations: PRO genuinely has none, while wh-traces, although they are not pronounced, have enough of a presence at PF to intervene and block wanna-contraction. However, the reason for this difference is unclear.

In the theory presented in this thesis, the intervening element in a wanna-contraction is not PRO, but instead it is the trace of a Case-chain. There is independent evidence showing that traces of Case-chains also do not block phenomena similar wanna-contraction. Consider the following data:

(32) a. Jamie has supposed Alex to be the most intelligent.
    b. Alex is supposed to be the most intelligent.
c. Alex is sposta be the most intelligent.
d. Who has Jamie supposed to be the most intelligent?
e. *Who has Jamie sposta be the most intelligent.

This evidence shows that traces of Case-chains allow wanna-contraction (here, with supposed instead of want). One could argue against the use of this data, because the sentences do not form a minimal pair: (32a, d, e) are in the active voice, while (32b, c) are passives (although whether there is a real difference between these occurrences of supposed in (32) is unknown). Nevertheless, we will take (32) at its face value. Therefore we need not make a special condition for PRO, which does not exist anyway; we simply reduce the distinction in cases of wanna-contraction to a already-needed distinction between Case-chain traces and wh-traces.

However, this may not be the correct generalization. In this theory, we have another generalization available: that intermediate traces in a Case-chain allow wanna-contraction—as in the relation between (31a) and (31b) and in (31c). On the other hand, the head of a Case-chain blocks wanna-contraction: in (31d), the position between want and to in SpecIP is both the tail of a wh-chain but also the head of a Case-chain, and SpecIP is the position of the head of the Case-chain at Spell-Out. Of course, the Spell-Out position of constituents is the position relevant to PF interpretation, and in PF it is determined how chains are pronounced. PF certainly treats the head of a Case-chain differently from its other positions. Therefore, we argue that wanna-contraction is explained because the head of a Case-chain, even if void of phonological material, has enough content at PF to prevent wanna-contraction.

2.7 Conclusion
In this chapter we have proposed a theory of theta role assignment at LF. This theory, along with independently motivated requirements of movement and feature checking at
LF in the syntax, is used to derive the properties of PRO—in particular, we can account for raising and obligatory control constructions in English. If PRO is redundant, then Control Theory may be trivially eliminated.

In particular, we looked at six broad classes of verbs in this chapter. The following table summarizes the data which has been considered:

<table>
<thead>
<tr>
<th></th>
<th>no object</th>
<th>object</th>
<th>passive</th>
</tr>
</thead>
<tbody>
<tr>
<td>I seemed to leave.</td>
<td>*I seemed Pat to leave.</td>
<td>*Pat was seemed to leave.</td>
<td></td>
</tr>
<tr>
<td>*I believed to leave.</td>
<td>I believed Pat to leave.</td>
<td>Pat was believed to leave.</td>
<td></td>
</tr>
<tr>
<td>I tried to leave.</td>
<td>*I tried Pat to leave.</td>
<td>*Pat was tried to leave.</td>
<td></td>
</tr>
<tr>
<td>I wanted to leave.</td>
<td>I wanted Pat to leave.</td>
<td>*Pat was wanted to leave.</td>
<td></td>
</tr>
<tr>
<td>*I persuaded to leave.</td>
<td>I persuaded Pat to leave.</td>
<td>Pat was persuaded to leave.</td>
<td></td>
</tr>
<tr>
<td>I promised to leave.</td>
<td>I promised Pat to leave.</td>
<td>*Pat was promised to leave.</td>
<td></td>
</tr>
</tbody>
</table>

These verbs are characterized by different Case, theta and selectional features:

- Raising verbs like *seem* do not assign a theta role (other than to their IP complement), select a nonfinite IP complement and do not have Case-features. As we see in (33), they cannot take objects because they do not check Case, and cannot be passivized because they have neither Case-features nor the appropriate theta role. However, raising to subject is grammatical.

- ECM verbs like *believe* assign one theta role to a DP (and one theme theta role to an IP), select a nonfinite IP complement (with null tense features) and have Case features. In (33), subject raising (which for this verb would be the same as control) is ungrammatical for the reasons given in §2.3 and to be elaborated on in chapter 3. However, raising to object is grammatical and its Case features allow it to be passivized.

- Subject control verbs like *try* assign a theta role (and another to their IP complement),
select a nonfinite IP complement and do not have Case features. Their behavior as control verbs was discussed in §2.3. Without Case features, they cannot take an object and also cannot be passivized.

- Subject control verbs like *want* with optional ECM assign a theta role (and a theme theta role to their IP complement), select for a nonfinite IP complement and have Case features. With Case features they can act either as a subject control verb or as an ECM verb, as seen in §2.3. However, this verb cannot be passivized.6

- Object control verbs like *persuade* assign two theta roles (and another theme theta role to their IP complement), select for a nonfinite IP complement, and have Case features. In (33), they cannot be detransitivized because they lack the selection feature for a DP (as discussed in §2.5), but they do check Case on objects and can be passivized. Finally;

- Subject control verbs like *promise* assign two theta roles (and another theme theta role to their IP complement), select for two complements—a DP and a nonfinite IP complement—and have Case features. Their selection features allow them to be detransitivized, but the detail of their structure, as discussed in §2.4, make it impossible for them to be passivized.

These combinations of features, interacting with the grammatical principles introduced in §2.2, account for the variation we have seen in these verbs: whether they are raising or control verbs, whether they take objects or not and whether they can be passivized or not.

However, in this chapter we have considered only a subset of cases of obligatory control. Building on the work in this chapter and in chapter 1, we will now consider other constructions in English in chapter 3, and “control” constructions in other languages in chapter 4. Non-obligatory control constructions are examined in chapter 5. Assuming that

---

6And to be perfectly honest, I do not have an explanation for the impossibility of passivization with *want*. One possibility is that, when *want* has an object, it is not an ECM verb like *believe*, but a subject control verb like *promise*. If this analysis is true, *want* would be impossible to passivize as *promise* is. How likely this possibility is, I leave to the reader to decide.
the results presented in this chapter hold up, then the need for a separate Theory of Control has been eliminated, and by reducing a stipulated module of the grammar to independently motivated interface conditions, the results of this paper simplify the grammar.