Abstract:

While we agree that the base-generated phonologically null nominal known as PRO has Case, we dispute the recent contention of Chomsky & Lasnik (1995) and Martin (2001) that there is a special Null Case, assigned in English by certain instances of infinitival to. We show that, contrary to Martin’s claims, there is no consistent distinction between instances of to that license PRO, and instances of to that require other types of subjects (DP-trace and lexical subjects), either in temporal interpretation or in the licensing of VP-ellipsis. Furthermore, a closer examination of for-infinitives reveals an inability of the null Case theory to predict when a subject will be PRO, and when it will be lexical.
Introduction

A central preoccupation of modern formal syntactic theory has been the analysis of missing subjects of infinitives, as in the underlined sequence in (1):

(1) To visit Sally would bother me.

In this sentence, visit has an understood subject, most likely the same referent as the matrix dative, so that it is understood that it is my visiting Sally that would bother me.

There have been two main analyses to the missing-subject phenomenon:

(A) the plug-in approach, named as such by Baltin (1995) (see references cited there), in which the missing subject is never present syntactically, but is supplied somehow in the semantics, perhaps by rules of inference;

and

(B) the postulation of a subject in syntactic representation (Chomsky (1973), and many others).

We assume the second view. One argument for the existence of a structural subject of a superficially subjectless infinitive, at least at some derivational level, was proposed in Baltin (1995), based on the behavior of floating quantifiers, and was based on the paradigm in (2):

(2) a. *All to leave would be difficult for them.

    b. To all leave would be difficult for them.

    c. For these people all to leave would be difficult.

    d. For these people to all leave would be difficult.
The argument was based on the proposal that floating quantifiers needed immediately c-commanding subjects. In (2)(c), the lexical subject of the infinitive fulfills this requirement. If we assume that subjects are generated VP-externally, the subject will always be able to appear after the infinitive marker to, and if subjects move out of the VP for Case reasons, they can only move if their movement will satisfy this requirement (Chomsky (1995a), Chomsky & Lasnik (1995)). If to, by itself, does not have a Case feature to check a Case feature on the unexpressed subject, known as PRO, PRO will not move, and will remain in its VP-internal position; consequently, the floating quantifier will not have an immediately c-commanding subject in cases such as (2)(a), and the data in (2) is accounted for (see Baltin (1995) for details).

This pattern of data, and its account, relies upon the existence of a structural subject within the infinitive which can occupy one position but not another. If we say that the infinitives never have structural subjects, we have no account of why floating quantifiers can precede to in for-infinitives, but not in infinitives lacking lexical subjects, while floating quantifiers can immediately follow to in all infinitives. At the very least, no such account has ever appeared within frameworks that assume the plug-in view of unexpressed subjects.

This argument depends upon nothing more than the lack of an EPP feature on to to check Case (what Pesetsky & Torrego (2001) call a “sub-feature of a feature”—in this instance a Case feature on to.) The inability of to to host PRO in its Spec could be due to any of the factors enumerated in (3), which we believe exhausts the set of possibilities:
(3)(a) the lack of a Case feature on PRO;

(b) a Case feature on to, but ability of this feature to be satisfied by Agree (Chomsky (2000, 2001)), rather than Move, allowing the Case feature to be satisfied by a nominal’s remaining in situ;

(c) the lack of a Case feature on to.

A recent proposal concerning PRO that was originally made by Chomsky & Lasnik (1995) has been developed by Martin (2001) which denies any of the states of affairs enumerated in (3), and which posits an EPP feature on to which checks a Case that is associated with PRO. This case is not one of the standard Cases (nominative, accusative, etc.) but rather is a Case that these authors dub “null Case”, assigned to elements that lack phonological features.

In this article, we will examine Martin’s arguments for the existence of a null Case that is associated with PRO¹, and his implementation of this view. We will show that none of his arguments, when examined more closely and with a wider range of data considered, support a special null Case for PRO, and his implementation of an analysis which posits this null Case is problematic. In a certain sense, given the conditions on the occurrence of floating quantifiers in infinitives that were analyzed in Baltin (1995), it would be extremely unfortunate to have another set of grammatical phenomena that required PRO to be in the Spec of to. This article shows that, independent of the

---

¹ For present purposes, we need not enter the debate about whether Case is assigned or checked, even though we believe that the ultimate correct account of Case is a checking view (see Baltin & Barrett (forthcoming) for a fuller account.)
considerations in Baltin (1995), there does not seem to be a conditioning factor for PRO’s occurrence in [Spec, to].

This article is quite restricted in scope. We show that PRO does not have a null Case. We do not show that PRO does not have Case. We believe, in fact, that PRO does have Case, and that this Case, following, e.g. Sigurdsson (1991), is nominative. We do have some ideas as to how to implement this analysis, but to do so here would take us too far afield, and hence we will defer our positive proposals for the distribution of PRO to Baltin & Barrett (forthcoming).

This reply is organized as follows. Section 1 offers a brief history of Case and its role in determining PRO’s distribution. Section 2 discusses Martin’s account of null Case and his distinction between which instances of infinitival to license null Case and which ones don’t. Section 3 discusses problems with Martin’s predicted correlations between null Case-licensing to and meaning. Section 4 examines Martin’s predictions about the possibility of VP-ellipsis in infinitives, and show that those predictions are not systematically realized. Section 5 explores Martin’s account of the occurrence of lexical subjects in for-infinitives, and shows that his account of why lexical subjects occur in for-infinitives is based on untenable and empirically unsustainable assumptions. Section 6 discusses the occurrence of PRO in environments other than infinitival subject position. Section 7 sketches a preliminary version of our own view of missing subjects. Section 8 concludes.

1. Case and the Distribution of PRO

In this section, we will discuss the original rationale for Case in what can broadly be called Chomskyan formal syntax (the period beginning with the Revised
Extended Standard Theory)(Chomsky (1980)), and extending through Government-
Binding Theory (Chomsky (1981) and the current Minimalist Program(Chomsky
(1995b)). If we agree that superficially subjectless infinitives have a PRO subject at all
levels of derivation (originally argued, to the best of our knowledge, in Chomsky (1973)),
we must answer the question of why PRO appears where it does-i.e., why it appears as
the subject of infinitives or gerunds, but not as the subject of a finite clause, or as the
object of a verb or preposition, and why lexical nominals and PRO seem to appear in
complementary distribution. Jean-Roger Vergnaud, in an unpublished letter to
Chomsky and Lasnik following the publication of Chomsky & Lasnik (1977), proposed
that there was a requirement that all lexical nominals (NPs at the time, prior to the
reanalysis as DPs in Abney (1987) and earlier work by Brame) possess Case, which
was assumed to be assigned in the following environments:

(3) a. Nominative Case is assigned in [Spec, T”]
   b. Accusative Case is assigned to an NP c-commanded by V.
   c. Oblique Case is assigned to an NP c-commanded by P.

There was assumed to be a Case Filter that applied at LF that would rule out
structures containing a lexical NP that lacked Case:

(4) Case Filter : * [NP +lex] that lacks Case.

PRO, then, is simply an NP that, because it lacks lexical content, is permitted in
environments in which NPs are required but in which Case is not assigned. If we
assume that Case assignment is obligatory (see, for example, Fukui & Speas (1986)
for an example of such a proposal), we could not allow PRO in an environment of a
Case-assigner (i.e., object of a verb or preposition, or subject of a finite clause).
As Martin discusses, the view that PRO is Case-less is problematic. Freidin & Lasnik (1981) show that null operators require Case, for example, a requirement which is inexplicable if Case is inextricably bound to lexical content. Such considerations led Chomsky (1986) to adopt a proposal by Aoun (1979). According to Aoun’s proposal, a nominal needed Case in order to make the noun-phrase headed chain (see Chomsky (1981) for a discussion of the notion of chain) visible for theta-role assignment, and, by the theta-criterion (Chomsky (1981)) each chain must receive one and only one theta-role. However, chains can be headed by PRO, and can obviously receive theta-roles, as in (5) (being neutral for the moment about the position of PRO):

(5) \([\text{DP PRO}], \text{to be visited by werewolves} \)…

Therefore, in order to account for PRO heading chains, as Chomsky & Lasnik (1995) discuss, there must be a disjunctive statement for which chains are visible for theta-marking:

(6) (Martin’s (12)) A chain is visible for theta-marking if it contains a Case position or is headed by PRO.

Given the general conceptual dissatisfaction with disjunctive statements, together with empirical problems with the notion that PRO lacks Case, Chomsky & Lasnik (1995)) propose that PRO indeed has Case. However, allowing PRO to have Case leads back to the original problem that led Vergnaud (apparently) to introduce the notion of Case into this theoretical problem: how do we account for the different
distributions of PRO and lexical nominals? If they all have Case, and Case accounts for the distribution of nominals, they should all appear in the same environments.

To answer this question, Chomsky & Lasnik (1995) propose that PRO’s Case is not one of the standard Cases (nominative, accusative, oblique) but rather is a special Case, which they call “null Case”, assigned by the infinitive marker *to* to nominals that lack phonetic content (in English, or its counterpart in other languages).

2. Martin’s View of Null Case

Martin discusses the fact that not all infinitives can allow their subjects to have null Case. In particular, ECM infinitives and infinitives whose subjects have been raised cannot allow their subjects to have null Case. The subject of an ECM infinitive is lexical in English. If the subject of a raising infinitive received null Case, then raising, which is thought of as being movement to a Case position, would cause the resulting Chain to have two Cases, the null Case and a matrix nominative Case. This would violate the theta-criterion and the requirement that movement must be forced by some morphological requirement, such as Case (Chomsky (1991), reprinted in Chomsky (1995b)).

(7)a. I believe Sally to like pizza.

b. Sally seems to like pizza.

Therefore, it seems that some principled distinction must be made between those instances of infinitival *to* which do license null Case, and those instances which do not.
Following Stowell (1982), Martin proposes that certain instances of infinitival to, those in which the infinitive subject is controlled, are associated with an unrealized future time denoting the infinitive with respect to the matrix. For example, in (8)(Martin’s 22), the infinitive is always future with respect to the time of the matrix verb:

(8) a. Ginny remembered to bring the beer.

   b. Sara convinced Bill to go to the party.

   c. Bob wants to buy a new camera.

Martin follows Stowell’s original observation that raising infinitivals and ECM infinitivals, on the other hand, denote times that are simultaneous with the times of the matrix predicates:

(9) (Martin’s (23)):

   a. Zagallo believed Ronaldo to be the best.

   b. The doctor showed Bill to be sick.

   c. The defendant seemed to the DA to be guilty.

Again following Stowell, Martin analyzes control infinitives as being [+Tense, -Finite], while ECM and raising infinitives are taken to have the features [-Tense, -Finite]. The principles of Case association with Tense and infinitives are as follows:

(10) a. [+Tense, +Finite] checks nominative Case.

   b. [+Tense, -Finite] checks null Case.

   c. [-Tense, -Finite] checks no Case at all.
The implementation of this analysis is somewhat unclear. Martin provides no explicit phrase-markers for infinitives, and therefore does not tell us what the categorial status of to is taken to be, let alone whether tensed and untensed to are taken to be in different structural positions in the phrase-marker, or simply in the same structural position, differing solely in feature content. For example, it is generally assumed (see Pollock (1989)) that infinitives and finite clauses have the same structure\(^2\), and that there is a clausal projection headed by T. One possibility is that to itself is T when it checks null Case. Hence, the initial structure of an infinitive with a PRO subject would be (11):

\[
(11) \quad [T' [T \text{ to }] \quad [V'' [D'' \text{ PRO }] [V' [V \text{ leave}]]
\]

\[
[+\text{Case}] \quad [+\text{Case}]
\]

Presumably, finiteness would be a feature on T (although see Rizzi (1997) for arguments that finiteness heads its own projection). As far as we can tell, Martin is assuming this, given his discussion of verb raising in Icelandic control infinitives on pages 152-153.\(^3\) However, non-tensed to would also have to be assigned to a

\(^2\) although we do not subscribe to this assumption (Baltin (1992), Barrett (1997)). We will provide our reasons for rejecting it below.

\(^3\) Briefly, the facts seem to be the following: verb-raising in control infinitives takes place, given that the infinitive verb precedes sentential negation in these infinitives, but not in raising or ECM infinitives, as discussed in Thrainsson (1984) (Thrainsson 1984 #530) and Holmberg (1986) (Holmberg 1986 #520). The infinitive verb follows the sentential negation marker in the latter type of infinitive:

(i) (Martin’s (44)(a) Maria lofadi [ad lesa ekki bokina].

Maria promised COMP read not the-book.

(ii)(Martin’s (45)(b) ) * Eg taldi [ Mariu lesa ekki bokina].
grammatical category, and what would this grammatical category be? It would make no sense to place non-tensed *to* under T, and therefore give it the contradictory feature specifications [+T] (for the categorial feature to which it is assigned) and [-T]. In order to relate (in Martin’s terms) non-tensed *to* and tensed *to*, one might consider assigning tense and the absence of tense to a kind of “archi-category”, much like Laka’s(1990) [] phrase. Laka proposed that there is a category [] which ranges over negation and affirmation, so that what had been considered a NegP is actually a []P, a category which ranges over both negation and a category which entails the absence of negation.

In any event, it is clear that the infinitive marker *to* is not located in T, given that sentential negation follows T in English ((Pollock 1989)). In sententially negated infinitives, the negative precedes *to*, in contradistinction to sententially negated finite clauses:

(12) a. For Fred not to have eaten.

b. Fred has not eaten.

Categorically placing *to* in T is clearly the wrong move. The other possible instantiation

| I believed Maria read not the book. |
| (iii) (Martin’s (46)(b)* Maria virtist [ t lesa ekki bokina]. |
| Maria seemed read not the book. |

While we agree that this is an argument for a higher projection in these control infinitives, we do not feel that it shows this higher projection to be T. As our discussion immediately below this point in the text makes clear, it could be M(ood). It is also not clear that this higher projection is a null-Case assigner. We are exploring another possibility in Baltin & Barrett (in preparation).
of Stowell’s proposed distinction would make T a feature on to, which could then be a member of some other category, so that some instances of to would be [+T], while others are [-T]. Again, this seems to have some precedent, ironically within the analysis of negation, in that both Zanuttini (1997) and Landau (2002) have proposed that some instances of C have a Neg feature, while others don’t. Additionally, Neg is assumed by these authors to head its own projection in the languages that have C [+Neg].

It would still be helpful to determine which category to belongs to. Barrett (1997) proposes that to is the head of a Mood Phrase, a Mood0, and that all modals originate in that projection.

The semantic motivation for identifying mood, rather than tense, with the phenomenon of the unrealized future seems overwhelming. Martin himself suggests this when he notes that the finite paraphrase of a control infinitive does not contain will, but rather would, so that (8)(a) is paraphrasable not as (13)(a), but rather (13)(b) (Martin’s (24)(a) and (25)(a) respectively):

(13)(a) Ginny remembered that she will bring the beer.

(b) Ginny remembered that she would bring the beer.

The difference between (8)(a) and (13)(b), on the one hand, and (13)(a), on the other, is that the time in the complement in the former is future with respect to the matrix, but could be past with respect to the time of utterance of the sentence, while the time in the complement in the latter ((13)(a)) must be future with respect to the time of utterance of the sentence.
It is far from clear, however, that English has a future tense. Most grammarians
distinguish tense and time, and *will* is unquestionably a modal. Martin apparently
concedes that Tense may be the inappropriate notion of the relevant semantic properties
when he states (p. 147) that “I will argue that tense in control infinitivals is invariably a
modal element, but that it corresponds most closely to *would* or, in certain contexts,
*should* (see Martin 1996).” On page 149-150, he states:

“The discussion here suggests that a uniform analysis of the interpretation of
nonfinite tense may indeed be possible—at least as uniform as saying that it can be one of
a restricted number of modals, where the choice of modal seems to depend on the
embedding predicate. Here it should be kept in mind that all that is crucial to my proposal
is that control infinitivals contain some sort of tenselike element: whether it is a modal or
a pure tense is irrelevant.”

Two points can be made about this proposal. First, the lack of precision for the null
Case-checker makes the proposal and its implications somewhat unclear. Second, and
perhaps related to the first point, Martin seems to be trying to create a parallelism
between the checking properties of *to* and the checking properties of Tense. Nominative
is checked by [+Tense, +Finite], while Null Case is checked by [+Tense, - Finite].
When one examines [+Tense, -Finite] more closely, it starts looking much more remote
from finite Tense. Calling a modal a tenselike element, in our view, is just wordplay—it
has no theoretical content.
When the properties of elements traditionally associated with the Tense projection are considered, Tense does not seem to be the right notion for the distinction between infinitives with PRO subjects and infinitives with other types of subjects. However, the essence of Martin’s claim is that there is some feature which both licenses null Case, and has a constant, specifiable semantic consequence.

In the next section, we will examine the semantic consequence that Martin specifies, and we will show that infinitives with PRO subjects do not uniformly exhibit the semantic property which he associates with the feature licensing null Case.

3. Problems with Martin’s correlations.

If Martin replaces [+Tense,-Finite] with [ +M, -Finite] as the Case checker for null Case, are there any consequences, beyond those that we mentioned at the end of the last section? Recall that Martin is claiming a stable correlation between the interpretation of the infinitive, a semantic property, and the possibility or requirement of a PRO subject, a syntactic property. The unrealized future for the infinitive is supposed to correlate with the PRO subject, and this correlation is achieved by the presence of a feature on to which has both syntactic (checking null Case) and semantic (unrealized future) properties.

Interestingly enough, Enc (1996), whom Martin cites, notes that the modals generally have multiple interpretations. Martin, following Enc, notes that will can have a nonfuture sense that encompasses “a range of modalities, such as “epistemic necessity”….and “dispositional necessity”…” (p.149):
(14) (Martin’s (33)): Pat will be sleeping right now.

(15) (Martin’s (34)): Sarah will play music loud to annoy her mother.

How are these two senses of *will* to be distinguished? Are they distinguished by a feature that has syntactic consequences as well? We cannot see any syntactic consequences to the distinction between epistemic necessity and dispositional necessity.

When we look at gerunds, we find that the understood time for the gerund depends on the matrix predicate that takes the gerund as a complement. For example, the gerund is necessarily future with respect to the time of avoidance in (16), present with respect to the time of continuation in (17)(a) and (b), and past with respect to the time of stopping in (18). Do these gerunds have different syntactic representations? This seems highly implausible.

(16) He avoided bringing the wine.

(17) (a) He continued bringing the wine.

(b) He kept on bringing the wine.

(18) He stopped bringing the wine.

In short, variability in interpretation seems to be characteristic of modals, tenses, as shown by Enc, and gerunds. Syntactic manifestation of this variability does not seem to be generally indicated, however.
In any event, the correlation between futurity and null Case-checking (in Martin’s terms) does not hold, as Martin himself notes. Martin observes that factive infinitives check null Case, and are simultaneous with the matrix, and notes the following example (his (31a)):

(19) Kerry was surprised to solve the problem.

To Martin’s example, we can add the following, all complements of psychological adjectives:

(20) Kerry was \{glad, sad, sorry, upset\} to see her best friend walk out the door.

Indeed, infinitive complements of degree words can generally be simultaneous with the matrix, and contain PRO subjects:

(21) John was too obtuse to be aware of his effect on others.

(22) John was thin enough to be able to fit through the turnstile.

Noting these examples, Martin states (p. 149) that “Modals are typically future-oriented, yet they differ from pure tenses in that they can have a variety of interpretations, depending on factors such as the presence of other operators, presuppositions, and the prevalence of other context variables.”

In short, Martin is retreating from the claim that there is a stable semantic property that correlates with the syntactic property of checking null Case. He and we believe that an arbitrary diacritic on to that checks null Case, with no other motivation, would be unacceptable, and would lead to the view that to does not check null Case at all. We believe this to be the case, and we will present other evidence to this effect.
4. The argument from ellipsis for [+Tense] to

Martin argues for a distinction between [+Tense] to and [-Tense] to by assuming a generalization to the effect that only agreeing functional categories allow ellipsis of their complement (Saito & Murasugi (1990), Lobeck (1991)). Lobeck was concerned with unifying N”-ellipsis, as in (23), Sluicing, as in (24), and VP-ellipsis, as in (25).

(23) I don’t like Bill’s books, but I do like Sally’s____.

(24) Sally wanted to go out with somebody, but I don’t know who____.

(25) Sally likes George Eliot, and Fred does____, too.

Putting aside VP-ellipsis for the moment, Lobeck noted that other instances of D do not license N”-ellipsis, and other instances of C do not license sluicing.

(26) John brought a book to school, and Sally borrowed the *([N” book]).

(27) John said that Sally wanted to go out with some guy, but I’m not sure who he thought that * ([IP Sally wanted to go out with t]).

Lobeck’s generalization was that the affixal D’s requires a filled Spec with which it agrees, while the D the does not, and that this distinction between the two instances of D correlates with the ability of the D to license ellipsis of its N” complement. Similarly, a +wh C agrees with its filled wh-specifier, while the complementizer that does not, and this distinction correlates with the C’s ability to license sluicing.

Martin applies this distinction between agreeing and non-agreeing functional heads to his proposed distinction between two types of infinitive markers. VP-ellipsis does not apply in ECM complements, consonant with Martin’s analysis of the infinitives and this generalization: (examples taken from Martin 2001):
(28)(Martin’s (51)) a. *I consider Pam to [VP like soccer], and I believe Rebecca [T to] IVP e]] as well.
   b. *Bill believes Sarah to be [AP honest], and he believes [Kim [T to][VP e]] as well.

However, VP-ellipsis is possible in control infinitivals (Saito & Murasugi (1990)):

(29)(Martin’s (52)) a. Kim isn’t sure she can [VP solve the problem], but she will try [PRO [T to][VP e]].
   b. Rebecca wanted Jill to [VP join the team], so Pam persuaded her [PRO [T to][VP e]].

Martin’s analysis predicts that VP-ellipsis is impossible in infinitive complements of *seem and *appear, cardinal raising verbs, as in (30) (Martin’s (90)):

(30)(Martin’s judgements) a. *John does not like math but Mary seems to [VP e].
   b. *Harry may not be as happy as he appears to [VP e].

While at least one of us agrees with Martin’s judgments about (30)(b) (and the other finds it marginal), we both find 30(a) completely acceptable. Furthermore, when the data set is expanded, more acceptable raising predicates with VP-ellipsis emerge.

(31) Bob tries not to gain weight but he tends to.

Supposed to and have to can be shown to be raising constructions, as in (32), and ellipsis is perfectly acceptable:

(32)a. There is supposed to be a discussion.
i. There has to be a negotiation.

(33)a. Although she hasn’t spoken to the President yet, she is supposed to___.
   b. Although he doesn’t want to speak to the President, he has to___.

Martin himself notes (p. 160) that infinitival complements of the adjectives *likely* and *certain* allow VP-ellipsis:

(34)a. John isn’t likely to win, but Sally is likely to___.
   b. We don’t think that John will win, but Sally is certain to___.

Martin’s account of the possibility of ellipsis in cases such as (34) points the way to his treatment of cases such as (31)-(33), and suggests how he would handle judgements of (30) that are viewed as acceptable by native speakers. He analyzes *likely* and *certain* as having dual classifications as both raising and control predicates. Given that we know that these adjectives induce raising, Martin must show that there are sentences with *likely* and *certain* that can only be analyzed with these adjectives inducing control of their infinitival complements. His evidence for this comes from a contrast originally noted by Baltin in 1987:

(35)a. How likely to win is John?
   b.*How likely to be a riot is there?

Baltin’s account ran roughly as follows. The evidence that *likely* is a raising predicate comes from its ability to host expletives in its subject position, unlike control predicates:

(36)a. There is likely to be a riot.
   b.*There is eager to be a riot.
However, when the entire infinitival complement of *likely*, including the trace of the expletive, is fronted, as in (35)(b), the trace of the expletive fails to be properly bound, violating the Proper Binding Condition on traces (Fiengo (1974, 1977). We have accounted for the unacceptability of (35)(b), but we do not have an account of the acceptability of (35)(a). Baltin analyzed this case as a case of control, and since the empty category that is the subject of the infinitive is not a trace, it is not subject to the Proper Binding Condition.

Hence, this argument for *likely* having a dual classification rests on the Proper Binding Condition. However, there is a great deal of evidence today that the Proper Binding Condition is inaccurate (Huang (1993)). For example, Huang notes sentences such as (37), in which a VP containing a trace of a fronted object occurs:

(37) [VP Criticized t, by his boss] he, never was.

Given this evidence, we do not believe that the Proper Binding Condition can be the culprit in the deviance of (35)(b). Although it is not incumbent upon us, for the purposes of this reply, to indict the true guilty party, we would like to suggest a possible suspect. If we adopt Chomsky’s (2000, 2001) analysis of expletive-associate pairs, the expletive has a defective phi-set, and a T which contains the expletive within its specifier position cannot therefore check its phi-features against those of the expletive. T must therefore check its phi-features against the associate, which it can do by entering into the relation Agree. Agree simply requires that the two elements be in a c-command relationship, and if a larger phrase that contains the associate is wh-moved, neither T nor the associate will c-command one another. Hence, the ungrammaticality of (35)(b) would be due to the failure of T, occupied by *is*, to have its phi-features checked.
For our purposes, we have shown that the contrast in (35) is not evidence for a dual classification of *likely* and *certain* as being both control and raising predicates. Martin presents a second argument for *likely* and *certain* occurring as control predicates, based on quantifier scope. Citing Barss (1986), he discusses the contrast in scope possibilities of (38) (his (80)):

(38) a. Some senator is likely to lie to every member of the committee.

   b. How likely is some senator to lie to every member of the committee.

   c. How likely to lie to every member of the committee is some senator?

   d.

Barss and Martin take (38)(a) to be ambiguous with respect to scope, as is (38)(b); (38)(c), on the other hand, only allows the subject to take wide scope. We agree with these judgements. Using May’s (1985) account of quantifier lowering, in which a raised quantified subject can lower at LF into the infinitive from whence it came, the possibility of inverse scope in (38)(a) and (b) is due to the quantifier lowering back into the infinitive. This is impossible in (38)(c) because there would be an unbound trace; (38)(c) could therefore, under this account, only be a case of control. Requiring some senator to take wide scope, in other words, is claimed to be inconsistent with a raising analysis of *likely*.

We do not agree with the assumptions upon which this argument is based; specifically, we do not agree that (38)(a) involves quantifier lowering into the infinitive to derive inverse scope of the infinitive object over the matrix subject. Johnson (2000) notes that inverse scope between infinitive objects and matrix subjects is possible when the matrix predicate could only be a control construction, as in (39) (Johnson’s (7)): 
(39) Someone wanted to visit everyone.

Therefore, the possibility of an infinitive object scoping over a matrix subject does not necessarily indicate that the subject has lowered into the infinitive. In fact, Johnson’s analysis relies on the object moving into the matrix in the overt syntax, and quantifier scope is determined by the subject and object being clause-mates at the point at which scope is determined, perhaps along the lines of Kayne (1998). The inverse scope possibilities therefore do not differentiate raising predicates and control predicates. There must be some mechanism that prevents the object from raising into the matrix overtly, but note that just as (38) (c) does not allow inverse scope, neither does (40)(b), in contrast to (40)(a):

(40)(a) Some student was criticized by every professor.

(b) Criticized by every professor, some student was.

Hence, there is no independent evidence for a control analysis of these predicates, leaving the acceptability of ellipsis in the infinitival complements of these predicates as a problem for Martin’s analysis.4

---

4 See Sauerland & Elbourne (2002) for a recent analysis of the scope facts exemplified in (38) which does not rely on a dual analysis of raising predicates such as likely.
The problems for Martin’s analysis are much worse when one more fully considers the data from infinitive complement VP-ellipsis, as can be seen by considering (30)(a), which we take to be acceptable, and (31) and (32). If these infinitives are control cases, the infinitive marker to should have the feature [+Tense]. However, they do not differ in interpretation from the raising cases—they are all simultaneous with the matrix. Therefore, if we accepted Martin’s analysis of these infinitive complements based on his interpretation of the ellipsis possibilities, we would have to give up his correlation of the feature [+Tense] with any semantic property, obviating his claim to have isolated a simple, precise characterization of the null Case-checker.

In this section, we have shown that the ellipsis possibilities in infinitives do not provide evidence for a subset of infinitive markers that are specified as [+Tense]. We have not provided an account of those infinitives which do not allow ellipsis after to—specifically, we do not have an account of the unacceptability of (30)(b). Parenthetically, we are skeptical of the viability of Lobeck’s generalization about the class of ellipsis-licensing heads being those which agree with an element in its specifier, upon which Martin’s argument rests. While it is true that the determiner the does not license N” ellipsis, demonstratives do, as in (41), and we are not aware of any account that places an agreeing element in the demonstrative’s specifier:

(41) These pictures of Sally are quite good, but I am not crazy about those__.

This seems to be an account where Lobeck’s generalization is too strong. There seems to be another case where it is too strong, as well-namely, the case where a complementizer that has had an element in its Spec through successive-cyclic wh-movement:

(42) * Sally says that John wants to go out with some girl, but I’m not sure
(43) who, she thinks t\textsubscript{i} that\textsubscript{___}. (John wants to go out with t\textsubscript{i}.).

Assuming movement as a last resort, a guiding principle of minimalism, successive-cyclic movement is thought to arise from checking some but not all of a moving element’s features (Chomsky (1995a)). Therefore, a complementizer through which an element has successive-cyclically moved should count as an agreeing functional head, and should therefore license ellipsis, in this case sluicing.

5. For-infinitives

Recall that the original motivation for the introduction of Case into generative grammar was the desire for an explanatory account of the distribution of lexical DPs versus the distribution of non-lexical DPs. It was initially thought that Case could be a requirement for lexical DPs, while non-Case environments could be reserved for non-lexical DPs, i.e. PRO. As evidence began to accumulate that PRO also needed to be specified for Case, it looked as though Case was no longer able to provide an answer to the question of where lexical DPs had to occur, and where PRO had to occur. The idea of a null Case, which Martin develops, seemed to provide the right partition: null Case could only appear on nominals that are base-generated as phonetically null, and the other cases (nominative, accusative, etc.) could only appear on nominals with lexical content.

For-infinitives, however, disturb this neat division of labor. These infinitives have the same temporal/modal/aspectual properties as infinitives with PRO subjects. For example, there is no difference in the relevant properties (which we believe are modal) in (43)(a), with a for-infinitive complement, and (43)(b), with a control complement:

(43)(a) I would prefer for Sally to leave.
(b) I would prefer to leave.

Martin analyzes *for*-infinitives as having [+Tense] infinitive markers. He therefore takes the subject of a *for*-infinitive to bear null Case. *In other words, phonetically contentful DPs can bear null Case.* On the face of it, it would appear that we are back where we started—without an account of which nominals had to be lexical, and which could be empty. Martin proposes an account of *for*-infinitives that tries to answer this objection to his treatment of null Case. He proposes that phi-features differentiate the subjects of *for*-infinitives from PRO, by taking the former to have phi-features, which he takes to be uninterpretable on both nouns and verbs, contrary to, e.g. Chomsky (1995). Martin takes *to* to lack phi-features against which to check a subject’s phi-features; therefore, a lexical subject of an infinitive must have some other checker, and Martin takes *for* to provide this service. PRO, on the other hand, is taken to lack phi-features entirely, and therefore does not have any to check. In this way, Martin accounts for the impossibility of *for* with PRO subjects, as in (44)(his (71)):

\[(44)\]  
*John wants [for PRO to win].*

The impossibility of an ordinary lexical subject in place of PRO, then, as in (45), is accounted for not by Case, but by agreement; the infinitive’s subject has nothing against which to check its phi-features:

\[(45)\]  
*[Sally to leave] would be horrible.*

The assumption that phi-features on nominals are uninterpretable, however, and must therefore be checked, is decidedly a non-standard assumption. It is generally thought that phi-features, the features of person, number, and gender, are uninterpretable on predicates, but interpretable on argument nouns. Hence, the assumption that the
subject’s phi-features must be checked runs contrary to general thinking in these circles. Moreover, the assumption that PRO lacks phi-features is extremely implausible.

Agreement in phi-features doesn’t only exist between subjects and matrix verbal elements, such as elements in T; such agreement is also generally thought to hold (in English at any rate) between antecedents and reflexives, as in myself, yourself, himself, etc. Under the assumption that PRO lacks phi-features, how do we then account for the distinction between (46)(a) and (46)(b)?

(46)  
(a) To shave myself would be difficult for me.

(b) *To shave himself would be difficult for me.

Also, how do we then account for predicate adjective agreement in infinitives, presumably relying on phi-features of the PRO subject, as in Spanish (47)?

(47)  
(a) Estar contentos seria dificil para ellos.

To be happy (3rd masc. pl.) would be difficult for them (3rd masc. pl.)

a. *Estar contentos seria dificil para mi.

To be happy (3rd masc. pl.) would be difficult for me (1st sg.)

The idea of PRO lacking phi-features, therefore seems to be a non-starter, but it is crucial to Martin’s account of the different distributions of PRO and lexical DPs.

If PRO and lexical DPs both have phi-features, and both can have null Case, PRO should be able to be the subject of a for-infinitive, and should not be able to be the subject of a non-ECM infinitive without for.

6. PRO in environments other than infinitives.

Martin suggests in a footnote that “for reasons of space” he is concentrating on PRO in
infinitives. Given the postulation of PRO in other environments, particularly gerunds and nominals, this delimitation of focus seems arbitrary. In this section, we will briefly review the evidence for PRO in these latter constructions, in which there is no evidence for a [+Tense] configuration.

Martin’s proposal assumes a correlation between null Case assignment and the presence of a TP projection.

“…there are two types of infinitives. T in infinitivals selected by control predicates is [-tense] and checks Case, whereas T in infinitivals selected by raising predicates is [-tense] and does not check Case.” (Martin 2001, p.152)

Even though, as previous sections have discussed, it is the feature [tense], not necessarily any particular projection, which is crucial for Martin’s claims to be valid, there is evidence for the presence of PRO without the presence of [tense]. In particular, both gerunds and nominalizations provide counterevidence to the claim that the presence of [tense] is correlated with the presence of PRO.

6.1 Gerunds

Martin discusses in a footnote the case of gerunds, which can have PRO subjects, but which Stowell (1982) claims are [-tense]. Martin argues that gerunds are actually
[+tense], but provides no explanation of how the feature [+tense] can be assigned by a head [N] (or even a [V], if PRO-gerunds are taken to be bare VP projections). 5

In Martin’s argument against Stowell’s proposal that gerunds are [-tense], it is claimed that in (1) below (Martin’s footnoted (ib)) the fact that the “bringing” is interpreted as prior to “remembering” entails [+tense], since simultaneity is associated with [-tense].

(48) Jenny remembered bringing the wine

However, assuming that the gerund is [+tense] makes sentences like (2) below, with the same tense reading as (1), mysterious:

(49) Jenny remembered Bill bringing the wine

Furthermore, as we recall from 16-18, repeated here, gerunds can be interpreted as simultaneous, as in 17:

(16) He avoided bringing the wine.

(17)(a) He continued bringing the wine.

(b) He kept on bringing the wine.

(18) He stopped bringing the wine.

---

5 Martin does not claim that the head of TP is a necessary condition for the assignment of null Case. However, we note in passing that allowing the feature [+tense] by itself to be the defining environment for the assignment of case to PRO is problematic, since if all that is needed for null Case assignment is [+tense], rather than [+tense][-finite], we should expect PRO to turn up in tensed clauses as well. To argue instead that gerunds are dominated by TP would be contrary to the distributional evidence: gerunds have the distribution of NPs not TPs.
A closer scrutiny of 16-18 would indicate that the gerund itself does not have any tense specification since its temporal relation to that of the matrix seems to be entirely specified by the matrix predicate.

In fact, the syntactic projection of gerunds is only part of the problem. The case-assigning properties of gerunds are in fact fairly complex.

There are three basic types of “clausal” gerunds, or gerunds where a subject is represented: POSS-gerunds, ACC-gerunds and PRO gerunds.

A hierarchy of arguments may in fact exist within gerund DPs, as discussed in Longobardi (2001), although space does not allow a detailed discussion here.

If gerunds are [+tense] as Martin claims, then all three subject types would be assigned the same null Case, just as “for-“ infinitives discussed above necessitate an analysis whereby null Case can be assigned to overt DPs in the presence of “for”. It is unclear in the above cases, however, what additional element would be responsible for assigning the null Case to the lexical DPs in POSS-ing and ACC-ing gerunds. We note further that all three types appear to have the same simultaneous reading found in the case with PRO:

(50) John remembers PRO selling books

(51) John remembers Bill selling books

(52) John remembers Bill’s selling books.

---

6 Reuland (1983) compares ACC-ing and PRO-ing alternations to want-type verbs, considering them both clausal. Reuland argues that –ing is under Infl.
6.2 Nominals

Other similar problems arise when we consider nominalizations. Chomsky (1986) proposed the existence of PRO in nominals based on evidence from binding facts. Arguing that the Sentential Subject Condition (SSC) applies in noun phrases, Chomsky cited the following (Chomsky’s example 219)

(53) (a) they (i) told [stories about each other(i)]
   (b) *they (i) heard [my stories about each other(i)]
   (c) *they (i) told [stories about them(i)]
   (d) they (i) heard [my stories about them(i)]

The subject my in (ii) blocks binding by condition A but permits binding in (iv) by condition B. The latter condition also blocks binding in (iii). Mysteriously, however, in (54) (Chomsky’s 220(iii)), binding is allowed without a condition B violation where it is expected:

(53) they (i) heard [stories about them(i)]

To account for this, Chomsky proposes the representations in (55) and (56) (Chomsky’s 221(i) and 221(ii)).

(54) *they (i) told [PRO(i) stories about them(i)]
(55) they (i) heard [PRO(j) stories about them(i)]
The SSC applies as expected assuming that the DET position (spec) can include an implicit argument with the properties of PRO.

As further evidence in support of PRO in NP, Chomsky (1986) cites examples (noted by Howard Lasnik) repeated below (Chomsky’s 223(i) and 223(ii)):

(56) [The knowledge that John might fail] bothered him

(57) [The possibility that John might fail] bothered him

In (57), John can be the antecedent of him but in (56) it cannot. Chomsky argues that this distinction is accounted for by assuming that the syntactic representation of (56) includes a PRO controlled by him.

(58) [PRO knowing that John might fail] bothered him

In both (56) and (58) there is a violation of condition C if him refers to John, since John is now bound by PRO.

Recently, Longobardi (2001) notes that PRO appears in nominalizations based on cross-linguistic evidence showing cases where the phonetically null subject argument role of the noun is understood as coreferential with the anaphor/antecedent. Evidence from Italian (Longobardi’s example 19) is shown below:

(59) La descrizione di se stessa inviata a quella ditta e stata de grande giovamento alla carriera di Anna.

The description of herself submitted to that firm was very helpful for Anna’s career. Furthermore, it is noted that the environments where this arises are exactly those in which infinitives with PRO-controlled subjects can replace the head noun in question:
(60) Descivere se stessa in quell modo e stato di grande giovamento alla carriera di Anna

Describing herself that way was very helpful for Anna’s career.

Assuming that PRO is possible in nominalizations, it is interesting to note that there are a variety of temporal interpretations in these constructions. Some have the simultaneous interpretation while others, have a future-orientation to the nominal complement:

(61) IBM(i) launched a PRO(i) takeover of Xerox(j)

In fact, the situation with such nominalizations closely parallels the gerund examples in (16)-(18) presented earlier:

(62) IBM avoided the takeover of Xerox
(63) IBM continued the takeover of Xerox
(64) IBM stopped the takeover of Xerox

Thus, given the evidence that PRO exists in both nominalizations and gerunds, it would appear that these contexts of null Case fall outside the analysis presented by Martin, and considerably weaken the claim that [tense] has any bearing on the assignment of such case.

7. A Preview Of Something Positive

So far, we have been extremely negative about the idea that PRO has a special null Case. We do, however, believe that PRO has Case, being convinced by the arguments in Chomsky & Lasnik (1995) to this effect. However, we are also convinced by the arguments of Sigurddson (1991) that PRO has one of the standard Cases-specifically,
nominative. However, once one accepts the idea of PRO having a nominative Case, one is left with the problem of explaining why lexical DPs cannot occur in the position of PRO. After all, as we stated in Section 1, Case was originally invoked to partition the environments for PRO and lexical DPs. If both lexical subjects of finite clauses and PRO, which occurs as the subject of infinitives, have nominative Case, we apparently fail to explain this partitioning.

We believe that the answer comes from taking seriously the view of Case-licensing as being Case-checking rather than Case-assignment. Case-assignment takes the nominal to lack Case features on its own, and views the process whereby nominals bear Case-features as a process of the nominals getting them from the Case-assigner—something like a process of transfer. Case-checking, on the other hand, assumes that the nominal bears Case at the outset, and its particular Case features must be licensed by some other element in its local domain that has the same Case features—V, T, or P, for instance. The checking view is simply inspection for a feature and feature value that are already present.

The checking view thus allows for the logical possibility that a nominal will enter the derivation with a Case feature, but that it might lack another element in its local domain with the same Case feature. If Case is checked at PF, a lexical nominal with a nominative Case feature might occur in an environment that lacks any other element with this nominative Case feature.

We believe that infinitives lack this nominative Case feature. Recently, Haeberli (1999), and Pesetsky & Torrego (2001) have proposed that what people call nominative is really an uninterpretable categorial T (Tense) feature on the determiner of
the nominal. Suppose that infinitives lack Tense. It would follow, then, that a nominative subject of an infinitive would then carry an uninterpretable T feature with nothing against which to check this feature.

One solution would be to delete the nominal that carries this T feature at PF. Lasnik (1995) has proposed other instances of deletion rescuing a derivation by deleting an element with an uninterpretable feature, in his analysis of pseudo-gapping. We would suggest this to be at work here. In other words, what people call PRO is really a nominal that has been deleted at PF because it has an uninterpretable T feature. In a sense, then, there really is no PRO, but rather, along the lines of Rosenbaum (1967), one of the earliest treatments of missing subjects of infinitives, the nominal has been deleted.

This is a rather sketchy treatment, and we will be presenting the details and implications in Baltin & Barrett (forthcoming).

8. Conclusion

Starting from the view that PRO has a Case, but that this Case seems to be reserved for elements that are phonetically null, Martin tried to develop Chomsky & Lasnik’s idea that PRO received a special null Case. He attempted to develop a principled account of the distribution of elements that receive this null Case, viewing an account of null Case that relied on arbitrary diacritics for the specification of null Case’s distribution as intolerable. We share Martin’s view that null Case’s distribution should be specified in a principled manner, but we do not share the view that this is achieved in Martin (2001). On closer examination, each of Martin’s arguments for a [+Tense, -Finite] specification for a Case-
checking environment for PRO dissolves, and given the mechanisms that this analysis
invokes for for-infinitives, there is no natural way to account for the distribution of PRO
and lexical DPs. Given the variety of pieces of evidence and interesting insights in this
article, we feel that the problems for this analysis are not due to this particular
implementation of it, but rather are due to the view of PRO as having a null Case.

References:


MIT.

Baltin, Mark. 1992. Negation and Clausal Structure. Talk Presented at the 23rd Regional Meeting of the


Baltin, Mark & Leslie Barrett. forthcoming. A Deletion Alternative to PRO.


Cambridge, Mass., Massachusetts Institute of Technology.


