Notes for Linguistics 610
Howard Lasnik

Case and Passives (revised October 2, 2002)

Case theory

The virtue of the X-bar theory introduced in Chomsky's (1970) "Remarks on Nominalizations" is that it captured a lot of generalizations that were not captured before. Phrasal categories appear to be projections of heads almost all the time. If we say "all the time", then we get the strong version of the X-bar theory. Further, when you introduce them into the structure they will project phrasal structures. That goes a long way towards solving many of the learnability problems inherent in phrase structure rules, so let us assume something like it is right. The problem is that descriptively it seems quite incorrect. Let us look at some of the predictions it makes and let us see to what extent they are correct and to what extent they are incorrect.

In English we have:

(1) V' ⇒ V { NP / PP / CP }

prove the theorem
rely on Bill
stated that John is here

N' ⇒ N { NP / PP / CP }

*proof the theorem
reliance on Bill
statement that John is here

We do not necessarily expect that every verb has a corresponding nominal (that was part of Chomsky's (1970) point in *Remarks*...), but "prove" clearly has a nominal: "proof". Since "prove" has a nominal and since the X-bar theory says that the complement system is general, and since "prove" by its semantic nature takes an NP complement, we would certainly expect "proof" to take an NP as well.

This gap seems very general. Let us collect some other transitive verbs that nominalize and see what happens. I am going to claim that the gap is even more general: it is not just that there

are no transitive verbs that nominalize in this way, there are no nouns at all that have NPs as their complements, in flagrant violation of the prediction of the theory laid out here thus far.

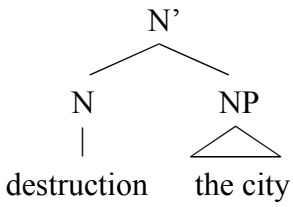
Consider some other examples:

(2) destroy the city
*destruction the city
write the book
*writer the book

To account for this paradigmatic gap, Chomsky (1970) proposed an obligatory rule that says: insert "of" in the context: N __ NP:

(3) *Of*-insertion rule:
Ø ⇒ of / N ___NP

(4)



Under this proposal NPs like "*destruction the city" and "*writer the book" exist at D-structure, but there is a transformational rule that wipes out the evidence that they exist, by changing them to "destruction of the city" and "writer of the book".

Look now at the following:

(5) A'⇒ A $\begin{Bmatrix} *\text{NP} \\ \text{PP} \\ \text{IP / CP} \end{Bmatrix}$ proud Bill
happy about that result
proud that Harry won the race

Even though we cannot say "proud Bill", we can say "proud of Bill". Following Chomsky, we

might generalize that "of" is inserted also in the context: A__ NP. In this regard, the rule of *Of*-insertion would be:

(6) *Of*-insertion rule:

$$\text{Ø} \Rightarrow \text{of} / \left\{ \begin{matrix} \text{N} \\ \text{A} \end{matrix} \right\} \_\_ \text{NP}$$

This looks more and more like a phonological rule. In fact when anyone puts forward a rule like this in phonology, they always have a disclaimer that would say, in this case, that N and A are put together for ease of exposition, not because N and A have any reality. N and A are just bundles of features, and the reason why the rule works this way is because N and A share certain features.

This is a classic kind of argument in phonology, and one would expect a similar argument in syntax to be appropriate. In fact, in the late 60s and early 70s linguists started exploring ways of representing syntactic categories as bundles of features.

We will look at a proposal that is always attributed to "Remarks on Nominalizations" and which almost certainly dates from that era, but which I have not been able to find in that work. The proposal is the following. We have four basic lexical categories with the following representations in terms of two binary features:

(7)

NOUN	[+N, -V]
VERB	[-N, +V]
ADJECTIVE	[+N, +V]
PREPOSITION	[-N, -V]

In the mid to late 70's when Chomsky was talking about this analysis, he gave a sort of justification. Verbs and adjectives ought to share a feature because they are the major types of predicates, that is, [+V]. Similarly, nouns and adjectives ought to share a feature because they are the major kind of substantive categories, that is, [+N].

Chomsky suggested that prepositions are not like either nouns or verbs in these respects and

that this feature system explains this. But of course, we know from phonology, that that is incorrect. Prepositions share with nouns [-V] and they share with verbs [-N]. That is, the minus value of a feature is just as available as the plus value of a feature. I put aside this technical question.

The "*of*-insertion" rule characterizes a natural class. Nouns and adjectives share the feature [+N], so the rule of *of*-insertion applies to [+N] categories.

Jean-Roger Vergnaud, in a personal letter to Noam Chomsky and me in early 1977, was concerned with this phenomenon and a whole class of arguably related phenomena that I will return to. Vergnaud's basic idea was that in a richly inflected language, (like Latin, Greek, etc...) there are characteristic positions in which NPs with particular 'cases' occur. These positions are:

- Subject of a finite clause --- Nominative Case
- Object of a transitive verb --- Accusative Case
- Specifier of an NP --- Genitive Case
- Complement of a preposition or of certain designated verbs -- Oblique Case

Suppose that this is true even in languages (like English) with little overt case morphology. Further, notice what we had before:

(8) *N NP
*A NP

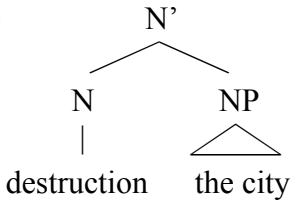
These configurations are not any of those enumerated above. Chomsky's interpretation of what Vergnaud said is the following (which he developed in the paper *"*On Binding*"* (1980) and later on in *Lectures on Government and Binding* (1981), henceforth *LGB*). A noun phrase needs case (which Chomsky began to call 'Case') to be morphologically complete. If it finds itself as a subject of a finite clause it will get Nominative Case, if it finds itself as an object of a transitive verb, it will get Accusative, if it finds itself as a specifier of an NP, it will get Genitive, and if it finds itself as the complement of a preposition or of certain designated verbs, it will get Oblique Case. Arguably, if an NP finds itself as the complement of N or A, it will get no Case at all,

violating some requirement. In particular, Chomsky's proposed the Case Filter:

(9) *The Case Filter:*
*NP with no Case

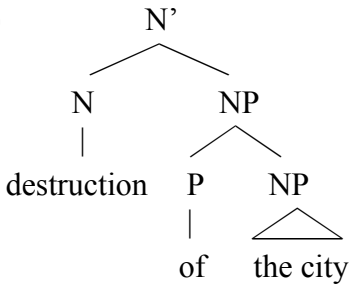
Structure ((4)), repeated here, is, therefore, excluded, even though the X-bar theory permits it.

(10)



As noted earlier, in place of "*destruction the city" we find "destruction of the city". It seems that "of" is a pleonastic preposition, one that really does not add anything to the meaning. "Destruction of the city" means exactly what "*destruction the city" would mean if you could say it. That "of" is pleonastic is then straightforward. As for its prepositional status, well, it sounds like a preposition, so let us just say it is one.

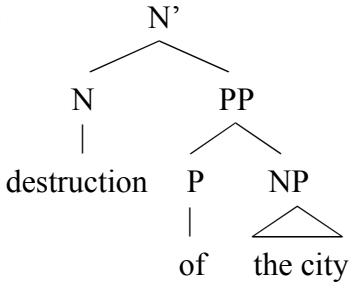
(11)



We adjoin the "of" to the NP "the city", creating another NP node. "of the city" behaves like a constituent: "Of which city did you witness the destruction?" is good. If Chomsky is right about one general principle of derived constituent structure, then this NP created by adjunction is

automatically changed into a PP.[1]

(12)



We have not come close to solving all the problems created by "*of*-insertion" but we have gotten pretty close to solving one of these problems, that of its apparent obligatoriness. "*Of*- insertion" is "obligatory" because if you do not do it, you will end up violating the Case Filter.

Let us look at a couple of other configurations relevant for Case theory:

(13) John is likely to lose the race

In (13) "the race" is the semantic complement of "lose". "Lose" also, by its semantic nature, needs a subject. In other words, "lose" has a subject theta role that it has to assign. On the face of it, sentence (13) looks like a violation of the Theta Criterion. There is a related mystery to this. Look at the following sentence, which is very similar in meaning to (13):

(14) It is likely that John will lose the race

In (14), the expletive "it" is pleonastic, not referential. We can see then that (13) also seems to have another violation of the Theta-Criterion because "John" in (13) is in a position where no theta-role is assigned, as we can see in (14), where we have an expletive.

---

[1] Recall from *Syntactic Structures* that one of the things that passive does is to insert "by" in front of the NP that used to be the subject ("The man was arrested by the police"). Chomsky in *LSLT* noted that "by the police" behaves for all later purposes as if it were a PP. He argued that this follows from a principle of derived constituent structure. Suppose you have created "by the police". It looks for all the world like something that could have been created by the PS rules as a PP, as in something like: "I stood by the lake", hence it is a PP.

So, in (13) we have the verb "lose" which needs to assign a subject theta role but there does not seem to be anything to get it, and we have an NP "John" which needs a theta role, but it is not in a position to receive on. We have two problems, but they are complementary problems. We have an "extra" argument and we have an "extra" theta-role assigner. Putting the two problems together leads to a solution. What is really going on is that "John" gets the theta role that "lose the race" has to assign. We have to ask two things: whether that gives us the right interpretation of the sentence, and if it does, how we can mechanically implement that. It seems that it does give us the right interpretation of the sentence. How can we implement that? In part, it depends on our theory of theta-role assignment. In the classic theory of theta-role assignment there is only one thing that we can do, and that is to say that in D-Structure "John" is the subject of "lose the race". By S-Structure it has become the subject of the higher clause. In *LGB*, D-Structure is defined as a pure representation of GF (Grammatical Function) theta (structural positions relevant to theta-role assignment). In a theory like that there is no alternative but to say that "John" is the subject of the lower clause.

At D-Structure, then, we have the following:

(15) $[_{IP}\ [_{I'}$ is likely [John to lose the race]]]

Yet at S-Structure "John" is higher in the clause, in particular, in the higher subject position, as evidenced by the agreement properties displayed in, e.g., "John is likely to win the race" and "John and Bill ARE likely to win the race".

Similarly, the Subject-Auxiliary inversion transformation treats this 'raised' NP as a subject:

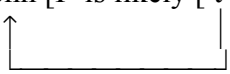
(16) a. Is John likely to win the race?
b. Are John and Bill likely to win the race?

Every test you can think of for subjecthood will be passed by "John" in "John is likely to win the race", except one, the Theta Criterion. With a movement analysis, we can have our cake and eat it too. At the level of representation relevant for satisfying the Theta Criterion, "John" is not the subject of "is likely to lose the race". But at the level of representation relevant to everything

else, “John” is the subject of “is likely to lose the race”.

To complete our implementation of this insight, we need a rule that moves “John” from subject of “lose” to subject of “is likely”:

(17) $[_{IP}$ John [I’ is likely [t to lose the race]]]



The rule that moves “John” to specifier of IP must be a rule that says something like this:

(18) Look for an NP that does not have anything in it. Find a lower NP and move it into that position.

The specifier of IP does not have anything in it (since it is not a theta-position), so it is a possible place to move. This is a rule in the same bag of rules as the verb raising rule in Lasnik (1981) (though not the verb raising rule in Pollock (1989)). Remember what that rule said: Look for a V position that does not have anything in it. We might call rule (18) *NP raising*.[2] In a simple restrictive framework for transformations, (18) would translate into: Substitute NP for NP.

Structure Preservation will demand that if you are going to substitute an NP for something, that thing has to be an NP. Similarly, you do not have to say in the rule to substitute an NP for an NP that is empty, because if you try substituting an NP for an NP that already has something in it, that would violate recoverability of deletion. All we really have to say is: Substitute NP, or even more simply, Move NP.

That kind of suggestion in the early 1970's led within a couple of years to a radical simplification of the the movement part of the transformational component:

(19) Move α

(19) means: Find any constituent you like, and move it wherever you like, but subject to such

[2] Note also that it is a classic instance of substitution.

constraints as recoverability, structure preservation, etc...

So, we have a nice general rule, but it looks like we have to call it obligatory, because of bad sentences like the unraised source of "John is likely to lose the race":

(20) *Is likely John to lose the race

The Case Filter of Vergnaud and Chomsky provides an alternative to stipulated obligatoriness: if "John" does not raise, it will not have Case. "John" in the D-structure of "John is likely to lose the race" is not the subject of a finite clause. Nor is "John" the object of a transitive verb. Further, "John" is not the specifier of an NP. Finally, "John" is not the complement of a preposition.

Hence, "John" is without Case, violating the Case Filter. That forces the raising to take place. We thus do not have to say that the raising is obligatory, just that if you do not do it, you end up violating the Case Filter. (This state of affairs is quite parallel to that of Affix Hopping. In effect, it is obligatory, but its obligatoriness did not have to be stipulated, as it followed from the Stranded Affix filter.)

Vukić: Why couldn't we salvage the ungrammaticality of "It is likely John to lose the race" by inserting "of" as in "It is likely of John to lose the race".
Lasnik: That is a very good question. So far that is unexplained. As we will see, in a couple of other situations two solutions for remedying a Case Filter violation, raising and *of*-insertion, seem to be available:

(21) *destruction the city
destruction of the city *of*-insertion
the city's destruction *t* raising

But the D-Structure of (20) only permits the raising option. I will return to this problem when I

present Chomsky's refinement of Case theory in his *Knowledge of Language* (1986). ■

Returning to general properties of raising, one might expect that alongside raising adjectives, there are raising verbs. And in fact one does find raising verbs.

(22) John seems to be clever

(23) It seems that John is clever

These two sentences are parallel with respect to thematic properties. The "it" in (23) is an expletive; it does not refer to anything. This indicates that "seem" does not have any subject theta-role to assign.

In (22) we have a familiar mystery: "seems" does not have any subject theta-role to assign and nevertheless "John" (an argument) is in that position. In addition, "to be clever" has a subject theta role to assign, but there is nothing there to receive its theta role. This problem can be solved in the now familiar way: by raising "John" from the position of subject of "to be clever" to the position of subject of "seems". As usual, if "John" does not move, it will violate the Case Filter.

At D-structure we have:

(24) $[_{IP}$ seems [John to be clever]]

And at S-structure we have:

(25) $[_{IP}$ John seems [t to be clever]

In fact, if (24) were an S-structure (see also (20) above), the sentence would additionally violate another principle which Chomsky in his *Lectures on Government and Binding* (1981) (henceforth *LGB*) called the Extended Projection Principle (EPP). EPP says the following: a clause must have a subject. We can control for that violation, however, by providing an expletive subject:

(26) *It seems [John to be clever]

Gutierrez: How about a sentence like: “John wants to solve the problem”, where “to solve” does not seem to have a subject either?
Lasnik: Yes, ”solve” clearly has a subject theta-role to assign, but we cannot give the same explanation that we gave about “likely” and “seems”. If we say that “John” moves from the subject position of “to solve” to the subject position of “want”, we would expect to be able to have an expletive as the subject of “want”, which is not possible. This is in accord with the clear semantic fact that “want” is attributed to an individual, that is, it has a subject theta-role to assign. I will return to this issue shortly.

■

To conclude our present preliminary discussion of Case theory, a small technical problem should be noted regarding the Case Filter. Recall that we have been assuming, following Chomsky’s work since the early 70's, that when something moves, it leaves behind a trace, where a trace has to be of the same category as the thing that it is a trace of.

Consider the following example:

(27) John seems [t to be crazy]

The trace in (27) is a NP. But this NP is not in a Case position, so apparently it violates the Case Filter. Given that the sentence is grammatical, we have to change the Case Filter to something like this:

(28) *The Case Filter:*

* NP [lexical] with no Case

Notice that this goes along very nicely with the idea that what is going on is something about morphology. If you have to pronounce the NP, then it has to have certain morphological

properties, like Case.

With this much new background about the Case theory we now turn to discussion of passive constructions and passive-related phenomena, investigations of which have been an important and lively issue since the earliest days of transformational grammar.

Passive

Evidence for modularity

The original Passive Transformation as in *Syntactic Structures* (T12) and *LSLT* does at least three major things:

1. Moves the "object" NP to the left
2. Inserts *be* and *-en*
3. Moves the "subject" NP to the right and inserts *by* forming a *by*-phrase

In other words, in this formulation the passive transformation combines three different operations. Let us call this a "unified" approach to passive. An alternative analysis might treat all five operations as functioning as independent operations of the syntax. This alternative analysis might be called a "modular" approach.

An overwhelming argument for the modular analysis and against a unified analysis is that we find all the modules separately. When they all happen to converge on one sentence, then that will be a classic "passive sentence". But you can find them all independently. To say that there is a phenomenon of passive would be like saying there is a phenomenon of negative questions because there are sentences that are both negative and questions, but no one ever proposed a negative question transformation because the modular analysis of *Syntactic Structures* was so appealing in that case. Recall that a negative question was just the result of independent application of the negative transformation T16 and the question transformation T18.

I will now illustrate what I have claimed: that every one of the properties of passive sentences can occur independently of every other one. For the moment, let us represent insertion of *be* and *-en,* that is, operations 2 and 3 above, together, as *be* + *-en,* and operations 4 and 5 together as "*by*-phrase". Operation 1 is encoded as "NP movement". Then we have the following situation with regard to occurrence of those operations in different contexts:

(29)

By-phrase	BE+ -EN	NP-movement	Example
✓	✓	✓	John was arrested by the police
X	✓	✓	John was arrested
✓	✓	X	It was proved that 2+2=4 by John
X	✓	X	It is believed that John is crazy
X	X	✓	John seems to be crazy
✓	X	X	The destruction of Rome by the barbarians
✓	X	✓	Rome's destruction by the barbarians
X	X	X	The police arrested John

We can infer from the chart in (29) that "passive sentences" are just a coincidental co-occurrence of three properties. This represents overwhelming evidence against treating passive as a unified phenomenon. The three fundamental properties of passives can show up independently of each other. They can occur in any possible combination.

Argument positions

In classic passive sentences, the apparent subject of a passive is the understood object of the verb, as in the following example:

(30) John was arrested by the police

Movement has taken place from a theta-position to a Case position, in our terms. The next question is: Was it from a Caseless position?

How can we find out whether that position is a caseless position? Let us hold everything else constant (in particular, we control for the EPP by putting an expletive in the subject position of "was arrested" and not doing the movement of "John"):

(31) *It was arrested John by the police

Since the sentence is bad, the position where "John" is is presumably not a Case position.

Now we have to figure out why it is not a Case position. The position of "John" in (31) looks like it is a position where Case could be assigned, since it is the object of the verb. More accurately, it is the object of the passive participle of the verb. Evidently, we have to say that the passive participle of a verb is not a Case assigner, even if the active form is. Let us try to justify this claim.

Here is one conceivable way of approaching this: It had been noted for a long time that passive participles have a lot in common with adjectives. This can be seen, for example, from pairs like "the angry man" and "the arrested man" etc. Further, adjectives are not Case assigners (cf. "*proud John" vs. "proud of John"). So it is not astonishing that passive participles are not Case assigners.

However, passive participles are not identical in their properties to adjectives. If "arrested" were identical to adjectives in being [+V; +N] we would expect that we would be able to do *of*-insertion to save the sentence "It was arrested John by the police" from violating the Case Filter. But this is not possible:

(32) *It was arrested of John by the police

Recall that we had concluded that lexical Case assigners are verbs and prepositions, namely categories that are [-N] and the categories that are *of*-triggers are nouns and adjectives, that is, [+N] categories.

So, how do we analyze passive participles? In Chomsky and Lasnik (1977) a passive participle is taken to be a neutralized verb/adjective. That is, if you look at what a verb is: [+V, -N], and you look at what an adjective is: [+V, +N], you see that they are both [+V]. Chomsky and Lasnik (1977) suggested that when you make a verb into a passive participle, you neutralize the difference between the verbal properties and the adjectival properties. You create a category that is [+V] with no marking at all for N. It will not then trigger *of*-insertion, because it is not +N, and it will not be able to assign Case because it is not -N.

We have now reduced one fragment of passive to Move α. In this instance, the movement is from a theta-position that is not a Case position to a Case position that is not a theta-position. There is a difficulty here though. Semantically one would expect the subject of "arrested" to be a

theta-position, and then movement to it should not be possible. However, there is evidence, independent of movement, that the subject of a passive is not, in fact, a theta-position. To see this, note that there are verbs that can take clausal complements and nominal complements as well, such as "prove".

(33) a. I proved the theorem
b. I proved that 2+2=4

Consider now potential "passives" but without movement. First, with an NP complement:

(34) *It was proved the theorem

As before, (34) is bad, in violation of the Case Filter. This analysis makes a prediction that if we put as the complement of "proved" something that does not need Case, in particular, a clause instead of an NP, then the construction should be fine without movement, at least as far as the Case Filter is concerned. This prediction is correct:

(35) It was proved that 2+2=4

But now we also have evidence that the subject position of a passive predicate is not a theta position, because we <u>are</u> able to get an expletive subject in that position. By definition, you cannot get pleonastic elements in theta-positions.

Gutierrez: We still have to explain <u>why</u> the subject position in passives is not a theta-position.
Lasnik: That is exactly right. In fact, nobody has fully succeeded in doing that. People have fancier and fancier ways of saying it. The phenomenon is that passive predicates, even when they are passives of verbs that have subject theta roles to assign, do not have subject theta roles to assign. There is only one thing in the literature that I know of that has the feel of an explanation, and that is Baker, Johnson and Roberts (1987), where it is proposed that the passive morpheme on the verbs is a sort of a subject clitic, an argument bearing the subject theta-role. That frees up

the structural subject position which is then a non-theta position.[3]

■

By-*phrase*

We are concerned here with elements like the bracketed part in (36):

(36) John was arrested [by the police]

In the *Syntactic Structures* framework, the Passive transformation had the effect of moving the subject to the end of the sentence, inserting *by*, and then we had an optional *by*-phrase deletion rule for "John was arrested". This account was very standard in the 50's and 60's. By the mid 60's people began to worry a lot about *by*-phrase deletion, as to why that does not violate recoverability of deletion. Well, suppose it is always "by someone", then it will not violate recoverability of deletion, as formulated in *Aspects*, because you are specifying in the rule just what you will delete.

In *Aspects,* Chomsky discusses several ways in which deletion might be recoverable. One way is to delete something under identity. A second way, relevant to our problem, is if the rule tells you exactly what you are going to delete. But there were still some problems with that. For example, in a sentence like: "John was run over by a car", we would have to have another rule, namely, of "*by*-something"' deletion. And that still will not quite do it. Suppose we have a situation where John was injured by some collection of people, and we say "John was injured". Now we will need to have a "*by*-some collection of people" deletion rule. And as we construct more and more examples, we need more and more *by*-phrases that can be deleted.

The answer to that is of course to say that we do not really need a *by*-phrase in the first

---

[3]There is also another related mystery. We have seen that you do not have to move the clause "that 2+2=4" to subject position in (35) because clauses do not need Case, and it is the Case Filter that is driving movement. But now (i) is also a good sentence:
(i) That 2+2=4 was proved
It appears that you do not have to move the clause, but you are allowed to move it. In more recent, "economy-driven" versions of the theory, this is problematic since you only do things that you have to do. So, if you do not have to move "that 2+2=4:" in (35), then the question arises as to why you are even allowed to move it as in (i). The EPP could be relevant here.

place, and what that means is that we either have a violation of the Theta-Criterion in "short" passive sentences (with no *by*-phrases) or we say that it is just a property of passive verbs that they do not have a subject theta role to assign.

Vukić: So, what is going on in the long passives (with *by*-phrases)?
Lasnik: I have not come across in the literature a really convincing line on that either. One approach is to say that *by,* like other prepositions, has a particular semantic force. *By* assigns a role to its complement, the way *with* does in a sentence like "build with a hammer", when we have instrumental role. So the question is what role does *by* assign? The likeliest candidate is agent role. In fact, *by*-phrases are often called "agent phrases". That is possible for the examples we have seen so far, but if you look at a wider range of cases it is not adequate.

Note first that "classic" subjects in subject position are not always agents. The prediction of the hypothesized theory is that when you have an active sentence with a non-agent subject, then either you will not be able to make a passive or if you make a passive, this passive will not be allowed to have a *by*-phrase. But that is not true of English:

(37) The news surprised John

"The news" is not an agent in any semantically coherent sense, but the passive sentence is good:

(38) John was surprised by the news

Similarly, in (39) "John" is not an agent, but (40) is still good:

(39) John received the package
(40) The package was received by John

We see that at least in English a *by*-phrase need not be an Agent phrase. Rather, a *by*-phrase in a passive sentence can be anything that a subject in the corresponding active sentence can be. Any role that a subject can have in English in an active sentence will be the same as the role of

the *by*-phrase when you make a passive sentence. A *by*-phrase is, then, like a subject phrase with respect to the Theta-Criterion when it is present, but, strangely, it need not be present.

That is the really curious thing about the *by*-phrase: that it is optional. Given that it is optional, you would expect the *by*-phrase to be just a regular prepositional phrase, meaning whatever it ought to mean based on the semantics of the preposition, but that is not right. In fact, you cannot pin any particular meaning on *by*. Anything the noun phrase could have meant in subject position is still meant in the *by*-phrase, and that is not explained. Notice that it was explained in a sense in the *Syntactic Structures* theory, but there it was more enumerated than explained, listed together with all the other properties of passives.

■

Gutierrez: Could we say that the *by*-phrase is more like *of*-insertion?
Lasnik: Indeed, the presence of *by* looks more like the presence of *of* than anything else. If what I just said is right, the *by* is not making any contribution to the meaning, just like the *of* is not making any contribution to the meaning. But the problem for the Theta Criterion caused by the optionality of the *by*-phrase still remains.

■

Consider now a sentence like:

(41) John is believed to be crazy

Is (41) an instance of passive or raising? In (41) "John" **raises** from the subject position of the predicate "to be crazy" to the subject position of "is believed" which has **passive** morphology. In the mid to late 60's there were intense debates regarding this question. The question now vanishes because there are no longer computational operations like "passive" or "raising". There is just Move α, and (41) is an instance of movement which takes an NP that is not in a Case position and moves it to a Case position. Whether we choose to call it "passive" or "raising" is not a matter of any theoretical or empirical import.

Further aspects of NP movement

Distribution of NPs

So far we have been discussing various aspects of the formal distribution of NPs. According to the X-bar theory, complements to all categories should in principle be the same. Thus, something that is a complement to a V, should also be able to be a complement to an N, or an A, but we found a big gap there. We never found NPs to be complements of N's or A's.

From the point of view of Theta theory, we found that when you look at a verb and a noun with essentially the same meaning, like "prove" and "proof", Theta theory tells you that they ought to have the same complements, but they do not.

Case theory was called on to explain the gaps in those paradigms. Thus, over and above what is predicted by the X-bar theory and by Theta-theory there is a further requirement that has to be satisfied by lexical NPs, and the reasoning is analogous to what we saw when we were looking at verbal morphology. Various principles of phrase structure and transformations tells us where various things might go, but what we discovered was that something we had good reason to analyze as a head could not occur free-standing as a head, so we proposed the Stranded Affix Filter which explains that gap in the paradigm. Similarly, the Case Filter was proposed to explains various gaps in another paradigm.

Case theory, as presented above, consists of two parts. One part is principles of Case assignment, a list of environments where Case is licensed; the other part is the Case Filter (star a lexical NP with no Case). As for the latter, the exemption for non-lexical NPs is crucial for the trace in (42). Without an exemption for trace, movement would never be able to rescue a Case Filter violation.

(42) Bill seems [t to be careful]

An exemption is also needed in (43), but, as we will see, not for a trace:

(43) Bill tries to be careful

Item by item (42) and (43) seem virtually identical. However they are actually very

different. “Try” clearly has a theta-role to assign to its subject. Hence, a movement analysis is not available. Indeed, (44) would violate the Theta Criterion in deep structure or it would violate the principle of recoverability of deletion. If the subject position of “try” had been empty at D-structure, that would have violated the Theta Criterion at D-structure, because “try” has a subject theta-role to assign and there would not have been anything there to get it.

(44) Bill tries t to be careful

Alternatively, if there was a subject of “try” at D-structure, we would have moved something on top of something with content.

In (42), we know that “seems” does not assign a subject theta role, so the movement analysis is available. We know that “seems” does not have a subject theta role because we can have sentences like (45) with expletive subject.

(45) It seems that Bill is careful.

You can also see this in other ways:

(46) It seems to be raining

(47) There seems to be a solution

On the analysis we have been developing, (45)-(47) share a property. In (45) “Bill” started off as the subject of the embedded clause; by analogy then, “it” in (46) started out as the subject of the embedded clause, and in (47) “there” started out as the subject of the embedded clause. So, this makes a prediction: there is no necessary connection between “seems” and “Bill” or “it” or “there”, but there is a necessary connection between “Bill” and “to be careful”, we can say “Bill is careful”, but not, say, “there is careful”. Similarly, there is a connection between “it” and “to be raining”, since you can say “it is raining”, and you cannot say “there is raining”. Finally there is a connection between “there” and “to be a solution”, since you can say “There is a solution”, you cannot say “It is a solution” (with expletive ”it”).

Thus, all the requirements of the surface subject of "seems" were satisfied by the embedded predicate. The matrix predicate imposes no requirements whatsoever. That generalization is overwhelming evidence for the raising analysis. Notice that nothing like that happens with "try":

(48) a. *It tries to be careful,
b. *There tries to be a solution

Idioms provide further evidence. Take the following construction:

(49) The cat is out of the bag

This sentence has two readings. One of them is the literal reading, and the other one is the idiomatic reading ("the secret has been revealed")

(50) The cat seems to be out of the bag

In (50) we still have both readings available. Why should that be? "The cat" clearly has to go with the predicate "to be out of the bag", and even more so for the idiomatic reading. Idioms are sort of lexical items, they have very idiosyncratic meanings, but they are bigger than lexical items, they are phrases or clauses, and ones that are (sometimes) susceptible to syntactic operations . "The cat" starts as the subject of "to be out of the bag" and moves to the subject position of "seems."

Consider now the counterpart of (50) involving "try" instead of "seems":

(51) The cat tried to be out of the bag

In (51) we only get the literal reading. The idiomatic reading is gone, just as one expects if "the cat" must have been the D-structure subject of "tried".

So, what <u>is</u> the analysis of sentences with "try"? We know that "Bill" in (44) is generated as the subject of "try", as that is the only way that the Theta-Criterion is going to be satisfied,

and the only way we can explain all these differences between "try" and "seem". We also know that "to be careful" has a subject theta-role to assign, so there must be an argument there as well. We will just call it ARGUMENT for the time being.

(52) Bill tried [ARGUMENT to be careful]

We have to figure out what this ARGUMENT is. The traditional approach in the 50's and 60s to sentences like this is that this ARGUMENT is "Bill", and a rule deletes it under identity with the first "Bill". Paul Postal (1976) gave the deletion operation the name "Equi-NP deletion".

Let us stick to the classical description for a minute longer. That means there are two ways in which an NP can vacate a position that we know it was in it at D-structure. One is that it can move to another position. Another is that it can be deleted, but only under identity.

Early on it was realized that Equi as deletion under absolute identity was not going to be enough. It is rather tricky to make the argument against this view, because it is going to rely on how you would interpret an ungrammatical sentence if it were grammatical. Nevertheless, let's give it a try. Consider (53):

(53) Everyone tries to be careful

We might paraphrase the interpretation of this sentence in the following way (which is ungrammatical, but we will disregard that for the purpose of this argument)

(54) Everyone tries for himself/herself to be careful

Notice that the following would <u>not</u> be an accurate paraphrase of (53):

(55) Everyone tries for everyone to be careful

So a pure Equi-NP deletion rule will not work. Another view of "Equi" (as in Chomsky and Lasnik (1977)) was that the thing that gets deleted is not a full identical NP. Rather, it is a

reflexive-like element. Then the D-structure of (53) is something like (56):

(56) Everyone tries SELF to be careful

(56) pretty transparently captures the meaning. Then we can say that Equi is SELF-deletion. This analysis was never widely adopted, even though there has never been an argument against it, as far as I know. What was adopted instead was an analysis that said that the argument did not have to be deleted because it did not have phonetic features in the first place. The name Chomsky gave to this element was PRO.

The properties of PRO then are the following:

1) It is an argument;
2) It has no phonetic matrix;
3) It is a sort of reflexive.

Assuming the PRO analysis, the following context indicates that PRO, like trace, does not have to have Case.

(57) a. John tried [PRO to solve the problem]
b. *John tried [Mary to solve the problem]

"Mary" in (57)b is not in any of the positions where Case is assigned, so the example is straightforwardly ruled out.

Note in passing that (as expected so far) along with referential NPs such as "John", expletives like "there" and "it" also must obey the Case Filter, as the following shows:

(58) *It seems <u>there</u> to be a solution
(59) *Bill tried <u>it</u> to rain

Summarizing, this is what we have so far as far as the distribution of NP is concerned:

(60)	*NPs that obey the Case Filter*	*NPs that ignore the Case Filter*
	lexical NPs	PRO
	expletives	NP traces

Case assignment by a complementizer

Chomsky (1980) (*"*On Binding*"*) presented the Case Filter as being relevant to phonetically overt NPs (often called "lexical") as opposed to "silent" NPs. Case in "On Binding" was taken to be a morphological feature, and so things that have to be morphologically realized need it. The Case Filter then, gets to be more and more like the stranded affix filter, just a property of morphology. That sounds really plausible and coherent, so let's pursue the theory further.

Consider next how "John" gets Case in (61):

(61) For John to win would be nice

One possibility is that "For John" is a prepositional phrase and "John" gets Case from the preposition "for". Notice that we can have "For John" by itself in a sentence like the following:

(62) It would be nice for John

Given this it is tempting to think of "for John" in (61) as a prepositional phrase. But consider now (63):

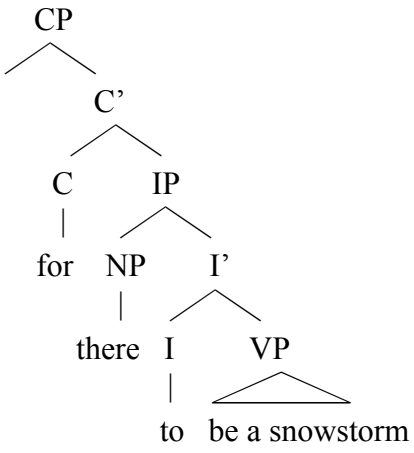
(63) For there to be a snowstorm would be nice

Recall that we have established that "there" has to satisfy the Case Filter. So, once again we might think that the way "there" satisfies the Case Filter is because "for there" is a prepositional phrase. But we cannot find any sentence where "for there" is clearly a prepositional phrase. The reason intuitively is that "for", when it is a preposition, has some theta-role to assign, but 'there' is a pleonastic, so it can never get that theta-role. But if "for there" cannot be a prepositional phrase, then we have two mysteries: why is (63) grammatical;

and what is “for” in that example.

Suppose that “for” is a sentence introducer, a complementizer. The (relevant part of the) structure of (63) would then be as below:

(64)



There is a generalization to be made here, that “that” goes with finite I and that “for” goes with infinitive I. That is just the kind of generalization that our technology can capture. It is a selectional restriction, relating the head of CP to the head of its complement, IP.

That is all very well, but it does not solve the first problem, the apparent violation of the Case Filter. We still do not know how the NP following “for”, as in (61), gets Case.

Government

To attempt to come to grips with the problem above, let us look at Case theory in more detail. So far we just have a list of the configurations where Case can be received.

(65) A, a Case assigner, assigns Case to NP only if:

1) A is a Case assigner

2) A bears a certain structural relation to NP

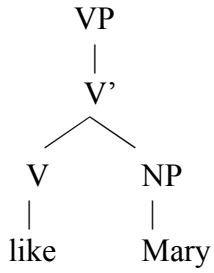
(66)

Case Assigners	*Non-Case Assigners*
Finite Infl	Non-finite Infl
Certain verbs	Passive verbs
Prepositions	Adjectives
	Nouns

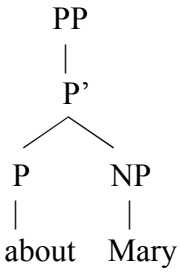
We will try to figure out what this structural relation is about. Even if you have a Case assigner in the sentence, obviously, that does not mean any NP anywhere in the sentence can get Case from it. There are structural positions crucial to Case assignment.

One relation is the head complement relation as in the following:

(67) a.

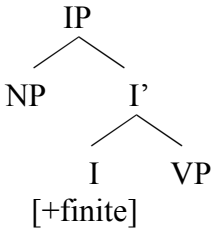


b.



Another relation is the one involving Case of the subject NP of a finite clause, in the position of the specifier of IP:

(68)

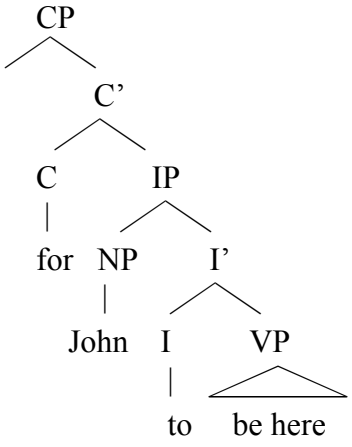


These two relations, head-complement and spec-head relations, are the two core X-bar theoretic relations, the only ones in the X-bar theory.

Given that, it is evident that there is no relation whatsoever between “for” and “John” in

(61) in terms of X'-theory. The (relevant part of the) structure of (61) is:

(69)



This relation is a huge mystery. To address this mystery Chomsky proposed in *LGB* that there is a unified structural relation encompassing all three of these relations that we have seen so far. Chomsky called this relation *government*.

A working definition of government for the moment is:

(70) A governs B only if every maximal projection dominating A also dominates B and conversely.

That is, A and B are contained in all the same maximal projections. In (67)a V and its complement are both within the same maximal projection, namely VP. In (67)b P and its complement are both within the same maximal projection, namely PP. In (68) I and its specifier are both within the same maximal projection, namely IP.

In this system, Case assignment can be stated as follows:

(71) A assigns Case to B only if
1) A is a Case assigner
2) A governs B

This does not yet solve the problem of (69). Even if we say that "for" is a complementizer and a Case assigner, we still do not explain how "John" gets Case, because according to our definitions so far "for" does not govern "John" and government is necessary for Case assignment.[4]

In order to account for (69) we need to change the definition of government to the following:

(72) A governs B only if every XP dominating A dominates B and conversely (unless IP dominates B and not A, in which case ignore IP).

With this definition, we allow "for" to assign Case in (69) to "John" which is the specifier of IP.

It appears that we have a fairly good theory of Case assignment, based on the notion of government, and the notion of government is pretty natural except for the parenthetical. The parenthetical is there for two types of examples, one is sentences like (61), and the other type, which we will look at now, is the following:

(73) I believe [John to be a liar]

Here "John" is the subject of the predicate "to be a liar" and "John to be a liar" is the complement of "believe", just as in (74):

(74) I believe [John is a liar]

If the clausal complement of "believe" is IP, then the same extension of government we used for "for" complements is available here. And since "believe" is a Case assigner (cf. "I believe your story") it will assign Case to "John" in (73). This type of Case assignment is ususally called *Exceptional Case Marking* (ECM).

---

[4] In *LGB* Chomsky treated IP as S and CP as S', so that "for" would govern "John" since the boundary separating them, S, is not a maximal projection. But, as we saw, these structural proposals were very problematic, and were soon abandoned.

Case and PRO

Earlier we posited the existence of an "empty category" (silent category) called PRO. It satisfies a phrase structure principle, the EPP, and it satisfies the Theta Criterion:

(75) I tried [PRO to solve the problem]

Clearly the distribution of PRO cannot be determined solely by its argument character. We have seen PRO as subject of a non-finite clause. But it cannot occur as subject of a finite clause:

(76) *John says [PRO is clever]

Nor can it be the object of a verb or preposition:

(77) a. *John injured PRO
b. *John talked about PRO

One theory of this distribution is that PRO must not be Case-marked. This was suggested by Bouchard (1982) and Manzini (1983). However, there are argument positions where no Case is assigned but where PRO is still impossible, as in the following instance of passive:

(78) It was arrested PRO

Note, though, that if PRO is a sort of an anaphor (as hinted earlier) (78) will be out for that reason: PRO has no antecedent, Also, (79) might be excluded by a locality constraint on this antecedent-anaphor relation:

(79) *$John_i$ said it was arrested PRO_i

Let us look at another example that will be ruled out by this theory of PRO:

(80) *John believes [PRO to be clever]

The reason (80) is bad is not because PRO is an argument, nor because PRO is an anaphor, since we can say "John believes himself to be clever". (80) can be excluded because PRO is Case-marked, "exceptionally" in this instance.

Though this proposal is attractive, it has difficulties. The verb "believe" has the property that it nominalizes into "belief", and when it does, it seems to take semantically the same kinds of complements that "believe" takes.

(81) a. John believes that Mary is clever
b. John's belief that Mary is clever

(82) a. John believes Mary to be clever
b. *John's belief Mary to be clever

(82)b can be ruled out since "Mary" has no Case in that configuration. According to the *LGB* theory, infinitival Infl is not a Case assigner and "belief" is not a Case assigner either.

Now consider (83):

(83) *John's belief [PRO to be clever]

The reason why (83) is bad is not because PRO is an argument, nor because PRO is an anaphor, nor because PRO must <u>not</u> be Case marked, since PRO in (83) is not in a position where Case can be assigned. Now we are ready for the final refinement that will give us Chomsky's fundamental descriptive generalization presented in *LGB*. Suppose we strengthen the statement that PRO must not be Case marked to the following statement:

(84) PRO must not be governed

Case marking entails government but not conversely.

It is very hard to come up with a direct test as to whether we get the government relation in (83). However, there is an indirect argument that we in fact have such a relation in (83). The argument deals with examples like (82)a above, where "believe" governs "Mary" (as evidenced by the fact that it assigns Case to it). If "believe" governs "Mary" in (82)a, then it is reasonable to assume that "belief" governs "Mary" in (82)b and, therefore, PRO in (83). The null hypothesis is that whatever property "believe" has, "belief" also has, as was argued in "Remarks on Nominalizations". Recall that the property that "believe" had that allowed it to govern "Mary" in a sentence like (82)a was that the complement of "believe" is merely an IP, not a CP.

This is not a complete explanation yet. So far our only basis for positing (84) is example (83). Although descriptively it works, the statement (84) is not as conceptually natural as the previous statement (PRO must not be Case-marked), since we had a sort of intuition for why PRO must not be Case marked (namely, because it must not be morphologically realized, if Case is indeed like a morphological affix).[5]

Inherent Case

There is another question that arises about IP complements to nouns. Why can't (82)b be rescued by *of*-insertion. In *Knowledge of Language* (1986) Chomsky provides an account of the ungrammaticality of (85):

(85) *John's belief of Mary to be clever

Contrary to *LGB*, in *K of L*, nouns and adjectives have the capability to assign Case, a special sort of Case which Chomsky called *inherent.* By this, he meant that it is associated with theta-marking, as stated in (86).

(86) X inherently assigns Case to Y only if X theta-marks Y

[5]In *LGB*, Chomsky proposed a way to deduce (84). At this moment, I will not go into it, since it involves a module of the grammar that we have not yet explored, that concerning anaphora.

"Belief" then cannot assign (inherent) Case to "Mary" because it does not theta-mark it.

Chomsky does not discuss this notion in any detail, but his intuition is that in some languages (like the Slavic languages) where some verbs assign accusative, and others assign dative or instrumental, the accusative is "structural", but the dative or instrumental is associated with a particular theta role, and is, therefore, inherent. There is a fair amount of evidence for a fundamental distinction between structural Case on the one hand and inherent Case on the other. Some of it goes back to work of the late 70's and early 80's, in particular, by R. Freidin and L. Babby (1984). They noticed that in Russian a verb that assigns normal accusative Case can be passivized and then the former object gets nominative Case. However, if you have a verb that assigns dative or instrumental or other oblique Cases, you either cannot passivize the verb at all or if you do passivize it, its former complement still has dative or instrumental Case. In that sense, the term "inherent" seems right. It appears to be a fundamental property of the verb that it has to assign this Case.

According to Chomsky's proposal, the realization of the inherent Case in English is "of". So in (87)a "destruction" assigns inherent Case to "the city". Once the Case is assigned, the sentence gets pronounced as (87)b.

(87) a. destruction the city
b. destruction of the city

Notice, by the way, that (85) was probably not the appropriate example after all. "Belief" is apparently unable to assign Case even to its complement:

(88) a. John believes Mary
b. *John's belief of Mary

It is actually very difficult to find a clear instance. Among the few is the following:

(89) *John's proof of Mary to be clever

(89) is clear, since "proof" is capable of assigning inherent Case to a complement:

(90) a. John proved the theorem
b. John's proof of the theorem

Case of wh-*traces*

Recall that at the moment we have the following distribution of NPs with regard to the Case Filter:

(91)	*NPs that obey the Case Filter*	*NPs that ignore the Case Filter*
	lexical NPs	PRO
	expletives	NP traces

It seems that we have an elegant theory of which NPs need Case and which ones do not. In particular,"noisy", that is, morphologically realized, elements have to obey the Case Filter and "silent" things do not. Unfortunately, there is a big descriptive problem with this theory that led Chomsky to abandon it in *LGB*. To understand this problem, we need to consider the phenomenon of *wh*-movement.

Wh-movement is an operation that takes an interrogative NP and moves it to the front of the sentence. In particular, it moves this NP into the position of specifier of CP, which is a non-argument position, usually referred to as A'('A-bar')-position. Movement to an A'-position is called A'-movement.

(92) $[_{CP}$ What $[_{C'}$ will $[_{IP}$ you read t]]]?

It turns out that a *wh*-trace must have Case, contrary to expectations in (91).

Let us look at configurations where you cannot have Case and see if we can rescue the sentence by doing *wh*-movement.

(93) *It was arrested John

(94) *It was arrested who

(95) *Who was it arrested t

(95) is problematic for (91), since the *wh*-trace is "silent" but it seems to need Case. We surely want to say that (95) violates the Case Filter. One plausible possibility is that it is not the trace, but, rather, "who" that needs Case. Then, we can still maintain that Case is needed only for "noisy" NPs. That was the proposal of Chomsky in "On Binding" (1980). We still have a question of how exactly that *wh*-phrase gets Case, and that was the least elegant part of Chomsky's Case account in "On Binding". In particular, Chomsky suggested that although normally Case is assigned at S-structure (in our familiar list of ways), part of the rule of *wh*-movement is such that when you find a *wh*-phrase that you are about to move, first apply the whole Case assignment algorithm, even though it is not S-structure yet, and then do your *wh*-movement. Obviously, you replicate all of Case assignment as part of the *wh*-movement rule, not a very explanatory move. The obvious alternative to this analysis is that *wh*-traces indeed need Case.

There is an empirical argument for this alternative. Relativization in virtually all syntactic respects works just like interrogation. So, along side (96), you have (97).

(96) a. Who was arrested?
b. Who did they arrest?

(97) a. The man who was arrested
b. The man who they arrested

Both constructions clearly involve movement by any test we have ever seen. Superficially, it looks like the trace of relativization has to obey the Case Filter, so we cannot have (98):

(98) *The man [who it was arrested t]

So, here too it looks on the face of it that although the *wh*-trace is silent, it needs Case. Again,

according to Chomsky (1980) it would be "who", a "noisy" element, rather than the trace, that needs Case.

The Chomsky (1980) theory can actually be simplified in a certain respect. Lori Davis (1984) proposed that Case marking can occur whenever you like rather than just at S-structure. Let us assume Davis's theory because it looks quite simple. If you find yourself in a Case position you can get Case. If English were like many other languages, then we would be done and Chomsky's (1980) theory would look great. However, English very peculiarly has a variety of ways to make relative clauses. For example, instead of (97)b you can say (99), without the "who":

(99) The man they arrested

Now, we have got a clear control for our experiment. We can now see whether it is the *wh*-trace that needs Case or whether it is the *wh*-phrase that needs Case. Recall that (98) was ruled out by the Chomsky (1980) theory on the basis that the "who" needs Case. But now notice that (100) is also ungrammatical:

(100) *The man it was arrested

Based on data like this Freidin and Lasnik (1981) concluded that *wh*-traces do need Case.

Case and Visibility

Attempting to solve the problem of having a "silent" NP, namely, *wh*-trace, needing Case, Chomsky in *LGB* advances an alternative theory of Case. He proposes that arguments need Case. In particular, Chomsky deduces the fact that arguments need Case in the following manner:

A) Theta-roles must be assigned not just at D-structure but also at LF
B) To receive a theta-role at LF, an argument must be visible
C) An argument without Case is not visible.

This approach to the Case Filter is standardly called the *Visibility* approach.

In sentence (93), repeated here, we then have a violation of the Theta Criterion. "John" is not Case marked, so it becomes invisible to theta-role assignment.

(93) *It was arrested John

Similarly, in sentence (95) the trace of "who" is an argument, but it is not visible because it has no Case. Consequently, it cannot get its theta-role, and, therefore, violates the Theta Criterion at LF.

(95) *Who was it arrested t

Now, consider a good sentence like the following.

(101) Who did they arrest t?

If you try to translate (102) into some sort of logical language you will get (103), and you can clearly see that there is an argument position. The *wh*-trace gets realized as an argument variable "x", bound by an operator.

(102) For which x, they arrested x

This is not the case for an NP trace in (101), where "John" is the argument, and its trace is not. The trace of "John" marks the position it moved from. If the trace of "John" was also an argument, then we would incorrectly have two arguments, in violation of the Theta Criterion.

Within the theory where arguments need Case, we can thus explain why *wh*-traces need Case. The explanation is straightforward: *wh*-traces are arguments. NP traces are not arguments and so they do not need Case.

Although this theory solves the problem as to why *wh*-traces need Case, it effectively creates two new problems. One problem is that PRO is an argument but apparently does not need

Case, as we have seen. The other problem is that expletives need Case even though they are not arguments.

There are ways people have tried, ever since *LGB*, to deal with the fact that expletives need Case. In my opinion, all of those ways were failures. People did not try to deal with why PRO does not need Case until recently. Chomsky and Lasnik (reprinted in Chomsky (1995)) propose that PRO does need Case, but the Case it needs is a special 'null Case' different from the ones we are familiar with (see more on that in Lasnik (1993) *Lectures on Minimalist Syntax* (reprinted in *Minimalist Analysis*) and Martin (1992, 1996)).