
**Keywords:** microparametric syntax, minimalism, Binding Theory and PRO, I-to-C movement, quantifier scope and word order

1. Introduction

Richard Kayne's *Parameters and Universals* is a collection of articles that he has written in the past fifteen years. The main theme of the book is to examine how best to characterize the properties of human language that are not universal. Referring to the investigation into a large number of closely related languages as 'microparametric syntax,' Kayne suggests as follows: works in microparametric syntax will play a privileged role in the future in answering the question of the form that syntactic parameters may take and the question of how many irreducible parameters there really are, and they will also give us a powerful tool in uncovering the nature of the invariant principles of UG.

On the basis of this foundation, Kayne organizes this book by dividing his fifteen articles into three subparts: Part I consists of eight articles on Romance languages, Part II consists of three articles on some varieties of English, and Part III consists of three articles that open up...
a new approach to UG by focusing on quantifier scope, prepositional complementizers that select IP, and word order universals relevant to prepositions and complementizers, respectively.

This book is organized as follows:1 Chapter 1, "Microparametric Syntax: Some Introductory Remarks" [1996], argues for the importance of 'microparametric syntax' in developing the theory of parameters and universal grammar. Chapter 2, "Past Participle Agreement in French and Italian" [1985], argues that past participle agreement in Romance languages is the local relation between the functional head AGR and its Spec. Chapter 3, "Facets of Romance Past Participle Agreement" [1989], proposes an explanation of the absence of past participle agreement in the impersonal construction, in the ECM construction, and when object NPs undergo wh-movement in some languages. Chapter 4, "Null Subjects and Clitic Climbing" [1989], explains the correlation between the (im)possibility of clitic climbing and the (im)possibility of null subjects within Chomsky's (1986b) Barriers framework. Chapter 5, "Romance Clitics, Verb Movement, and PRO" [1991], shows the generalization that control with if is permitted in a Romance language if and only if infinitive verbs can precede clitics in the language, and proposes an explanation of the generalization by extending Chomsky's (1986a) Binding Theory to PRO. Chapter 6, "Italian Negative Infinitival Imperatives and Clitic Climbing" [1992], proposes an account of the fact that in Italian, infinitive verbs must precede clitics, whereas verbs in the negative infinitive imperative construction can follow clitics. Chapter 7, "Toward a Modular Theory of Auxiliary Selection" [1993], extends to auxiliary have Freeze's (1992) analysis of possessive have as be+incorporated preposition. Chapter 8, "Person Morpheme and Reflexives in Italian, French, and Related Languages" [to appear], analyzes the morphological structure of Romance clitic and nonclitic pronouns. Chapter 9, "A Note on Clitic Doubling in French" [to appear], provides a new analysis of a certain restriction on French nonclitic personal pronoun in terms of a requirement on structurally Case-marked pronouns. Part I is closed here.

Part II consists of Chapter 10, "Notes on English Agreement" [1989], Chapter 11, "Agreement and Verb Morphology in Three Varieties of

---

1 I will refer to the articles in the volume by the chapter number with the year of the original publication in square brackets.
English” [1995], and Chapter 12, “The English Complementizer Of” [1997]. Starting with the argument that English -s form on verbs is only marked for number (rather than person), the first two chapters discuss the lack of wh-agreement in standard English and a certain restriction on wh-agreement in nonstandard English. Chapter 12 shows that in some varieties of English the preposition of has a usage which is strongly similar to the infinitive to. Assuming that to in for-to infinitives are complementizers, Kayne concludes that of is a kind of complementizer, too.

Part III develops three proposals about the architecture of universal grammar, which are closely related to one other. Chapter 13, “Overt Versus Covert Movement” [1998], proposes that every operator has its scope fixed by overt phrasal movement rather than LF movement. Chapter 14, “Prepositional Complementizers as Attractors” [1999], proposes that prepositional complementizers, such as de in French, di in Italian, and to in English, are not directly merged with the infinitive IP which they select but are merged above VP and attract the IP to their Spec position. Chapter 15, “A Note on Prepositions, Complementizers, and Word Order Universals” [a slightly revised version of a paper contributed to the Chomsky Virtual Celebration], is an attempt to extend his proposal in Chapter 14 to account for certain word order universals concerning prepositions and complementizers, proposed by Greenberg (1963) and Dryer (1992) (see also Kayne (2001a) and Kayne (2002) for an extension of the same idea). Behind the three proposals in Part III lies Kayne’s idea that certain selectional relations which appear to be satisfied by merger should be reanalyzed as being satisfied by the overt movement of the selected category to the Spec of the selecting category.

Throughout this book, we can see the insightfulness behind his combination of apparently unrelated phenomena across languages to propose cross-linguistic generalizations. His presentation of linguistic data in favor of a set of hypotheses is also convincing. However, the most salient factor which makes his arguments persuasive is his attitude to compare only closely related languages, such as French and Italian, rather than directly comparing languages which are unrelated either diachronically or synchronically, such as French and English. This is a version of comparative syntax which he calls the ‘microparametric syntax.’

Comparative syntax reveals not only the extent to which languages
are different but also what they have in common. The common properties should be captured by the principles of UG and differences by parameters. In this sense, comparative syntax makes a huge contribution to the development of the theory of UG as well as a proper understanding of the nature of parameters. Kayne points out, however, a methodological problem with the direct comparison of typologically unrelated languages. He argues as follows: ‘in searching for clusters of properties, one must make decisions about what syntactic differences can plausibly be linked to what other syntactic differences’ and in so doing, ‘the more syntactic differences there are to begin with, the harder it will be, all other things being equal, to figure out the correct linkings.’

He is referring to the matter of falsifiability. Generative syntax is a set of scientific hypotheses which must be falsifiable. The more easily a hypothesis could be falsified, the sooner we could propose an alternative hypothesis that overcome problems with the earlier one. A hypothesis of the form ‘these differences between the two languages are related to each other’ could be tested by examining a third language to see if the particular properties in question actually do systematically cluster together. And such a test is feasible more easily if there are fewer differences among the three languages, in which case there is less room to take into account other factors that might affect the grammaticality of relevant data. In this sense, the hypothesis testing task generated by the French-Italian comparison is more manageable than the one generated by the French-English comparison.

Thanks to its methodological advantage, Kayne’s microparametric syntax has uncovered a number of cross-linguistic generalizations that should be explained in any version of generative grammar and contributed to its development in the last fifteen years. Thus, no one would deny that the Agr-based Case Theory (which has been generalized to the Feature Checking Theory) is a notable result of Kayne’s investigation into the past participle agreement in Romance languages (as well as Pollock’s (1989) analysis of the French verb movement). The feature checking has been one of the most powerful tools to explore the nature of phrase structure and movement in the minimalist guidelines. For these reasons, we can safely conclude that the microparametric syntax has made considerable contributions to the development of the theory of UG up to the minimalist program.

We may note here that, although any theory of UG must meet ex-
planatory adequacy, the minimalist program of linguistic theory and Kayne's theory differ in their way to pursue it. The minimalist program has virtually limited the set of available syntactic principles and devices to those motivated on conceptual (i.e. theory-independent) necessity. Such a conceptually restricted approach is forced to limit the set of relevant linguistic data to a small number and put aside many important syntactic generalizations as peripheral phenomena that are irrelevant to the core computational system of human language. By contrast, Kayne seems to be tacitly assuming that a theory of UG can postulate as many syntactic principles and devices as are necessary for explaining the descriptive generalizations revealed through the study of microparametric syntax. One might argue that this approach tends to broaden the set of accessible syntactic principles and primitives that are void of conceptual motivations. However, this kind of criticism against Kayne's theories is unjustified in most cases. As the minimalist program just seeks to elucidate how close to virtual conceptual necessity the architecture of language is, it should tolerate departures from conceptual necessity if they are motivated on sufficient empirical grounds and if there is no alternative solution that is obviously more desirable on conceptual grounds. For these reasons, it will be a reasonable research strategy to try to reinterpret Kayne's insights and findings under the set of assumptions and devices admitted as minimally necessary in the minimalist conceptions and examine what aspect of his proposals is unjustified on empirical or theoretical grounds, case by case.

Space limitation prevents me from reviewing every proposal in the fifteen articles. Therefore, I will pick up and discuss only three issues that have a direct relevance to the minimalist approach to linguistic theory: (i) Kayne's account of the correlation between null subjects and clitic climbing within the Barriers framework (Chapter 4), (ii) his explanation of the (im)possibility of control with if within Chomsky's (1986a) Binding Theory, based on the notions of S-structure and government (Chapter 5), and (iii) his reanalysis of quantifier raising as overt phrasal movement (Chapter 13).

2. Clitic Climbing and the Null Subjects

In this section, I will critically review Kayne's proposals on clitic climbing and null subjects made in Chapter 4 and propose an alterna-
tive solution, exploiting his proposals in Chapter 9.

2.1. Kayne’s Proposals

The pronominal objects in Romance languages can appear in a position between a canonical subject position and a canonical object position. This phenomenon, called ‘cliticization,’ necessarily takes place clause-internally in French. However, the object clitics can be moved out of certain infinitive complements in Italian. This nonlocal cliticization is called ‘clitic climbing.’

(1) a. *Jean les veut voir. <French>
b. Jean li vuole vedere. <Italian>

Kayne proposes a cross-linguistic generalization on clitic climbing, as stated in (2):

(2) Clitic climbing is permitted only in a Romance language that licenses null subjects.

He explains (2) within Chomsky’s (1986b) Barriers framework. His argument proceeds as follows: For a clitic to move out of an infinitive complement (=IP), it must be able to attach to the infinitive I. For a clitic to attach to I, it must move out of VP. Since VP is a potential barrier to antecedent government, a clitic can move out of VP only if the barrierhood of the VP is voided by L-marking. I can L-mark VP if I is strong enough to license null subjects or if a lexical V moves and attaches to I. In Italian, which licenses null subjects, the infinitival I is strong enough to L-mark VP. Hence, clitic climbing is permitted. On the other hand, in French, null subjects are not licensed and the infinitival V does not raise to I, either. Hence, the infinitival I does not L-mark VP and French does not permit clitic climbing.

Along these lines of reasoning, we predict that, in a null subject language, the infinitival V does not have to raise to I in order to license clitic climbing. As expected, in Occitan, a null subject language, the infinitival V does not raise to I (though it does raise to a functional head between IP and VP; cf. Kayne (2000: 66)), but clitic climbing is permitted.

If the weak I (in the sense of being unable to license null subjects) blocks clitic climbing, clitic climbing should be possible out of a nonfinite complement which lacks the projection of I, even in French. This situation is observed in certain causative constructions as below:
(4) Jean la fait manger par/à Paul.
John it-makes eat by/to Paul

In (4), the complement of the causative V is VP rather than IP. The VP is L-marked by the matrix V. Hence, (4) is well-formed even if French infinitive verbs do not raise to I.

Moreover, the proposed account makes the prediction that even if I is strong enough to L-mark its complement, clitic climbing should be blocked if a functional category which does not L-mark VP intervenes between IP and VP. With this in mind, consider the following Italian examples:

(5) a. Gianni non li vuole vedere.
Gianni NEG them wants to-see

b. *Gianni li vuole non vedere.

In (5b), the infinitival complement contains a negative element non, which arguably heads NegP between IP and VP. I L-marks NegP but Neg does not L-mark VP. Hence, VP is taken as a barrier for clitic climbing and (5b) is ruled out. (5a) is ruled in because the NegP here resides above the landing site of clitic climbing.

In sum, Kayne exploits the notion of L-marking and argues that clitic climbing is available only in null subject languages, in which the strong I can L-mark its complement.

2.2. A Note on Parameters on the Functional Head ‘I’

If a set of languages manifests two or more properties and another set of languages manifests none of them, we are tempted to propose that the presence or absence of the two properties should be analyzed as a result of a particular setting of the value of a single parameter, because, other things being equal, a linguistic theory that assumes only one parameter is to be preferred to the one that assumes two or more parameters for the sake of simplicity. In fact, localization of apparently unrelated phenomena to a single parameter is one of the central aims of comparative syntax, of which Kayne’s microparametric syntax is a version. However, it must be examined whether the particular localization is both theoretically natural and empirically justified. In what follows, I will point out two arguments against his implementation on theoretical grounds.

First, Kayne has not actually localized the two phenomena to a single parameter (of I); rather, his proposal amounts to saying that there are two different parameters one of which has its value fixed dependent on
the value of the other. Since the possibility of null subjects in a finite clause is related to the richness of person agreement on finite verbs (Platzack and Holmberg (1989)) and since the infinitive verbs do not manifest person agreement at all, it is implausible to assume that the infinitive I can license pro, which we are referring to as ‘null subjects.’ PRO in (6) should be distinguished from pro:

(6) Jean li vuole [PRO/*pro vedere]. (= (1b))

Given the distinction between PRO and pro, Kayne’s claim actually implies that, if a finite I can license pro, an infinitive I can L-mark its VP complement, even if it does not license pro itself. However, there is no a priori reason why some property of the finite I and a different property of the infinitive I must be correlated in the way as Kayne proposes.

Second, there are several pieces of evidence showing that parametric values of the finite and infinitive Is for a single phenomenon in a single language are fixed independently. Thus, in French, V-raising to I is obligatory in finite clauses but impossible in infinitive complements (Pollock (1989)). Even different instances of infinitive Is in a single language may have different properties for a single phenomenon. Thus, in Icelandic, V-raising to I is possible in control complements but not in ECM or raising complements (Watanabe (1993: 37–38)).

For these reasons, it seems implausible, if not impossible, to assume that the finite and infinitive Is in a single language are related to each other in such a way that if the finite I can license null subjects, the infinitival I can L-mark its complement. In order to fully explain (2), we need to argue that null subjects and clitic climbing share a common property that can be localized to the category I (which covers both the finite and infinitive Is). Recall here that a necessary condition for the availability of clitic climbing is the possibility of the clitic to adjoin to the infinitive I. With this as presupposed, let us consider why adjunction of the clitic to the infinitive I is possible only in languages in which null subjects are licensed.

An idea that might answer this question has already been suggested on independent grounds in a different chapter of this book. Note first that the nonclitic personal pronouns in French can occur in the complement of a preposition but not a verb (there is no such restriction on lexical NPs):

(7) a. Jean parles de moi. (‘Jean speaks of me.’)
   b. *Jean connaît moi. (‘Jean knows me.’)
(7b) is saved if clitic doubling takes place:

(8) Jean me connaît moi. (cf. Jean (*la) connaît Marie.)

These facts suggest that pronominal arguments in French are subject to the following condition:

(9) Pronominal arguments that are structurally case-marked in French must be doubled by a clitic.

With this in mind, consider (10) and (11):

(10) a. Jean a parlé à tous. (‘Jean has spoken to all.’)
    b. *Jean a invité tous. (‘Jean has invited all.’)

(11) Jean les invitera tous. (‘Jean them will-invite all.’)

(10) and (11) show that the same pattern that is found in (7) and (8) is observed when the nonclitic personal pronoun is replaced by the quantifier tous. In fact, the object of à in (10a) is limited to third person interpretation. In consideration of these facts, Kayne proposes that the object of invité in (10b) is not simply tous but tous plus a covert nonclitic pronominal, where the latter is a third person pronominal argument that is subject to (9). This proposal can accommodate the third person restriction observed in the following sentences in French:

(12) a. Tous chantent. (‘all were-singing-3pl.’)
    b. *Tous chantiez. (‘all were-singing-2pl.’)

Suppose that the subject arguments in (12a, b) are not simply tous which itself has a particular inflectional specification for person and number but tous plus a covert nonclitic pronominal. Then, (12b) will be ruled out because of the mismatch between the third person specification of the covert nonclitic pronominal and the second person agreement suffix on the verb.

The question arises here as to how general this limitation to third person is. Kayne postulates that (13) holds at least in French, Italian, and related dialects (p. 175):

(13) Covert nonclitic pronouns can only be third person.

This means that we can no longer assume that Italian (14) involves first person empty pronominal subjects in the canonical subject position:

(14) Parlavamo. (‘we-spoke’)

Hence, Kayne (2000: Ch. 9) proposes that what appears to be the agreement suffix -no must be the (postverbal) realization of a pronominal subject.²

² Kayne suggests that the third person agreement suffix -no in (i) must be non-
I, this analysis of (14) implies that in Italian, the pronominal subjects can be adjoined to I. On the other hand, the ill-formedness of (12b) indicates that the pronominal subjects in French cannot be adjoined to I as agreement suffixes. If tout in (12b) were merged not with the third person covert nonclitic pronoun but with the agreement suffix -iez which would be moved and adjoined to I later in the derivation, then (12b) would be ruled in for the same reasons as (14), because there would be no mismatch in agreement in this analysis. We therefore conclude that what distinguishes between Italian (a null subject language) and French (a non-null subject language) is whether or not a pronominal subject can be adjoined to I as an agreement suffix.

Returning to clitic climbing, it was noted above that a necessary condition for an object clitic to move out of an infinitive complement is that it must be able to attach to infinitive I. Now, assuming that clitics are a kind of pronoun, we find a parallelism between null subjects and object clitics. In these terms, the generalization in (2) can be restated as in (15):

(15) Object pronouns can be adjoined to I only in a Romance language in which subject pronouns can be adjoined to I.

It is reasonable to assume that pronouns are of the category D in Abney’s (1987) sense rather than the lexical category N (Longobardi (1994)) and that adjunction of D to I takes place for the sake of the (strong) D-feature checking of I. It is important to note here that there is no distinction between finite and infinitive Is with respect to D-feature checking (Chomsky (1995)). On these assumptions, (15) follows from a particular setting of a single parameter which can be stated as follows:

(16) I can/cannot check its (strong) D-feature with an element merged with its minimal projection.\(^3\)

pronomin and that (i) must contain a covert nonclitic third-person pronoun (pro) in the canonical subject position:

(i) Parlavano. (‘they-spoke’)

For reasons why an agreement suffix having the properties of a pronoun can only be first or second person, see Kayne (2000: 183, note 45).

\(^3\) When the clitic is adjoined to the I of the controlled infinitive, the I in question has its D-feature checked against the clitic and the PRO. This is an instance of multiple D-feature checking. Hence, the text proposal implies that in languages that permit clitic climbing, the I must permit multiple D-feature checking.
If the parametric value is positively specified, both the D head merged with the light verb v (the subject pronoun) and the D head merged with V (the object pronoun) can be remerged with I, generating null subjects and clitic climbing, respectively, whereas if the parametric value is negatively specified, no D can be remerged with I and neither of them is permitted.

This line of reasoning arguably provides us with a more natural way to understand the nature of the correlation between null subjects and clitic climbing. Another advantage of our solution is that it does not refer to the notions of L-marking and barrierhood, which have no obvious conceptual necessity in the minimalist framework.

3. Control with *If* and Binding Theory

In Chapter 5, Kayne shows the generalization that control with *if* is permitted in a Romance language if and only if infinitive verbs can precede clitics in the language. He proposes to explain the generalization in terms of Chomsky's (1986a) Binding Theory and Chomsky's (1986b) minimality. In this section, I will first outline his proposals and discuss some problems and possible solutions.

3.1. Kayne's Proposals

Although English permits infinitive interrogative clauses, there is a notable contrast as in (17):

(17) a. He doesn't know whether to go to the movies.
    b. *He doesn't know if to go to the movies.

The contrast is not observed when *whether* or *if* appears in finite clauses, as in (18):

(18) a. He doesn't know whether he should go to the movies.
    b. He doesn't know if he should go to the movies.

---

4 The acceptability of the following sentences in French appears problematic at first sight:

(i) Jean me/te/se voit. (‘Jean me/you/refl. sees’)

We assume that clitics in these examples are adjoined not to I but to the V adjoined to I. See Kayne (1994: 42–46) for the same conclusion on totally different grounds.

5 What remains to be solved is the question why the intervening Neg blocks clitic climbing.
If we assume that _whether_ is the _wh_-counterpart of _either_ that occupies [Spec, C] (Katz and Postal (1964)), whereas _if_ is a lexical complementizer that heads CP, the ill-formedness of (17b) can be attributed to the PRO Theorem:

(19) The PRO Theorem:

PRO must be ungoverned.

The structures of (17a) and (17b) are schematized in (20a) and (20b), respectively:

(20) a. He doesn’t know [CP whether [C C(ϕ) [IP PRO [I to go to the movies]]]]

b. He doesn’t know [CP C(if) [IP PRO [I to go to the movies]]]

Assuming that a lexically filled C⁰ counts as a governor of PRO but a nonlexical C⁰ position does not, we can rule out only (17b) as a violation of (19).

Looking at Romance languages, we find the following contrast:

(21) a. *Marie ne sait pas si aller au cinéma. <French>

   ‘Marie does not know if to go to a movie.’

b. Gianni non sa se andare al cinema. <Italian>

Since there is good reason to believe that French _si_ and Italian _se_ are the exact counterparts of English _if_ rather than _whether_, (21) shows that the counterpart of (17b) is ruled out in French but not in Italian. In finite clauses, there is no such contrast, and both (22a) and (22b) are well-formed:

(22) a. Marie ne sait pas si elle devrait aller au cinéma.

   ‘Marie does not know if she should go to a movie.’

b. Gianni non sa se dovrebbe andare al cinema.

Since the PRO Theorem (19) should hold universally, the ill-formedness of (21a) follows from (19). The problem is: why is (21b) not ruled out by (19)?

Kayne examines Romance languages other than French and Italian and finds out that Occitan and Sardinian pattern with French, whereas Catalan and Spanish pattern with Italian. Since all the languages except for French are null subject languages, being a null subject language is not a sufficient condition for permitting control with _if_.

---

⁶ In French and Italian, there is no single word for _either_.
Rather, the possibility of control with *if* is correlated with the order between infinitives and clitics. Consider the contrast between French (23) and Italian (24):

(23) a. Lui parler serait une erreur.
   to-speak himDAT would-be an error
   b. *Parler-lui serait une erreur.

(24) a. Parlargli sarebbe un errore.
   to-speak-himDAT would-be an error
   b. *Gli parlare sarebbe un errore.

Examining four more dialects of Northern Italy, Kayne presents the following generalization:

(25) If control with *if* is permitted in a Romance language, the language shows infinitive-clitic order.

Kayne explains this generalization by assuming Chomsky’s (1986a) Binding Theory, Chomsky’s (1986b) minimality, and the proposal that the infinitive-clitic order is derived only if the verb is adjoined to some X’. Given the proposal, partial structures of the French example in (23a) and the Italian example in (24a) are schematized in (26a) and (26b), respectively, where Infn is the functional category that lies between TP and VP (the label is the abbreviation of ‘Infinitive’):

(26) a. ... V [CP C (si) [TP PRO [T' [Cl+[[V+Infn] T]] ... [Infn e] ... [VP [v e] ... 
   b. ... V [CP C (se) [TP PRO [T' [V+Infn] [T' [Cl+T] ... [Infn e] ... [VP [v e] ...

In (26a), since *si* governs PRO, the PRO Theorem accounts for the ill-formedness of (21a). In (26b), on the other hand, the infinitive (=V+Infn) adjoined to T’ blocks off government of PRO by *se*, because it induces a minimality barrier in the sense of Chomsky (1986b: 10). Therefore, the potential PRO Theorem violation does not emerge and (21b) is ruled in.

Why is it that the infinitive adjoined to T’ in (26b), being a closer governor for PRO, does not itself induce a PRO Theorem violation? Kayne replies to this question by applying Chomsky’s (1986a) Binding Theory (henceforth, the KL theory) to PRO. The basic idea that lies

---

7 Strictly speaking, a revision is needed on the definition of ‘minimality barrier’ given in Chomsky (1986b: 42). See Kayne (2000: 82) for a relevant discussion.
behind the KL theory is that the Governing Category (GC) for pronouns, which do not have to be bound, are not necessarily equated with and can be smaller than the GC for anaphors, which must be bound in a local domain. A support for the KL theory is that, in conflict with Chomsky’s (1981) Binding Theory, anaphors and pronouns sometimes do not manifest a complementary distribution:

(27) a. The children like [each other’s friends].
    b. The children like [their friends].

The KL theory accounts for the well-formedness of (27a,b) as follows: in (27a), the entire clause, which is the smallest domain in which it could satisfy the binding theory with some indexing, is qualified as the GC for each other, in which it satisfies BT (A). In (27b), the object NP, which is the smallest domain that contains the pronoun, its governor, and a subject (=the pronoun itself), is taken as the GC for their, in which it satisfies BT (B).

Kayne extends the KL theory to PRO, which he assumes to be a pronominal anaphor. It follows that the PRO in (26b) may be governed by the infinitive verb. His argument is as follows: Since the PRO is governed by the V adjoined to T', the TP in the embedded clause is taken as the GC for PRO qua pronoun, in which it satisfies BT (B). Since the GC for PRO qua anaphor is the next larger TP, it also satisfies BT (A). Since PRO in (26b) satisfies both BT (A) and BT (B), it may be governed. In (26a), on the other hand, PRO is governed by the lexical C0. In this case, the GC for the PRO qua anaphor and the GC of PRO qua pronoun are both fixed as the matrix TP, in which the PRO cannot be bound and free simultaneously. Hence, (26a) is illegitimate.

Kayne’s approach to PRO has an advantage over Chomsky’s (1981), which predicts that PRO is always ungoverned. If PRO qua anaphor is ungoverned, it cannot have a GC, and hence it cannot satisfy BT (A). This implies that the antecedent of PRO is not syntactically predictable. On the other hand, in Kayne’s approach to PRO, the GC for PRO qua anaphor in (26b) is fixed as the next larger TP. As predicted, in (28), Gianni but not Maria can be the antecedent of PRO:

(28) Maria pensa che Gianni non sappia se andare al cinema.

In the French example in (29), PRO appears not to be governed by
either the lexical $C^0$ or the infinitive, which does not raise overtly to $I$. Therefore, as its stands, we cannot determine its antecedent:

\[(29) \text{ Jean veut [PRO aller au cinema].}\]

Jean wants to go to the movies

Kayne solves this problem by proposing that even in French (and English, too), the infinitive moves and adjoins to $T'$ at LF, as a result of which the antecedent of PRO qua anaphor is determined in the same way as it is in Italian:

If the infinitive adjoins to $T'$ at LF in French, the PRO Theorem violation in (26a) might be canceled and (21a) should be ruled in.

\[(21) \text{ a. *Marie ne sait pas si aller au cinéma.}\]

Kayne solves this potential problem by proposing that a PRO Theorem violation given at S-structure cannot be neutralized at LF, because a given indexing must respect BT (B) at all levels. In other words, although (21a) satisfies BT (B) at LF, it violates BT (B) at S-structure, which is the reason why (21a) is ill-formed. Kayne contrasts this with the way BT (A) is satisfied in (29). BT (A) is satisfied here at LF, but not at S-structure. This does not cause a problem because BT (A) can be met at some level of representation (Belletti and Rizzi (1988)). Kayne suggests that BT (A) has an intrinsically existential character (that is, for a given anaphor, there must be some antecedent at some level), whereas BT (B) has an intrinsically universal character (that is, a given pronoun must be free from all antecedents at all levels).

3.2. Binding Theory as LF Condition without Government

Kayne’s proposal that PRO qua pronominal anaphor can be governed without violating either BT (A) or BT (B) is theoretically important in that it enables us to determine its antecedent in the same way as overt reflexives that are subject to BT (A). Also, I would like to respect his idea that the illegitimate examples of control with if are ruled out by BT (B) for reasons noted below. His proposals, however, are incompatible with a minimalist perspective in that they adopt the binding theory that depends on the notion of ‘government’ and that applies at S-structure (as well as LF). In this subsection, I will demonstrate that BT (A) and BT (B) should apply exclusively at LF, and in the next subsection, I will provide an alternative solution that maintains his empirical coverage.

In the minimalist program of linguistic theory (Chomsky (1993, 1995)), the faculty of human language is taken to be an optimal solu-
tion to the conditions imposed on the interface levels of PF and LF. Along these lines, D-structure (DS) and S-structure (SS) have been eliminated on the ground that they do not satisfy (virtual) conceptual necessity. Accordingly, conditions which were taken to apply at SS have been remodeled as conditions on derivation or conditions at the PF/LF interface.

On Binding Theory, Chomsky (1995: Ch. 3) provides a strong argument that BT (A) and BT (B) apply at LF but not at SS. First, consider the following sentence:

(30) John wondered which picture of himself Bill took.

Under the copy theory of movement, (30) is mapped onto either of the following two LF representations, depending on which subpart of the moved wh-phrase undergoes 'reconstruction':

(31) a. John wondered [which x, x a picture of himself] [Bill took x]

b. John wondered [which x] [Bill took [x picture of himself]]

When the option (31a) is chosen, himself takes John as antecedent by BT (A) applying at LF. When the option (31b) is chosen, himself takes Bill as antecedent by the same principle. Consequently, himself can be anaphoric to John or to Bill and (30) is at least two-way ambiguous. Chomsky (1995: 206-207) observes, however, that there is a further ambiguity only when himself is anaphoric to Bill: the phrase take ... picture can be interpreted idiomatically (in the sense of 'photograph') or literally ('pick up and walk away with') when himself is anaphoric to Bill, though the idiom interpretation disappears when himself is anaphoric to John. Since it is reasonable to assume that the idiom interpretation is available only if take ... picture is present as a unit at LF, only (31b) is the possible input for the idiom interpretation. Suppose now that BT (A) has an intrinsically existential character. Then, himself in (30) could be locally bound by John and satisfies BT (A) at SS, and the later LF process would choose either the idiomatic or the literal interpretation freely. As a result, (30) should be ambi-

8 See Fox (2000) and references therein for evidence that BT (C) does not apply at SS but at LF.
9 See Chomsky (1995: 206-212) for a more detailed argument for the copy theory.
guous even when *himself* is anaphoric to *John*. In order to avoid this undesirable conclusion, it must be concluded that BT (A) applies only at LF.

Next, consider the following sentence:

(32) John wondered which picture of him Bill took.  

(Chomsky (1995: 208))

In (32), *Bill* cannot be the antecedent of *him*, whether *take ... picture* can be interpreted idiomatically or literally. This means that BT (B) forces the ‘reconstruction’ that leads to the following LF representation:

(33) John wondered [which x] [Bill took [x picture of him]]

In (33), BT (B) is violated because *him* is bound by *Bill* in the same IP. It might be important to ask why the ‘reconstruction’ is forced in (32). However, what concerns our discussion is not this question but the fact that *John* can be the antecedent of *him* in (32). If BT (B) applied at SS, where *John* c-commands *him* in the same IP, the coreference should be impossible. Hence, the availability of this interpretation in (32) indicates that BT (B) applies only at LF.

We have seen that BT (A) and BT (B) apply only at LF. This conclusion is in conflict with Kayne’s proposal on Binding Theory. Recall his suggestion that BT (A) has intrinsically existential character, whereas BT (B) has intrinsically universal character. The hypothesis that BT (A) can be satisfied at some level of representation can be innocuously replaced by the one that BT (A) applies only at LF if appropriate LF representations are given to sentences involving anaphors. However, the hypothesis that BT (B) must be satisfied at all levels of representation makes predictions different from the hypothesis that BT (B) applies only at LF. Specifically, (32) should be ruled out when *John* is the antecedent of *him*, if BT (B) applied at SS. But if BT (B) applies only at LF and if the infinitive *I* raises covertly to *I* in (21a), (21a) will be predicted to be well-formed.

(21) a. *Marie ne sait pas si aller au cinéma.*

An independent problem with Kayne’s approach to PRO is that it crucially depends on the notion of ‘government,’ which has been abandoned in the minimalist program, due to the lack of conceptual necessity. Recall that the GC for PRO qua pronoun in (21a) is fixed as the matrix IP because the lexical CO (=si) that governs PRO is outside the embedded IP. If we simply remove the notion of ‘governor’ from the definition of BT (B), the GC for PRO qua pronoun will be the embed-
ded IP in (21a). Hence, Kayne’s theory crucially depends its formulation on the notion of ‘governor.’ For these reasons, Kayne’s approach to the contrast between (21a) and (21b) cannot be maintained as it is. We need a different approach to explain the generalization (25). In the next subsection, I will explore such an alternative solution.

3.3. I-to-C Movement and Governing Category

As the first approximation, let us formulate BT (A) and BT (B) as follows:

(34) a. An anaphor must be bound in the smallest domain which contains it and a potential antecedent for it (the relevant domain = IP or DP).

b. A pronoun must be free in the smallest domain which contains it and a subject.

Given (34a, b), the domain in which anaphors must be bound is not equated with the domain in which pronouns must be free. Hence, complementary distribution no longer need to hold between anaphors and pronouns. This is a desirable consequence in view of the acceptability of (27a, b). However, it does not follow from (34) that PRO is excluded from the subject position of a finite clause (Kayne (2000: 97, note 80)):

(35) *I thought [(that) [PRO would win the race]].

In order to exclude (35), Chomsky (1995: 119–120) proposes that PRO, as well as overt NPs, bears an abstract Case, called ‘Null Case,’ and that it must be checked in the Spec of the infinitival element (with null agreement) and the head Ing of gerundive nominals (cf. also Martin (2001)).

(36) a. PRO to VP (to be sick)

b. PRO Ing VP (being sick)

The Null Case theory has a number of welcome results. First, coupled with the principle of Last Resort, it excludes examples in which PRO moves from a Case position to the Spec of infinitive I. Second, it can simplify the formulation of the Visibility Condition.10 Third, it can

---

10 The simplified formulation is that a chain is visible for θ-marking if it contains a Case position. If PRO does not have Case, it must be formulated disjunctively as follows: a chain is visible for θ-marking if it contains a Case position or is headed by PRO.
dispense with the assumption that the control complement is CP headed by the nonlexical \( C^0 \). Rather, it can be bare IP (see Bošković (1995) for relevant arguments). Suppose that the control complement is IP, and let us see how (34a, b) work to provide the appropriate interpretation for (37) (=\((29)\)):

\[
(37) \quad \text{[ip Jean veut [ip PRO aller au cinema]]}
\]

In (37), the GC for PRO qua pronoun is the embedded IP, in which it is free; the GC for PRO qua anaphor is the matrix IP, in which it is bound by Jean. Hence, its antecedent is determined as Jean.

So far, no reference to the notion of government is necessary. We can also dispense with Kayne's nonstandard assumption that the infinitive V (a minimal projection) can adjoin to I' (a nonminimal projection).\(^{11}\)

With this in mind, let us next consider the contrast between (21a) and (21b), keeping it in mind that what distinguishes (21b) from (21a) is only whether the infinitive V overtly raises to I:

\[
(21) \quad \begin{align*}
\text{a.} & \quad *\text{Marie ne sait pas si (PRO) aller au cinéma.} \quad \text{<French>} \\
\text{b.} & \quad \text{Gianni non sa se (PRO) andare al cinema.} \quad \text{<Italian>}
\end{align*}
\]

We want to respect Kayne's idea that what rules out (21a) is BT (B), because there is a parallelism between (21a, b) and (38a, b). Just as PRO locally 'governed' by a lexical \( C^0 \) is ill-formed (=\((21a)\)), so is an overt pronoun locally 'governed' by a lexical \( C^0 \) (=\((38a)\)), and just as PRO locally 'governed' by the infinitive V adjoined to I (or I') is well-formed (=\((21b)\)), so is an overt pronoun locally 'governed' by AGR in IP (=\((38b)\)):

\[
(38) \quad \begin{align*}
\text{a.} & \quad *\text{I hope for me to help myself.} \\
\text{b.} & \quad \text{John thinks that he is a genius.}
\end{align*}
\]

Hence, all we have to explain is why it appears that the GC for (PRO qua) pronoun must be defined by recourse to the notion of government:

\(^{11}\) Kayne (2000: 66) argues that, since the Sardinian infinitive is adjoined to T rather than T', the language does not permit control with if. If we reanalyze all instances of adjunction to T' as adjunction to T, the infinitive-clitic order in Italian and the clitic-infinitive order in Sardinian are to be analyzed as multiple adjunction of the clitic and the infinitive to T (differing only in the order of adjunction). The unavailability of control with if in Sardinian will be left unexplained.
(39) the local governor of PRO/ pronoun
   the finite I(+AGR)
   the infinitive V adjoined to I
   the lexical C

   the GC for (PRO qua) pronoun
   the IP in which PRO/pronoun is subject
   the IP in which PRO/pronoun is subject
   the next higher IP

Now, in order to explain the paradigm in (39) without the notion of 'governor,' let us propose to replace (34b) with (40):

(40) A pronoun must be free in the smallest domain which contains it, its Case-licenser, and a subject (the domain=IP or DP).

Note that we incorporate the notion of 'Case-licenser' to the formulation of BT (B) but not BT (A) or BT (C). I propose to relate the peculiarity of BT (B) to that of pronouns compared with anaphors or r-expressions.

There are two peculiarities of pronouns, both being concerned with Case. First, pronouns cannot be merged in Case-less positions (without forming a chain with a Case-position):

(41) a. I (myself) saw it (myself).
    b. John (*him) thinks that Mary (*him) saw it (*him).
    c. I saw him, John.

(41a) shows that anaphors can occur in positions distinct from either subject or object position. These positions are obviously Case-less positions. (41c) shows that the r-expression occurs in a right-dislocated position, which is also a Case-less position. On the other hand, (41b) shows that pronouns cannot occur in a Case-less position.

Second, the morphological form of pronouns changes depending on what Case they receive. Thus, the third-person singular masculine pronouns, he, him, and his, are assigned Nominative, Accusative, and Genitive, respectively. This contrasts with the fact that there is only one form of anaphor with a third-person singular masculine (default) specification and that there is no way to discern the Case assigned to a reflexive from its morphological form:

(42) a. A picture of himself pleased John. (himself with Genitive Case)
    b. John likes himself. (himself with Accusative Case)
    c. It is impossible for himself to fix the car. (himself with Oblique Case)
Moreover, the pronominal component of a reflexive does not necessarily manifest the same Case that is assigned to the reflexive. Thus, *myself and *yourself can be Accusative as a whole, even if *my and *your are marked for Genitive. On these observations that only pronouns are rigidly sensitive to Case requirements, we have incorporated the notion of 'Case-licenser' to the formulation of BT (B).

(40) can account for the contrast between (38a) and (38b) in a straightforward way:

(38) a. *I hope for *me to help myself.
   b. Johni thinks that he is a genius.

Since *me in (38a) is Case-licensed by for, which is outside the embedded IP, the GC for *me is fixed as the matrix IP. Hence, it violates BT (B). On the other hand, since he in (38b) is Case-licensed by I(+AGR), the GC for he is fixed as the embedded IP. Hence, it satisfies BT (B).

Next, consider the contrast between (43a, b, c). Given the Null Case theory, PRO here is Case-checked by the infinitive I. Then, the GC for PRO qua pronoun would be fixed as the embedded IP:

(43) a. *John wonders if (PRO) to go to the movie.  <English>
   b. *Marie ne sait pas si (PRO) aller au cinéma.  <French>
   c. Gianni non sa se (PRO) andare al cinema.  <Italian>

However, we cannot a priori reject the possibility that the Case-licenser of PRO undergoes head-movement. Let us propose that the infinitive I in control with if has some feature [+F] that must be checked by either I-to-C raising or V-to-I raising. On this proposal, the contrast between (43a, b) and (43c) can be explained in the following way. When the infinitive V overtly raises to I, as in (43c), V checks [+F]. In this case, the GC for PRO qua pronoun in (43c) is fixed as the embedded IP, in which it satisfies BT (B). In contrast, when the infinitive V does not raise to I, as in (43a, b), [+F] of I must be checked by covert I-to-C raising. In this case, the Case-licenser of PRO is outside the embedded IP at LF, and hence the GC for PRO qua pronoun is fixed as the matrix IP, in which it is bound by the matrix subject. Consequently, BT (B) is violated in (43a, b).

An important question that needs to be answered is: what is [+F]? For space limitation, we will not provide a definite answer to it here but juxtapose two possible solutions. The first possibility is that [+F] is a morphological selectional feature (cf. Ogawa (2002)). Consider the fact that in Japanese, the pronominal subject of a conditional clause
must be disjoint from the subject of the next higher clause:

(44) a. *Taroo-wa, kare-ga maiku-o mot-u to,
Taroo-Top he-Nom mike-Acc hold-Nonpast Comp
hito-ga kawar-u.
man-Nom change-Nonpast
‘If he holds a mike, Taroo becomes quite another man.’
b. *Taroo-wa, kare-ga maiku-o mot-ta ra,
Taroo-Top he-Nom mike-Acc hold-Past Comp
hito-ga kawar-u.
man-Nom change-Nonpast

Kare ‘he’ in (44a, b) does not violate the KL formulation of BT (B) because it is free in the finite clause of which it is the subject. However, there are good reasons to believe that the ill-formedness of (44a, b) is due to BT (B). Assuming our formulation of BT (B) in (40), Ogawa (2002) argues that the GC for the pronominal subject in (44a, b) is fixed as the matrix clause and (44a, b) violate BT (B) because the T in the Japanese conditional clause raises covertly to C. Note that in the Japanese conditional clause, there is a morphological selectional relation between T and C: the complementizer to selects the nonpast tense and the complementizer ra selects the past tense. On the basis of this observation, Ogawa argues that T moves to C in (44a, b) in order to check the morphological selectional features of C (and T). We can extend the same analysis to the control with if in (43a, b), if we can assume that the unique realization of the nonfinite T (such as to in English) is morphologically selected by the unique realization of the conditional C (such as if in English).

As a second approach to identify the content of [+F], we may think much of the fact that all the examples against the KL formulation of BT (B) ((38a), (43a, b), and (44a, b)) involve a pronominal subject in the clause with the irrealis modality. That the controlled infinitive complement has the irrealis modality is shown by the fact that (45a) can be paraphrased as (45b):

(45) a. I would like to ask him about the painting.
b. I would like it if I could ask him about the painting.

The relevance of the irrealis modality to the matter in question is also indicated by (46), which shows that the pronominal subject of a subjunctive clause in French must be disjoint from the subject of the next
higher clause:\footnote{12} 

(46) Il\textsubscript{i} veut qu'\textsubscript{il\textsubscript{j}} vienne.  
he wants that he comes-SUBJ (Hestvik (1990: 211))

A salient property of the subjunctive clauses is that T-to-C movement (and topicalization to [Spec, C]) is obligatory in Germanic SOV languages (Ogawa (2001: sect. 2.9.3)). Taking these facts into account, we can propose that T-to-C movement in (43a, b) and (44a, b) (and, perhaps, in (46)) takes place to check off the [+irrealis] feature of the relevant T. A potential problem with this proposal is that the pronominal subject of the finite conditional clauses or subjunctive clauses in English may be coreferential with the subject of the next higher clause:

(47) a. Mary yells at Bill if she is hungry.  
b. I wonder if I could pass the exam.  
c. John proposed that he (should) take over my job.

However, it is possible to assume that the T in the controlled infinitive is inherently specified as [+irrealis] but the finite T in English is not. For these reasons, we will not make a decision about what triggers the T-to-C movement in the control with if (but see Ogawa (2002) for an argument that what triggers T-to-C movement in (44a, b) is the morphological selectional feature of C).

Although questions remain, it is important to note that two generalizations have loomed up with respect to the clauses with the irrealis modality: (i) the T with the irrealis modality has a closer connection to C than the one with the realis modality in terms of morphological selection or head-movement; and (ii) certain subclasses of the clauses with the irrealis modality are not qualified as the binding domain for their pronominal subjects. Whether and how these two generalizations can receive a unified account has to be explored in a future research.

4. Can Covert Syntax be Eliminated?

In Chapter 13, Kayne focuses on scope-taking operations which are

\footnote{12} Hestvik (1990) provides a unified solution, in terms of the LF-movement of pronominal subjects, for the ill-formedness of (46) and the anti-subject orientation effects of Norwegian pronouns. See also Kayne (2001b), who proposes a movement approach to pronoun-antecedent relations which renders BT (B) (and 'government') superfluous on independent grounds.
standardly analyzed as movement in the LF component, and proposes to reanalyze them as overt phrasal movement of operators, followed by the overt movement of the remnant predicative categories. In this section, I will overview alleged evidence for Kayne’s proposal and discuss its potential problems.

4.1. Kayne’s Proposals

(48) is ambiguous between the following two interpretations: (i) what I will force you is that you do not marry anyone (force \(\succ\) no one) and (ii) I will not force you to marry any (particular) person (no one \(\succ\) force):

(48) I will force you to marry no one.

In Kayne’s proposal, (48) is ambiguously interpreted because the derivations outlined in (49) and (50) are both available in the overt syntax (W in (49) and (50) is the abbreviation of ‘word order.’ Kayne notes (p. 239) that W is not one of the familiar functional heads such as T and its essential function is to derive the correct word order of English.):

(49) marry no one \(\rightarrow\) (merger of Neg and preposing no one to its Spec)
    no one; Neg marry \(t_i\) \(\rightarrow\) (merger of W and attraction of Neg by W)
    Neg;+W no one; \(t_j\) marry \(t_i\) \(\rightarrow\) (VP-preposing to [Spec, W])
    [marry \(t_i\);k Neg;+W no one; \(t_j\) \(t_k\) \(\rightarrow\) (merger of the matrix elements)
    I will force you to [marry \(t_i\);k Neg;+W no one; \(t_j\) \(t_k\)

(50) force you to marry no one \(\rightarrow\) (merger of Neg and preposing no one to its Spec)
    no one; Neg force you to marry \(t_i\) \(\rightarrow\) (merger of W and attraction of Neg by W)
    Neg;+W no one; \(t_j\) force you to marry \(t_i\) \(\rightarrow\) (VP-preposing to [Spec, W])
    [force you to marry \(t_i\);k Neg;+W no one; \(t_j\) \(t_k\) \(\rightarrow\) (merger of the matrix elements)
    I will [force you to marry \(t_i\);k Neg;+W no one; \(t_j\) \(t_k\)

In (49), force takes scope over no one since [Spec, Neg] here is c-commanded by force. In (50), on the other hand, no one takes scope over force since [Spec, Neg] here c-commands force. An essential point of this proposal is that the scope-taking requirement on operators is sat-
satisfied by a combination of overt movements, without the LF operation of quantifier raising (QR). Kayne proposes to extend the same analysis to sentences involving universal quantifiers such as *everyone* and focus operators such as *even John* and *only John* (in these cases, *even* or *only* attracts the NP to its Spec and they move to W, after which VP is raised to [Spec, W]). This proposal enables us to eliminate the so-called 'covert syntactic component.'

This new approach to scope phenomena makes an important prediction that when an operator is overtly contained in an embedded clause, it cannot take the matrix scope since there is no LF movement of QR. This prediction is borne out by the Swedish example in (51) (Kayne (2000: 254)):

(51) Hon har bett oss att inga böcker läsa.
    she has asked us to no books read
    ‘She asked us to read no books.’

The basic word order in Swedish is SVO, but the negative phrase object in (51) precedes V by overtly raising to [Spec, Neg] without being followed by the remnant VP movement. Since it is visibly contained in the embedded clause, it cannot take scope over the matrix verb *bett* ‘asked.’

A similar contrast holds between universal quantifiers in English and those in German:

(52) a. (since) someone has tried to cheat everyone
    \( \exists \forall / \forall \exists \)

b. weil jemand versucht hat jeden reinzulegen
    since someone tried has everyone to-cheat
    \( \exists \forall / * \forall \exists \)

The English sentence in (52a) shows scope ambiguity between *someone* and *everyone*, whereas in the corresponding German sentence in (52b), the existential quantifier *jemand* must take scope over the universal quantifier *jeden*. This is because *jeden* remains within the extraposed infinitive clause and there is no LF movement of QR. As for (52a), Kayne proposes to derive the wide scope reading of *everyone* by the following derivation, postulating the functional category ‘Dist(ributive)’ in the matrix clause (cf. Beghelli and Stowell (1997)):

(53) someone has tried to cheat everyone \( \rightarrow \) (merger of Dist & movement of *everyone*)
    everyone; Dist someone has tried to cheat \( t_i \) \( \rightarrow \) (merger of W and movement of Dist to W)
Dist\textsubscript{j}+W everyone\textsubscript{i} t\textsubscript{j} someone has tried to cheat t\textsubscript{i} \rightarrow (IP-movement to [Spec, W])

[IP someone has tried to cheat t\textsubscript{i}]\textsubscript{k} Dist\textsubscript{j}+W everyone\textsubscript{i} t\textsubscript{j} t\textsubscript{k}

Another argument for Kayne’s proposal comes from the fact that quantifiers in the embedded subject position cannot take scope in the matrix clause. Thus, (54a) contrasts with (54b):

(54) a. She has requested that they read not a single linguistic book.

b. She has requested that not a single student read our book.

Not a single linguistic book can take scope over requested in (54a), whereas not a single student cannot in (54b). If the negative phrase undergoes LF movement, it remains unclear why such a subject-object asymmetry occurs. On the other hand, in Kayne’s proposals, the wide scope reading of not a single linguistic book in (54a) can be derived in the same way as that of no one in (50), whereas there is no way to derive the surface word order of (54b) by movement of a single student to the matrix [Spec, Neg] and a subsequent movement of the remnant VP/IP to [Spec, W]. Hence, (54b) is unambiguous.

A third argument comes from the following contrast:

(55) a. I will force you to turn down no one. (force > no one, no one > force)

b. I will force you to turn no one down. (force > no one, ?no one > force)

In (55a), it is fairly natural for no one to obtain the wide scope reading, whereas it is less natural in (55b), where the negative phrase is followed by the particle down. What is important is the fact that (55b) is not so degraded with the wide scope reading of the negative phrase as (54b). Kayne proposes to derive the word order in (55b) by first preposing the particle, as in (56):

(56) forced us to turn no one down \rightarrow (particle preposing)

down\textsubscript{m} forced us to turn no one t\textsubscript{m} \rightarrow (neg phrase preposing to [Spec, Neg])

no one\textsubscript{i} Neg down\textsubscript{m} force us to turn t\textsubscript{i} t\textsubscript{m} \rightarrow (merger of W and movement of Neg to W)

Neg\textsubscript{j}+W no one\textsubscript{i} t\textsubscript{j} down\textsubscript{m} force us to turn t\textsubscript{i} t\textsubscript{m} \rightarrow (VP-preposing to [Spec, W])

[force us to turn t\textsubscript{i} t\textsubscript{m}]\textsubscript{k} Neg\textsubscript{j}+W no one\textsubscript{i} t\textsubscript{j} down\textsubscript{m} t\textsubscript{k} \rightarrow

I will [force us to turn t\textsubscript{i} t\textsubscript{m}]\textsubscript{k} Neg\textsubscript{j}+W no one\textsubscript{i} t\textsubscript{j} down\textsubscript{m} t\textsubscript{k}
He then attributes the deviance of the wide scope reading of *no one* in (55b) to that of particle preposing seen in the first step of (56). If the negative phrase were assigned its wide scope by LF-movement, there would be no reason to account for the unnaturalness of the wide scope reading of *no one* in (55b).

In sum, every piece of evidence for Kayne’s theory comes from the fact that a quantifier at the right edge of a sentence can take scope over another quantifier to its left more freely than the one which is not at the right edge.\(^{13}\)

### 4.2. Typological Variations

Kayne’s approach to quantifier scope should be evaluated as an excellent exploration in that it can explain data that could never be explained in the traditional approaches to quantifier scope. In the traditional approaches, quantifiers were assumed to have their scope fixed by their movement in the covert (LF) component. Since covert component is a component after Spell-Out, operations in the covert component have been assumed to be insensitive to phonological information such as word order. Consequently, no distinction can be made in LF between head-initial and head-final languages or between the V-DP-Particle order and the V-Particle-DP order. In such a theory, there is no way to account for a difference in scopal interpretation between (52a) and (52b) or between (55a) and (55b). On the other hand, Kayne’s approach to quantifier scope has opened up a way to explain the relation between word order and scopal interpretations. Not only can it make a natural solution to the contrasts just noted, but it can also provide a solution to some of the unresolved problems with the traditional QR-based approaches to quantifier scope.

The first issue concerns the fact that, when there are two quantifiers in a simple sentence, scope ambiguity results in the German example in (57), but not in the Japanese example in (58):

\[
\text{(57) weil irgendjemand auf jeden gespannt ist.} \\
\text{since someone for everyone anxious is} \\
\text{(Kayne (2000: 257))}
\]

\(^{13}\) Even if a quantifier is at the right edge, its wide scope reading is not always possible, due to the (general) clause-boundedness of QR.
(58) dareka-ga daremo-o sinpaisi-teiru
someone-Nom everyone-Acc anxious-is \(\exists \forall /\forall \exists\)

(57) is nondistinct from (58) in that the quantificational object is not at the right edge. However, it can take scope over the quantificational subject in (57) but not in (58). In a traditional approach, this asymmetry must be due to the stipulation that the trace of a moved subject is visible for the interpretation of quantifier scope in German but not in Japanese (cf. Aoun and Li (1989)) or that the quantificational object in German but not in Japanese can undergo LF-movement to an IP-adjoined position. These are stipulations in that they are mere descriptions of facts and have no falsifiability. On the other hand, in the spirit of Kayne’s proposals, the contrast in question can be ascribed to the difference between the two languages in the way the surface SOV order is derived (Kayne (1994)).

Kayne accounts for the fact that jeden in (57) can take scope over irgendjemand by assuming that the landing site of the movement of the universal quantifier ([Spec, Dist] in his system) lies between IP and VP and that traces are visible for scope interpretations. (59) becomes ambiguous because the existential quantifier subject in [Spec, I] c-commands the universal quantifier object in [Spec, Dist] and the latter c-commands the trace of the former in [Spec, V]:

\[
(59) \quad [\text{IP irgendjemand}_{i} [\text{DistP auf jeden}_{i} [\text{VP tj gespannt ti}] \text{Dist}] \text{ist}]
\]

If the surface order in (58) were derived in the same way, it would be predicted that a corresponding sentence in Japanese should be ambiguous. However, it is possible to argue that the SOV order in Japanese is derived in a different way. Specifically, let us propose that in Japanese, the verb is first raised and adjoined to Dist and the remnant VP is moved to [Spec, Dist], as in (60) (cf. Kayne (2000: 277, note 110)):

\[
(60) \quad [\text{IP [DistP [VP dareka-ga}_{j} [daremo-o_{i} tV]]i V(sinpaisi-teiru)+Dist t_i]\
\]

Here, the asymmetric c-command relation between a subject and an object is unchanged. Hence, daremo-o cannot take scope over darekaga.

A number of tests are conceivable that can justify or falsify the structural distinction between German and Japanese. First, we predict that the sequence of a subject DP and an object DP excluding the verb should behave as a constituent (VP) in Japanese but not in German (see Koizumi (2000) for arguments that this is the case for Japanese).
Second, we predict that the sequence of an object DP and the verb excluding the subject DP should behave as a constituent (DistP) in German but not in Japanese. Third, (58) and (60) show that, while the Spec of IP is empty in Japanese, it is filled by the subject quantifier in German. Hence, if an adverb is adjoined to DistP, we predict that it should be able to precede the subject quantifier in Japanese, whereas it cannot in German. If any one of these predictions is not borne out, we will be forced to propose a more complicated derivation of (59) to account for its unambiguity.

4.3. Inverse Linking: Alleged Evidence for LF

His theory of quantifier scope not only makes such an interesting prediction about the relation between word order and scope interpretation. It also prerequisites elimination of some currently assumed syntactic constraints from grammar or relaxation of them. In other words, it forces us to reorganize the theory of grammar in a nontrivial fashion.

Consideration of the so-called 'inverse linking' phenomena makes this point:

(61) Someone from every city despises it. (∀ > ∃/* ∃ > ∀)  
    (May (1985: 61))

In (61), every city in the subject NP must take scope over someone and can be the antecedent of the bound pronoun it, despite the fact that every city does not c-command someone or it in overt syntax. The inverse linking is taken as one of the strongest pieces of evidence for the existence of the LF component and the covert movement operation of QR, as a result of which every city moves to a position that does c-command someone and it:

(62) [NP [every city]i [NP someone from ti]] despises it.

Kayne's theory has a way to account for the inverse linking phenomena. Thus, (61) can be derived in the following way:

(63) someone from every city despises it → (VP-movement prior to QR)  
    [despises it]i [someone from every city ti] → (merger of Dist and movement of every city to [Spec, Dist])  
    every cityj Dist [despises it]i [someone from ti] tj → (merger of W and movement of Dist to W)  
    Distk+W every cityj tk [despises it]i [someone from ti] tj → (movement of DP to [Spec, W])
Throughout the derivation in (63), someone does not c-command every city, whereas every city c-commands someone between the third and fourth steps. Hence, if quantifier scope can be determined derivationally, it follows that (61) permits only the wide scope reading of every city.

However, in order to legitimize the derivation in (63), some of the currently assumed syntactic constraints must be eliminated from UG. Whether Kayne’s approach can be justified or not depends on whether the elimination is justified on empirical grounds. First, the final representation of (61), shown in the last line of (63), involves a trace of every city not c-commanded by its antecedent. This would violate Fiengo’s (1977) Proper Binding Condition (PBC) (cf. May (1977)). However, if Müller (1997) is correct in claiming that the PBC effects such as *Which picture of t do you wonder who are now on sale? are a violation of the Minimal Link Condition (MLC) and that the PBC is not a primitive constraint of UG, both the derivation and the final representation in (63) are legitimized, since attraction of every city by Dist and attraction of [someone from t] by W are triggered by different formal features and no violation of the MLC is involved here. In this sense, Kayne can deal with the inverse linking phenomena only if Müller’s study on the so-called PBC effects is on the right track.

Second, extraction of every city from the subject noun phrase in the third step in (63) should be ruled out by the Condition on Extraction Domain (CED). However, if the ill-formedness of examples such as *Who did a picture of t please Mary? could be attributed to a version of Kuno’s (1973) Constraint on Internal Clauses (CIC), which filters out a phonological gap at the right edge of a sentence-internal constituent, we can legitimize the final representation in (63), since no CIC violation is involved here. In this sense, Kayne can deal with the inverse linking phenomena only if the CED, a derivational constraint, can be eliminated in favor of the CIC, a phonological filter.

Third, although we have assumed that W is merged between VP and IP and only VP can be moved to [Spec, W], the legitimacy of the derivation in (63) depends on the possibility of merging W above IP and moving DP to [Spec, W]. However, if W could be merged anywhere and any category could be moved to [Spec, W], it should lead to the problem of overgeneration. In order to avoid this problem, we must define how to use W and its Spec (cf. sect. 4.5).
To sum up, Kayne’s theory of quantifier scope can accommodate the inverse linking phenomena only if certain constraints are eliminated or relaxed. Whether the elimination/relaxation is possible or not must be examined on empirical grounds in future research.

4.4. ACD in Wh-in-situ

A potential problem with Kayne’s approach to quantifier scope is concerned with a certain asymmetry between universal quantifiers and certain instances of wh-in-situ with respect to the possibility of licensing the so-called Antecedent Contained Deletion (ACD) in it.

A general problem with the ACD is that it can be properly interpreted without inducing an endless regression of the copying operation. Thus, (64) is well-formed with the interpretation that John will stare down everybody that Fred will stare down. This fact has been taken to be evidence for the existence of some operation that moves the quantificational NP (QNP) out of VP:

(64) John will stare down [everybody that Fred will [e]].

In May (1985), the relevant movement takes place at LF. In Kayne (2000), however, this movement of QNP is an overt syntactic operation followed by the movement of the remnant VP. Kayne provides the marginality of (65) as evidence for his analysis:

(65) ??John will stare [everybody that Fred will [e]] down.

Since the particle that follows the QNP must be preposed before the QNP undergoes movement, (65) will be ruled out for the same reasons as (55b) with the wide scope reading of no one.

Note also the following contrast, which Hornstein (1994: 459) observes:

(66) a. At least one person considered every senator to be smart. \((\exists > \forall / \forall > \exists)\)

b. At least one person considered every senator smart.

\((\exists > \forall / * \forall > \exists)\)

Kayne’s interpretation of the lack of the wide scope reading of every senator in (66b) is that raising of the AP smart before the QNP-movement causes stronger deviance than the movement of the particle in (65). In contrast, the well-formedness of (66a) with the wide scope reading of every senator suggests that preposing the to+infinitive phrase is more readily available.

With these in mind, consider the following contrast, taken from Pesetsky (2000: 30–31):
(67) a. I need to know which girl ordered [which boy that Mary (also) did [e]] to congratulate Sarah.
b. *I need to know which girl Sue ordered [which boy that Mary (also) did [e]] to congratulate ti.

The contrast between (67a) and (67b) shows that the wh-in-situ in multiple questions can license ACD only when it does not c-command the trace of another wh-phrase. We cannot argue that (67b) is ruled out by the Superiority Condition on wh-movement, because the D-linked wh-phrases do not show the superiority effect elsewhere:

(68) a. Which person bought which book?
b. Which book did which person buy ti?

Pesetsky (2000) proposes that the wh-in-situ c-commanded by the trace of a moved wh-phrase, such as which book in (68a), may undergo LF-movement out of VP, whereas the wh-in-situ that c-commands the trace of a moved wh-phrase, such as which person in (68b), does not undergo LF-movement. Rather, the option of moving only the wh-feature out of the wh-in-situ in overt syntax must be chosen in (67b) and (68b).14

Since the wh-in-situ in (67b) does not move out of VP at LF, the gap contained in it is not properly interpreted because of the regress problem. Thus, (67b) is ill-formed. The ill-formedness of (67b) contrasts with the well-formedness of (69), where the which-phrase is replaced by the universal quantifier phrase:

(69) I need to know which girl Sue ordered [every boy that Mary (also) did [e]] to congratulate ti.

(69) is ruled in since, in contrast to the wh-in-situ in (67b), the universal quantifier in (69) can undergo phrasal movement in LF.

Pesetsky's explanation of the paradigm is incompatible with Kayne's theory without covert syntax. (67a) must be derived in Kayne's theory as follows: first, the to+infinitive phrase is raised; second, [which boy that Mary (also) did [e]] is raised; and finally, the remnant VP is raised; all the three movement operations take place in overt syntax.15

14 Pesetsky (2000) argues that the Minimal Link Condition is satisfied in (67b) because C attracts the wh-feature of which person before it attracts the phrase which book to [Spec, C].

15 Kayne suggests (p. 256), following Watanabe (1992), that empty operator movement is involved with a wh-in-situ in multiple wh-questions. However, this suggestion cannot account for the well-formedness of (67a).
A similar derivation can be given to (67b) and (69). The question is why (67a) and (69) contrast with (67b). Since the Superiority Condition does not apply to which-phrases (= (68)), what makes (67b) deviant cannot be the overt movement of which girl. Since the overt movements of which boy that Mary (also) did [c] and the to+infinitive clause are both possible in (67a), they cannot be the cause of the ill-formedness of (67b). Since which boy that Mary (also) did [c] is overtly moved out of VP, the ACD should be properly interpreted. Thus, Kayne's theory has no way to exclude (67b) without excluding (67a) and (69).

4.5. Word Order in the Core Computation of Human Language

Another issue that needs to be discussed is the theoretical status of the functional category W. As Kayne notes (p. 239), its essential function is to derive the correct word order of English.

In Chomsky's (1995) conceptions of the computation of human language (C_{HL}), the postulation of an abstract functional head such as W raises several questions. The first question is this: Does W have interpretable features, providing 'instructions' at either or both interface levels? If it does, it is classified into the same group as T, C, and D, and if it does not, it should have the same fate as Agr. Chomsky has eliminated Agr, since it is nothing more than an indication of a position that must be occupied at once by overt operations, and since empirical effects of postulating it can be covered without it by allowing the light verb v and T to have a parametrized property of licensing multiple specifiers. Basically the same argument can be given to W. They are both postulated just in order to derive a specific word order.

One might argue that W does provide instructions at the PF interface, since its introduction triggers a different word order. However, if W could be chosen just in order to change word order, we would face the problem of overgeneration. Thus, if the structure of (48) is (70) and if we could merge another instance of W with the matrix IP, then we would have (71), which corresponds to the ill-formed sentence in (72):

(48) I will force you to marry no one.

(70) [IP I will force you to [WP [marry ti]k Negj+W no onei tj tk]]

(71) [WP [marry ti]k W [IP I will force you to t'k Negj+W no onei tj tk]]

(72) *Marry I will force you to no one.
In order to avoid the problem of overgeneration, we need to assume that providing instructions at the PF interface should not be a sufficient condition for the introduction of a new functional category such as W. This brings us back to the question of what instructions the introduction of W provides at either or both interface levels.

Another question with W is: Does word order play a role in CHL? If not, it may be unreasonable to postulate a SYNTACTIC category relevant only to word order. Chomsky (1995: 324) suggests that, since there is no clear evidence that order plays a role at LF or in C\textsubscript{HL}, it seems natural to assume that there is no linear order in the core computational system and that ordering is part of the phonological component after Spell-Out.\(^{16}\) Since Chomsky’s suggestion is based on the conceptual necessity, a departure from it, such as the postulation of W, needs to be motivated by sufficient empirical arguments; although the fact that it can dispense with the LF component may be theoretically important, it is hard to evaluate its merits without fully examining its empirical results.\(^{17}\)

Still another question that emerges with the postulation of W is related to the fact that his explanation of the generalization in (73) is not solely dependent on his postulation of W:

(73) If a quantifier is at the right edge of a sentence, it can take scope over another quantifier to its left more freely than when it is not at the right edge.

There are some exceptions to (73), as in (74a, b). For those who make these judgments (cf. p. 259), he proposes that, prior to the raising of the quantifiers, movement of the particle or the to+infinitive is either marginally or fully applicable:

(74) a. I will force you to turn no one down. (\(\text{no one} > \text{force}\)) (\(\text{=(55b)}\))

b. At least one person considered every senator to be smart. (\(\forall > \exists\)) (\(\text{=(66a)}\))

\(^{16}\) Arguments against this view have been made by Fukui (1993) and Saito and Fukui (1998). But see Kayne (2001b: note 37) for a potential problem with their view.

\(^{17}\) One theoretical advantage of eliminating the LF component is that we can reduce the number of cycles to one, which has the advantage of eliminating some of the problems induced by the possibility of countercyclic derivations such as covert raising or reconstruction as LF lowering.
He assumes that the movement is triggered by the functional category Pred\(^0\) (Kayne (2000: 229)). But it is left unclear why Pred can be merged in English only when the attractor of a quantifier phrase (call it ‘Q’ for convenience) and W are merged; otherwise, sentences like *John had down turn Bill/*John has to be smart considered Bill. would be ruled in. Therefore, Kayne must make an additional assumption that if Pred is selected into a numeration, Q and W must also be selected into the same numeration. It is a nontrivial question whether this assumption is necessary just in order to derive the correct word order of English or motivated on independent grounds.

5. Conclusion

In this review article, I argued that, although the cross-linguistic generalizations Kayne (2000) has proposed on microparametric syntax must be explained in one way or another, some of his particular explanations are based on mechanisms which are conceptually unmotivated under the minimalist perspective. I then suggested ways to incorporate some of Kayne’s observations and insights to the framework of the minimalist program.

REFERENCES


Pollock, Jean-Yves (1989) "Verb Movement, Universal Grammar, and the


Kitami Institute of Technology
165 Koen-cho, Kitami-shi
Hokkaido, 090–8507
e-mail: ogawayo@mail.kitami-it.ac.jp