

On Feature Strength: Three Minimalist Approaches to Overt Movement

Howard Lasnik

Procrastinate (Chomsky 1993) favors covert movement; therefore, when movement is overt, it must have been forced to operate “early” by some special requirement, one that Chomsky codes into “strong features.” I compare Chomsky’s three successive theories of strong features and argue that two ellipsis phenomena, pseudogapping and sluicing, provide evidence bearing on the nature of strong features. I argue that movement *or* ellipsis can rescue a derivation with a strong feature, and I conclude that PF crash is relevant either directly, as in Chomsky 1993, or indirectly, as in the theory presented in Chomsky 1995a augmented by the multiple-chain theory of pied-piping (especially as interpreted by Ochi (1998)).

Keywords: strong features, overt movement, ellipsis, pseudogapping, sluicing, pied-piping

Given an economy condition like Procrastinate (Chomsky 1993), which is designed to favor covert movement over overt, we expect no movement to be overt, all else being equal. When movement is overt, rather than covert, then, it must have been forced to operate “early” by some special requirement. Chomsky (1993, 1994, 1995a)¹ codes this requirement into “strong features” and presents three successive, distinct theories of precisely how strong features drive overt movement.

PF crash theory

A strong feature that is not checked in overt syntax causes a derivation to crash at PF. (Chomsky 1993)

LF crash theory

A strong feature that is not checked (and eliminated) in overt syntax causes a derivation to crash at LF. (Chomsky 1994)

Portions of this material were presented at the 1997 Open Linguistics Forum at the University of Ottawa. I am grateful to the audience there, to the participants in my 1997 seminar at the University of Connecticut, and to Željko Bošković, Bob Freidin, Masao Ochi, and an anonymous *Linguistic Inquiry* reviewer for very helpful suggestions.

¹Page references to Chomsky 1993 will be to pages in Chomsky 1995b, where the paper was reprinted. Page references to Chomsky 1994 will be to pages in Campos and Kempchinsky 1995, one of two books where the paper was published (the other being Webelhuth 1995). I use the 1993 and 1994 citations for ease of exposition and to keep the historical development of the ideas of I am exploring clear.

Virus theory

A strong feature must be eliminated (almost) immediately upon its introduction into the phrase marker; otherwise, the derivation cancels. (Chomsky 1995a)

In this article I will bring some ellipsis facts to bear on the question of the nature of strong features. I will begin my exploration by briefly summarizing Chomsky's successive justifications for these three proposals, and the technical implementations of them.

The justification for the PF crash theory is as follows:

... the position of Spell-Out in the derivation is determined by either PF or LF properties, these being the only levels, on minimalist assumptions. Furthermore, parametric differences must be reduced to morphological properties if the Minimalist Program is framed in the terms so far assumed. ... we expect that at the LF level there will be no relevant difference between languages with phrases overtly raised or in situ (e.g., *wh*-phrases or verbs). Hence, we are led to seek morphological properties that are reflected at PF. (1993:192)

In the text and an accompanying note, Chomsky suggests two possible implementations of this approach:

... "strong" features are visible at PF and "weak" features invisible at PF. These features are not legitimate objects at PF; they are not proper components of phonetic matrices. Therefore, if a strong feature remains after Spell-Out, the derivation crashes.³⁸ (1993:198)

38. Alternatively, weak features are deleted in the PF component so that PF rules can apply to the phonological matrix that remains; strong features are not deleted so that PF rules do not apply, causing the derivation to crash at PF. (1993:216)

Whereas the justification for the PF crash theory is conceptual, the justification for the change to the LF crash theory is, as far as I can tell, completely empirical. The relevant discussion is not fully explicit, but what is at issue is evidently the ungrammaticality of sentences like (1).

(1) *John read what?

Assuming that the strong feature forcing overt *wh*-movement in English resides in interrogative C,² the potential concern is that that C might be introduced in the LF component, where, checked or not, it couldn't possibly cause a PF crash. Yet (1) is bad, so such a derivation must be blocked. I quote Chomsky's discussion.

... Spell-Out can apply anywhere, the derivation crashing if a "wrong choice" is made. ... If the phonological component adds a lexical item at the root, it will introduce semantic features, and the derivation will crash at PF. If the covert component does the same, it will introduce phonological features, and the derivation will therefore crash at LF. ... Suppose that root C (complementizer) has

² Notice that the relevant strong feature could not reside in the *wh*-phrase, since in multiple interrogation all but one of the *wh*-elements remain in situ.

(i) Who gave what to who?

a strong feature that requires overt *wh*-movement. We now want to say that unless this feature is checked before Spell-Out it will cause the derivation to crash at LF to avoid the possibility of accessing C after Spell-Out in the covert component. (1994:60)

Note that for Chomsky, the problem is specifically limited to the root—that is, to examples like (1), rather than (2).

(2) *Mary wonders John read what.

This is so since Chomsky assumes that lexical material can only be added at the root. Consequently, a C in the complement of *wonder* in (2) must have been added in the overt syntax, prior to the merger of the complement with *wonder*. But then the PF account would suffice. In (1), on the other hand, C could be added at the root covertly. The new LF account is therefore necessary.³ Chomsky states it as follows:

Slightly adjusting the account in Chomsky (1993), we now say that a checked strong feature will be stripped away by Spell-Out, but is otherwise ineliminable. (1994:60)

Under this approach, it is not the checking operation itself that eliminates a (strong) feature. Rather, checking renders the strong feature eligible to be eliminated by Spell-Out, the latter now being construed as a sort of operation, instead of just the branch point of a derivation. On this account, whether interrogative C is introduced overtly, as in (3), or covertly, as in (4), the strong feature will persist to the LF interface level since it was not checked prior to Spell-Out.⁴

(3) Spell-Out: C [strong Q] John read what *LF

(4) Spell-Out: John read what

LF: C [strong Q] John read what *LF

Chomsky (1995a) rejects the PF crash theory on conceptual grounds, and the conceptual argument he gives applies equally to the LF crash theory. Thus, he rejects any such account as an ‘‘evasion’’ and proposes what he claims is a more straightforward statement of the phenomenon, here called the virus theory. (Juan Uriagereka (personal communication) suggests this felicitous term. This use of *virus theory* is distinct from Sobin’s (1997) use of the same term.)

... formulation of strength in terms of PF convergence is a restatement of the basic property, not a true explanation. In fact, there seems to be no way to improve upon the bare statement of the properties of strength. Suppose, then, that we put an end to evasion and simply define a strong feature as one that a derivation ‘‘cannot tolerate’’: a derivation $D \rightarrow \Sigma$ is canceled if Σ contains a strong feature ... (1995a:233)

³ Željko Bošković (personal communication) reminds me that Chomsky’s (1993) theory of the organization of the grammar, as actually stated, avoids this problem, since Chomsky simply stipulates that ‘‘[a]fter Spell-Out, the computational process ... has no further access to the lexicon ...’’ (p. 189). In Chomsky 1994 he was explicitly concerned to eliminate that stipulation.

⁴ Note that it is irrelevant whether the strong feature in C is checked by covert movement, as it will still not be ‘‘stripped away.’’

Chomsky summarizes this approach as follows:

A strong feature . . . triggers a rule that eliminates it: [strength] is associated with a pair of operations, one that introduces it into the derivation . . . a second that (quickly) eliminates it. (1995a:233)

Later I will discuss this approach in further detail, showing, in particular, that it does not adequately address the empirical argument Chomsky gave for rejecting the PF crash theory.

Given that the PF crash theory concerns PF, ellipsis potentially provides new evidence bearing on its correctness, if, as Chomsky has consistently maintained over the years, ellipsis involves a PF deletion process.⁵ I am aware that a PF analysis of ellipsis is not uncontroversial. In fact, it is quite widely rejected in favor of an LF copying theory, so to the extent that the arguments I will present assume a PF theory of ellipsis, they will have to be regarded as highly tentative. However, I might note that there is actually little in the way of conclusive evidence against (or for) the PF theory. Perhaps the most important argument for an LF approach is the one developed by May (1985 and other works). The essence of the argument is that an LF process, Quantifier Raising (QR), feeds ellipsis resolution in antecedent-contained deletion constructions. Therefore, ellipsis resolution must itself be an LF process. Although this is, on the face of it, a very powerful argument, I might note that Hornstein (1994) argues that the crucial process is not actually QR but raising to [Spec, Agr₀], and, as briefly discussed below (and at greater length in Lasnik 1995a,c), there is reason for thinking that that process operates in overt syntax.⁶

Although an LF copying theory is now rather standard, Chomsky and Lasnik (1993) suggest, and Tancredi (1992) develops, a PF theory, largely based on interpretive parallels between elliptical constituents and deaccented ones.⁷ In this, Chomsky and Lasnik and Tancredi were actually resurrecting an old account of Chomsky's from the late 1960s and early 1970s. For example,

⁵ Eventually, I will suggest a reinterpretation of the virus theory under which it, too, will have a significant PF aspect.

⁶ Actually, I do not believe that Hornstein's approach can be the entire solution. See Lasnik 1993, in press, and especially Kennedy 1997 for discussion. But even its partial success suggests that an alternative to the QR analysis might ultimately be possible. Note, by the way, that if covert movement processes affect only formal features, as is sometimes proposed on the basis of Chomsky 1995a, then an LF process such as QR could not possibly newly create a configuration licensing ellipsis, an argument made in Lasnik, in press.

Two other recent arguments are worthy of comment. First, Kennedy and Merchant (1997), following Haik (1987), note that certain ellipsis constructions show sensitivity to island effects, and they seem to tentatively conclude from that fact that ellipsis involves LF copying. This is somewhat ironic, since Ross (1969) presents a classic argument for deletion, and against copying, based on obedience to island constraints.

Next, Chung, Ladusaw, and McCloskey (1995) argue that sluicing must be an LF process since it is sensitive to purely semantic properties of the antecedent. However, Romero (1997) shows that the phenomenon discovered by Chung, Ladusaw, and McCloskey depends on focus properties and that it actually shows up in the nonelided counterpart of the example they consider, contrary to their claim.

⁷ Such parallels had already been noted in Lasnik 1972, where numerous examples were presented, including the following:

- (i) a. John wants to catch a fish.
- b. John wants to catch a fish and so does Bill.
- c. John wants to catch a fish and Bill wants to catch a fish also.

It is a standard observation that (ia) is ambiguous, *a fish* being specific or nonspecific on some accounts. Yet (ib) is just two-ways ambiguous, not four. The interpretive similarity between (ib) and (ic) shows that this missing ambiguity phenomenon is not limited to ellipsis. Chomsky and Lasnik (1993) note that some condition (which they call PARR; in

Wasow (1972) cites a 1971 lecture where Chomsky “suggest[ed] that VP deletion and Sluicing can be formulated as very late rules which delete unstressed strings.” As it happens, these are just the two ellipsis processes I will be considering: first pseudogapping (a variant of VP-ellipsis, I believe), then sluicing.

(5) presents a few examples of pseudogapping from the classic study by Levin (1978).

- (5) a. If you don't believe me, you will the weatherman.
 b. I rolled up a newspaper, and Lynn did a magazine.
 c. Kathy likes astronomy, but she doesn't meteorology.

Although in many instances it might appear that the process is simply elision of the main verb, there is considerable evidence that more is involved. In the examples in (6), the ellipsis site includes the main verb plus (a) the small clause predicate or (b) the second object in a double object construction.

- (6) a. The DA proved Jones guilty and the Assistant DA will **prove** Smith **guilty**.
 b. ?John gave Bill a lot of money, and Mary will **give** Susan ~~a lot of money~~.

Rejecting the possibility of an ellipsis rule affecting a discontinuous portion of the structure, Jayaseelan (1990) proposes that pseudogapping constructions result from VP-ellipsis, the remnant NP having moved out of the VP by heavy NP shift. In Lasnik 1995c I argue that this proposal is correct in its essentials, though wrong in certain details. In particular, I modify Jayaseelan's analysis by positing raising to [Spec, Agr₀], instead of heavy NP shift, as the process removing the remnant from the ellipsis site.

Before proceeding, I would like to discuss a bit further the general analysis of pseudogapping as a special case of VP-ellipsis. In the first detailed discussion of pseudogapping that I have seen, Levin (1978) notes certain apparent differences (to which I will return) between pseudogapping and VP-ellipsis, but nonetheless concludes that pseudogapping is in fact VP-ellipsis, suggesting that the differences might follow from properties outside of the syntax.⁸ Levin also cites Stump 1977, a work I have not seen, as arguing for a VP-ellipsis account. But later Levin (1979) expresses skepticism. She voices a number of concerns, perhaps most significant among them that “backward” pseudogapping is very severely degraded, unlike backward VP-ellipsis. The following pair is representative:⁹

Lasnik 1972 it was called the Parallel Principle) is needed to guarantee the parallelism observed in an example like (ic). They go on to suggest that under a PF account of ellipsis, nothing further need be said about an example like (ib). This constitutes an (admittedly minor) argument for the PF approach.

Bošković (to appear a) presents another type of argument that at least one ellipsis process is a PF phenomenon. Bošković provides compelling evidence that one crucial aspect of the distribution of Serbo-Croatian clitics is determined by a phonological constraint. He then shows that constructions that violate the requirement can be rescued by VP-ellipsis. This is abstractly very similar to an argument I will present below.

⁸ Sag (1976), too, had briefly considered pseudogapping (not yet known by that title) and had also tentatively suggested that it is VP-ellipsis, while, as Levin later did, acknowledging certain differences.

⁹ I am somewhat misrepresenting Levin's point here, since her claim is that the ungrammaticality of backward pseudogapping is simply a consequence of the ungrammaticality of pseudogapping in subordinate clauses. But here, I (and my informants) strongly disagree with the factual claim. (7) seems far worse than its forward counterpart.

- (i) ?Mary interviewed Gingrich because John did Clinton.

(7) *Because John did Clinton, Mary interviewed Gingrich.

(8) Because John didn't , Mary interviewed Gingrich.

Interestingly, though, in notes that Levin added for the published version of Levin 1979, she states,

... I now believe it doesn't make much difference whether [Pseudogapping] is given separate treatment or collapsed with VP Deletion. Some of the restrictions on Pseudogapping would not need to be reflected in the rule, but could be relegated to the discourse component. (1986:89)

This vague hint could, I believe, be turned into a plausible line of inquiry. The pseudogapping construction involves strong contrastive focus, and it is conceivable that that property conflicts with the backward version of the construction, though there is not space here to try to give that speculation substance. I might note, though, that a purely syntactic account making pseudogapping ellipsis of something other than VP does not seem particularly promising, on the face of it. Suppose pseudogapping turns out to be YP-ellipsis for some $Y \neq V$. And suppose ellipsis can operate backward, as evidenced by the facts of VP-ellipsis. What purely syntactic factor could then prevent YP-ellipsis from operating backward? In light of these considerations, I regard the (admittedly robust) contrast in (7)–(8) as rather weak evidence against a VP-ellipsis account.¹⁰

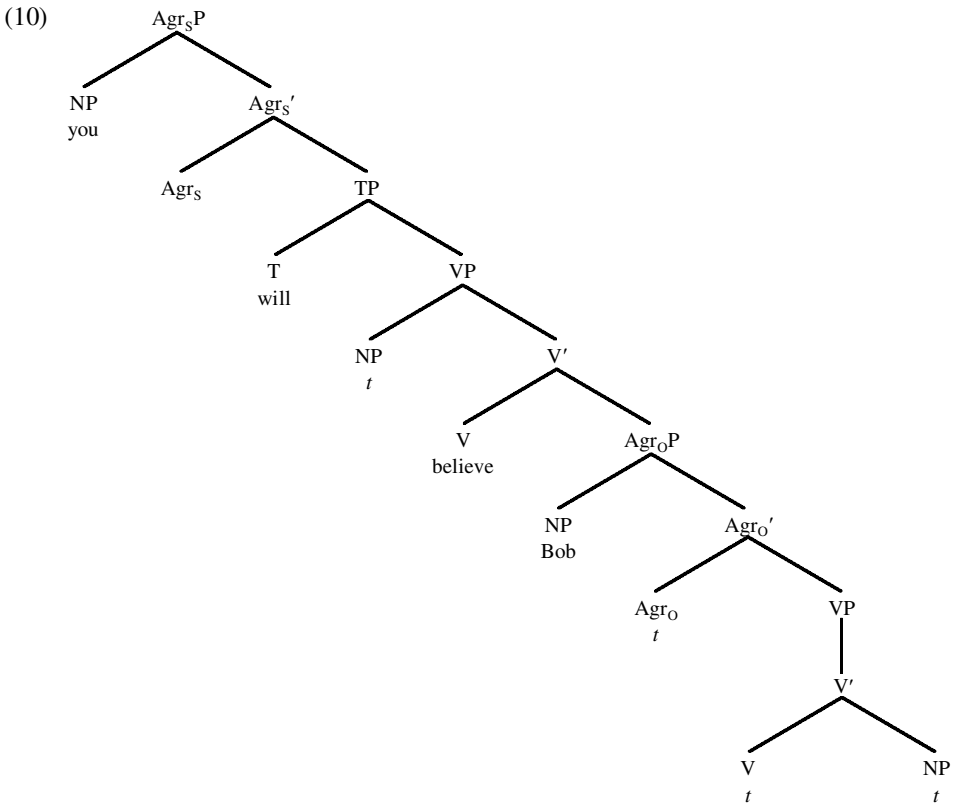
Returning to the raising process rescuing the remnant from deletion, I note that under standard assumptions, raising to [Spec, Agr₀] is covert, taking place in the LF component. Given Jayaseelan's (1990) goal, adopted also in Lasnik 1995c, of analyzing pseudogapping as affecting a constituent, the ellipsis process must then be analyzed as copying in the LF component, rather than deletion in the PF component. However, on the theory of LF movement advocated by Chomsky (1995a), and further defended in Lasnik 1995a,b, the necessary structure would not even be created in *covert* syntax. On that theory, since movement is invariably triggered by the need for formal features to be checked, all else being equal only formal features move. When movement is overt (triggered by a strong feature), PF requirements demand that an entire constituent move, via a sort of pied-piping. However, when movement is covert, PF requirements are irrelevant, so economy dictates that movement *not* affect the entire constituent. But then it is very difficult to see how covert raising of (the formal features of) accusative NP to [Spec, Agr₀] could possibly create an ellipsis-licensing configuration.

It seems then that if (feature-driven) movement newly creates a configuration licensing ellipsis, the movement must be overt rather than covert. Before I indicate how that is possible in the present instance, I note that if the movement is overt, then the conclusion above, that ellipsis must involve LF copying, no longer follows. If the licensing configuration must be created prior to the LF/PF split regardless, then ellipsis could just as easily be a PF deletion phenomenon.

¹⁰ As an anonymous *Linguistic Inquiry* reviewer observes, though, it will not be crucial to the following discussion that YP actually be VP. All that will be crucial is that the remnant move out of some phrase that undergoes subsequent deletion.

Now early Minimalist Program literature (e.g., Chomsky 1991, 1993, Chomsky and Lasnik 1993) did have accusative NP raising to [Spec, Agr_O], but covertly rather than overtly. However, Koizumi (1993, 1995), developing proposals by Johnson (1991), argues instead that that raising is always overt, driven as usual by a strong feature. In Lasnik 1995c I suggest that the strong feature in this instance is an “Extended Projection Principle feature” residing in Agr, hence the same feature that drives overt subject raising.¹¹ I will have little more to say here about this particular strong feature. I will, however, address another strong feature that must be involved in simple transitive sentences without ellipsis. Given that word order in English is V-O rather than O-V, if the complement raises out of VP, the verb must normally raise still higher. Koizumi’s specific proposal, which he calls the *split VP hypothesis*, is that V raises, via Agr_O, to a higher “shell” V position, as shown in (10) for the sentence in (9).

(9) You will believe Bob.

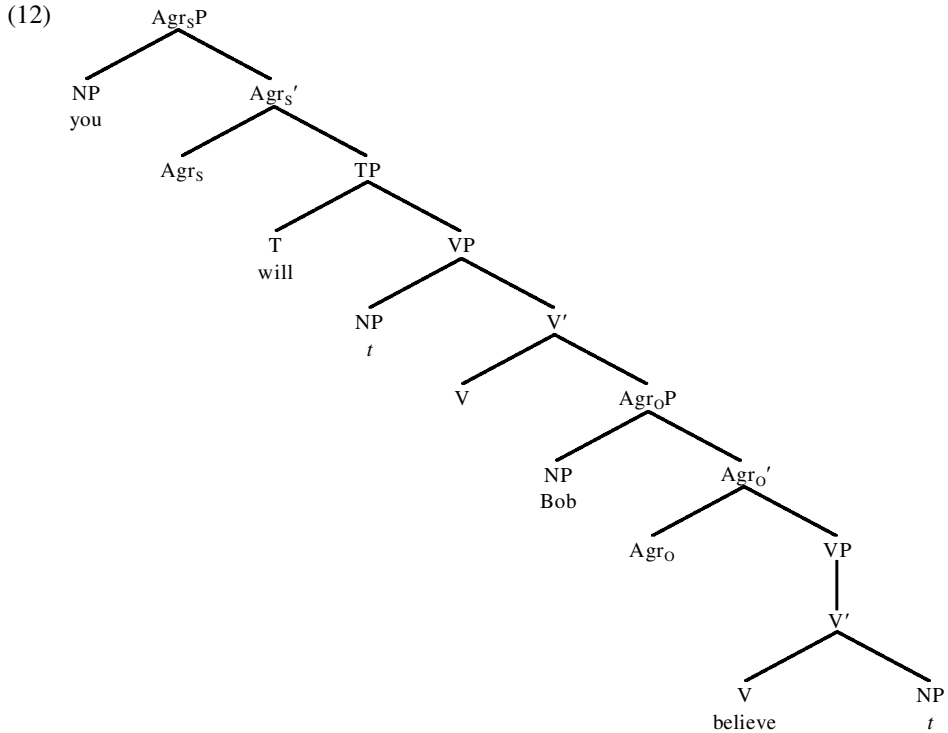


¹¹ In that paper I further suggest that the raising is not, in principle, limited to accusative NPs. Rather, all else being equal, complements more generally undergo such raising, including PP complements (an instance I discuss there) and clausal complements.

Under this general hypothesis, consider a simple pseudogapping example such as (11).

(11) You might not believe me but you will Bob.

If *Bob* overtly raises to [Spec, Agr_O] while *believe* remains in situ, then deletion of the residual VP produces (11). The relevant structure is shown in (12).

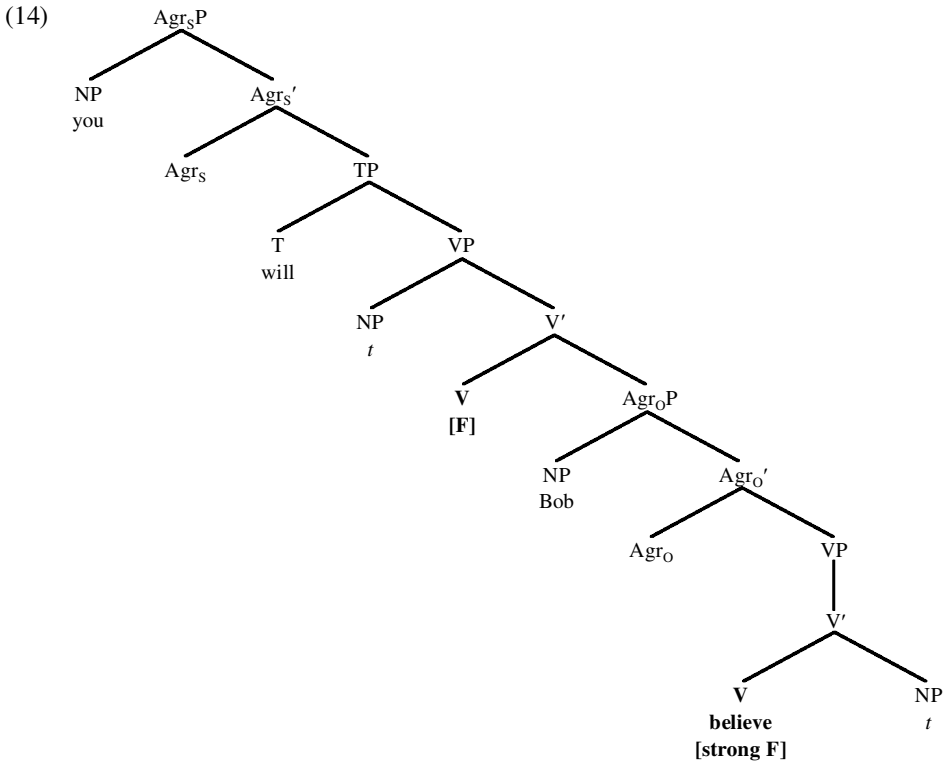


Deletion of the lower VP yields the pseudogapping example in (11).

The question that now arises is why the V need not raise in pseudogapping constructions, given that in nonelliptical sentences it must.

(13) *You will Bob believe.

By hypothesis, a strong feature is involved. Yet there seem to be two possibilities for a convergent derivation. The V can raise as in (10), presumably checking the relevant strong feature. Alternatively, the V can be deleted along with its containing VP as in (12). This mysterious state of affairs receives a rather straightforward account under the PF crash theory of strong features, under the new hypothesis that the strong feature forcing the V to raise overtly is a feature of the lexical V itself, rather than of the target position it raises to. The overt raising derivation is essentially unaffected by this change in perspective. The ellipsis structure is much more interesting. Consider (12) again, but from the point of view offered by (14).



If *believe* fails to raise, and no other relevant process takes place, the strong feature that is not overtly checked causes (14) to crash at PF. But if the lower VP containing *believe* is deleted in the PF component, then, patently, the strong feature cannot cause a PF crash, since the (category containing the) feature will be gone at that level.¹² It is not obvious how to capture this result

¹² An anonymous reviewer for *Linguistic Inquiry* provides several examples suggesting that sometimes the pseudogapping remnant must raise very high and that, therefore, sometimes considerably more than just a V must raise. Consider the following instances of pseudogapping:

(i) While she didn't want to read *War and Peace*, she did *Bleak House*.

(ii) I tried to steal the Rembrandt, but I didn't the Picasso.

The point is that in the nonelided versions (or, for that matter, in the antecedent clauses), seemingly what must be raised is *want to read* or *try to steal*, rather than just a simple verb, since those are the sequences that are missing in the elided versions. Actually, I am not certain just how accessible those readings are, as opposed to ones where just the lower verb is missing. Assuming that they are possible, I will briefly speculate on what the derivations might be. I note first that long A-movement out of control clauses is hardly unprecedented; see, for example, Nemoto 1993 for extensive discussion of such movement in Japanese. So the ellipsis itself is not necessarily problematic. As for the raising in the nonelided

under the LF crash theory of strong features or the virus theory (though later I will suggest that the latter might be rendered compatible). Thus, we apparently have an argument for the PF crash theory.

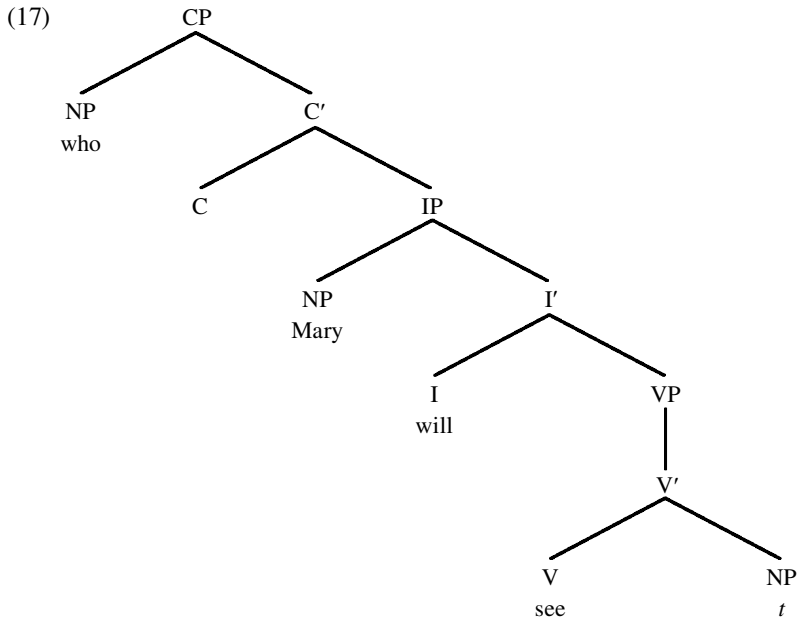
Sluicing, an ellipsis phenomenon first investigated in detail by Ross (1969), displays an abstractly similar paradigm. Saito and Murasugi (1990) and Lobeck (1990) very plausibly analyze sluicing as *wh*-movement followed by IP-ellipsis. (15) displays a representative example.

- (15) Speaker A: Mary will see someone.
 Speaker B: I wonder who ~~Mary will see~~.

Ross characterized sluicing as an embedded-question phenomenon, and the standard examples, like (15), accord with that characterization. However, sluicing also shows up in matrix interrogative contexts, as in (16).

- (16) Speaker A: Mary will see someone.
 Speaker B: Who ~~Mary will see~~?

The structure of the matrix sluicing example (16) is presumably (17) (with irrelevant details suppressed).



clauses, two possibilities come to mind. First, as a result of “restructuring,” verb sequences in control structures might be able to behave as if they constituted simple verbs and raise accordingly. Alternatively, suppose that such raising of complex verbs is *not* possible. Then one might assume that deletion is the only option available, once the remnant has raised high. Nonelided versions would then be the result of normal short raising of complement and verb, internal to the control clause. Needless to say, there is much more to be said about this type of optionality. I hope to explore it further in future work.

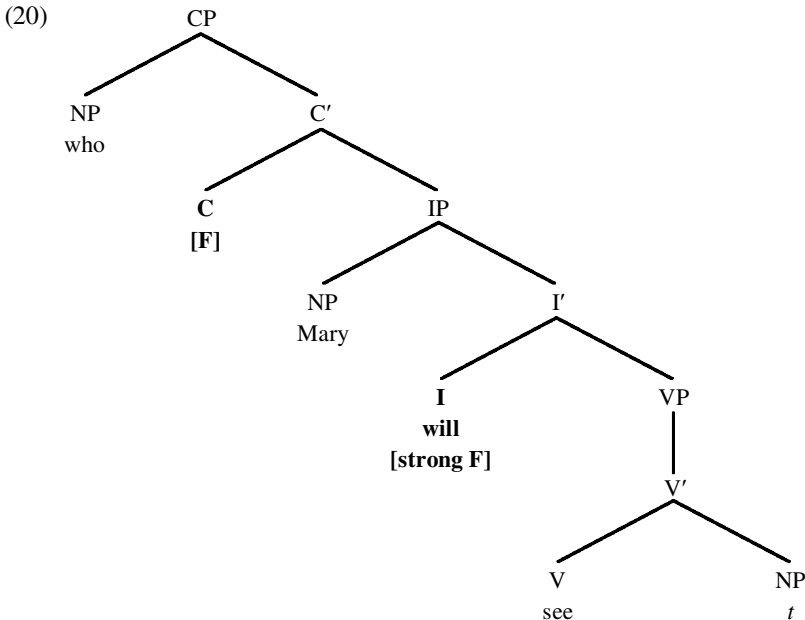
Note that under the assumption that sluicing is, as standardly assumed, IP-ellipsis,¹³ the source for the sluicing example must be as indicated above, rather than, for example, (18).

(18) Who will Mary see?

For (18) to be the source, the elided material would have to be C', and not IP. But now an interesting, and interestingly familiar, question arises. Since the source of matrix sluicing does not have I-raising to C, why is the source ungrammatical without sluicing?¹⁴

(19) *Who Mary will see?

Clearly, in matrix questions like (18) there must be a strong feature driving the overt raising of I. But in the matrix sluicing example (16) that strong feature has evidently not been checked in overt syntax. This is highly reminiscent of the situation with pseudogapping discussed above, where the lexical verb is normally required to raise to the higher shell V position, but does not have to raise if the VP containing it is elliptical. Similarly, in the matrix sluicing case I-raising to C, normally obligatory, does not take place if IP is elliptical. A parallel account is available. A priori, the strong feature forcing overt I-raising in matrix interrogatives might be a feature of C or of I. If it is a feature of I, and if we continue to assume that ellipsis is a PF deletion operation, then the facts fall nicely within the purview of the PF crash theory. To see this, consider again (17), but with the strong feature indicated as in (20).



¹³ A virtual necessity if an ellipsis site must be a maximal projection.

¹⁴ There is another mystery surrounding matrix sluicing. Suppose that I raises to C and that the ellipsis site is IP.

If I fails to raise and nothing further happens in the derivation, a PF crash results, (19), the strong feature of I remaining unchecked. But if the IP is deleted, then nothing remains of the strong feature, checked or unchecked, at PF, so well-formed sluicing, (16), results.

Thus, for two separate constructions, the generalization is the same: an instance of normally obligatory overt movement does not take place (overtly) if the moved item is contained in an ellipsis site. And for both, the same natural account emerges:

- (21) a. Ellipsis is PF deletion.
 b. An unchecked strong feature causes a PF crash.
 c. In the constructions investigated here, the relevant strong feature resides in the item that (potentially) moves, rather than in the target.

With respect to (21c), I hypothesize that the strong feature is in matrix interrogative I, for matrix interrogatives with or without sluicing; and in lexical V, for transitive constructions with or without pseudogapping. The hypothesis regarding lexical V is based on Koizumi's (1993, 1995) approach to clause structure whereby accusative NP always raises overtly. For Koizumi, V invariably raises overtly as well, but I have argued that there are circumstances where this does not happen.

Although I have claimed that the two situations are parallel, that does not reflect the more common view. I-raising to C is quite widely taken to be overt in normal matrix interrogatives in English. On the other hand, apart from Koizumi's work, NP-raising to [Spec, Agr₀], and hence V-raising to a higher position, is standardly assumed to be covert in English. However, in Lasnik 1995b, based on Lasnik and Saito 1991 (see also Postal 1974 and Vanden Wyngaerd 1989) and Den Dikken 1995, I have argued at length that such movement is indeed overt. Rather than repeating all of the arguments here, I will merely summarize their thrust and give a representative example. First, I have already argued above that if raising is to create a new ellipsis configuration, the raising must be overt, given the feature movement theory of covert movement. Second, there is an overwhelming generalization that with uncontroversial overt raising, the raised NP displays "high" behavior in all respects, whereas with covert raising (in existential constructions), the "associate" of the expletive displays "low" behavior. This contrast is shown for reciprocal binding in (22).

The result should be as in (i), but this is ungrammatical.

- (i) Speaker A: Mary will see someone.
 Speaker B: *Who will Mary see?

The proposal of Saito and Murasugi (1990) and Lobeck (1990) about the specific way ellipsis is licensed might be relevant here. They suggest that for a head to license ellipsis of its complement, that head must agree with its specifier. In the sluicing example in (i), the licensing head is C. Now the content of C is the raised I (T and Agr), which obviously agrees with the subject, but does not obviously agree with the specifier of CP. I leave for future investigation the task of making this speculation more precise. Note, by the way, that (ii) is much improved.

- (ii) Speaker A: Someone will see Mary.
 Speaker B: Who will see Mary?

But that is to be expected, since here there is a possible derivation via VP-ellipsis, one that does not involve sluicing at all.

- (22) a. Some linguists seem to each other [*t* to have been given good job offers].
 b. *There seem to each other [*t* to have been some linguists given good job offers].

The explanation offered in Lasnik 1995b for this contrast is as follows: When movement is covert—hence, following Chomsky (1995a), affecting only formal features—the referential and quantificational properties needed to create new binding and scope configurations are left behind, so no such new configurations are created.

Crucially, both simple direct objects and exceptional-Case-marked subjects pattern with the overtly raised NP in (22a) rather than with the covertly raised NP in (22b).

- (23) The DA questioned two men during each other’s trials.

- (24) The DA proved [two men to have been at the scene] during each other’s trials.

Arguably in (23), and almost certainly in (24), the base position of the antecedent of the reciprocal is too low for binding to obtain. Hence, raising is almost certainly involved. Given (22), there is strong reason to believe that the raising is overt. And then, given the word order of English, the verb must also have raised overtly. Thus, there is a reasonably firm basis for the proposals made in this article.

With the ellipsis argument for the PF approach to strong features in hand, at this point it is necessary to consider possible arguments against that approach. One argument is that “look-ahead” is needed. At a given point in the overt portion of a derivation, it is apparently necessary to inspect the PF representation to see whether Procrastinate can be evaded.¹⁵

A second argument, Chomsky’s original one, is that the ungrammatical (1), repeated here, has the possible well-formed derivation in (25).

- (1) *John read what?
 (25) Spell-Out: John read what
 LF: C [strong Q] John read what

As noted above, if C with its strong feature is inserted in the covert component, at the level of PF that strong feature will not exist, hence cannot possibly cause a PF crash.

Chomsky’s LF crash theory addresses the second of these arguments but has nothing to say about the first. The virus theory, repeated here, purports to deal with the first argument and indirectly with the second, as well as with the claimed conceptual problem that the PF and LF crash theories are just evasions.

Virus theory

A strong feature must be eliminated (almost) immediately upon its introduction into the phrase marker.

Momentarily, we will see how Chomsky makes the virus theory precise. Note first, though, that the virus theory demands something I have already necessarily rejected: that a strong feature is

¹⁵ The LF crash theory shares this problem.

always a property of the target of movement, never of the moved item. This is so since if an item that is to move were to have a strong feature, that feature could not, in general, be eliminated immediately. In some derivations the target that it would check against would be far, even indefinitely far, away. Thus, to the extent that the ellipsis analyses I have presented are well supported, there is already reason to reject the virus theory. However, Chomsky's instantiation of this theory is of considerable import, so I will turn to it now.

Chomsky makes the following suggestion:

The intuitive idea is that the strong feature merged at the root must be eliminated before it becomes part of a larger structure by further operations. (1995a:234)

After considering how derivations work in general, he indicates that

the descriptive property of strength is [(26)]. Suppose that the derivation D has formed Σ containing α with a strong feature F . Then

[(26)] D is canceled if α is in a category not headed by α . (1995a:234)

Chomsky observes two very interesting properties of this approach: (a) that cyclicity follows;¹⁶ (b) that a strong feature is checked by an overt operation.

We . . . virtually derive the conclusion that a strong feature triggers an *overt* operation to eliminate it by checking. This conclusion follows with a single exception: covert merger (at the root) of a lexical item that has a strong feature but no phonological features . . . (1995a:233)

This exception involves a kind of example we have seen before.

(1) *John read what?

Recall that it was this sort of example that led Chomsky to reject the PF crash theory in favor of the LF crash theory.¹⁷ But, as Chomsky in effect acknowledges, the problem now arises anew in the virus theory. How can derivation (25), repeated here, be blocked?¹⁸

(25) Spell-Out: John read what
 LF: C [strong Q] John read what

¹⁶ At least for overt movement, though Chomsky does not add this qualification.

¹⁷ The situation is actually more complicated, since there are languages, such as French, that have *wh*-movement of the English sort, but only optionally in matrix questions. See Bošković, to appear b, for extensive discussion.

¹⁸ As Máire Noonan (personal communication) has pointed out, even *overt* insertion of C in the matrix without overt *wh*-movement seems to be incorrectly allowed by Chomsky's formulation.

[A] strong feature merged at the root must be eliminated before it becomes part of a larger structure by further operations. (1995a:234)

Chomsky elaborates this as follows:

Suppose that the derivation D has formed Σ containing α with a strong feature F . Then . . . D is canceled if α is in a category not headed by α . (1995a:234)

When, as in the example now under consideration, the interrogative will not be embedded, hence will never be part of a larger structure, nothing demands that the strong feature be checked overtly.

To prevent this, covert insertion of strong features must be barred. Chomsky proposes to do this with the economy principle (27).

(27) α enters the numeration only if it has an effect on output.

Chomsky elaborates on this as follows:

With regard to the PF level, *effect* can be defined in terms of literal identity. . . . α is selected only if it changes the phonetic form. At the LF level the condition is perhaps slightly weaker, allowing a narrow and readily computable form of logical equivalence to be interpreted as identity. (1995a:294)

This immediately raises a question concerning the central argument for the virus theory—that it eliminates the look-ahead inherent in the PF and LF crash theories. There seems to be considerable look-ahead here, all the way from the very beginning of the derivation, the numeration, to the very end, phonetics and semantics.

Under [(27)], the reference set [for economy comparisons] is still determined by the numeration, but output conditions enter into determination of the numeration itself . . . (1995a:294)

Apart from this conceptual question, there is an empirical question about whether the correct result is in fact obtained. There is reason to think that it is not.

Consider the situation at issue, insertion in the LF component of interrogative C in English, a language in which C has a strong *wh*-feature. (27) purports to prevent this. The first question is whether this C has an effect on output. Clearly, covert insertion of a C will have no phonetic effect. Will it have an effect at the LF output? Either it will or it will not. If it will (apparently Chomsky's intention), then covert insertion is allowed, and we generate (1) with structure (28).

(28) C [_{IP} John read what]

Since this is not the correct result, suppose instead that C will not have a semantic effect. Then we cannot generate (1) with structure (28), so the problem is apparently solved under the assumption that insertion of interrogative C has no effect on semantic output. As Chomsky states the situation:¹⁹

. . . the interface representations (π , λ) are virtually identical whether the operation takes place or not. The PF representations are in fact identical, and the LF ones differ only trivially in form, and not at all in interpretation. (1995a:294)

But our goal is actually more general than just ruling out (1) with structure (28). Rather, it is ruling out (1) altogether. Under the assumptions just spelled out, (1) is successfully excluded

¹⁹ Here I am somewhat reinterpreting what Chomsky actually said, since prior context indicates that he was referring to the operation of "insertion of strong features." But I do not see how to fit such an operation (insertion of strong features independently of the item of which they are features) into the theory. Possibly I am missing something crucial.

with C covertly inserted. But what if C is not inserted at all? That is, what if the structure is the same at both LF and “S-Structure”?

(29) [_{IP} John read what]

(29) violates no morphological requirements, and, if C has no effect on output, the assumption that was necessary in order to exclude (1) *with* C inserted, then it should mean exactly *What did John read?* So if C has a semantic effect, inserting it in LF should be permitted. And if it does not have a semantic effect, *not* inserting it should be of no consequence. Thus, even given the new economy condition (27), (1) is allowed, and allowed as a standard interrogative, presumably an incorrect result. In this regard too, then, the PF crash theory of strong features fares no worse than the virus theory. Either one demands an additional stipulation, perhaps just that lexical insertion is prohibited in the covert component²⁰ (a result Chomsky was trying to deduce, but, as we have just seen, not completely successfully).

At this point there are no clear arguments in favor of the virus theory of strong features over the PF crash theory. Further, the ellipsis paradigms discussed above seem to provide considerable support for the latter approach. Interestingly, though, there is a potential way to reconcile the ellipsis facts with the virus theory. I will end my investigation by showing how such an account would work. Recall that Chomsky (1995a) proposes that strength is always a property of an “attracting” head, never a property of the item that moves. This is necessary under the virus theory (at least as Chomsky articulates it) since a strong feature in an item to be moved would never be checked quickly enough to keep the derivation from terminating. The above analyses of pseudogapping and sluicing are incompatible with that proposal, demanding, as they do, that the moved item sometimes have the strong feature. There is a possible alternative analysis, based on Chomsky’s (1995a) theory of pied-piping, particularly as explicated by Ochi (1998).²¹

Ochi, following Chomsky, considers the nature of pied-piping, the usual reflex of movement triggered by a strong feature. Chomsky (1995a) gives the following characterization:

For the most part—perhaps completely—it is properties of the phonological component that require such pied-piping. Isolated features and other scattered parts of words may not be subject to its rules, in which case the derivation is canceled; or the derivation might proceed to PF with elements that are “unpronounceable,” violating FI [Full Interpretation]. (1995a:262)

Overt movement consists of a complex of operations under this approach.

Applied to the feature F, the operation Move thus creates at least one and perhaps two “derivative chains” alongside the chain $CH_F = (F, t_F)$ constructed by the operation itself. One is $CH_{FF} = (FF[F], t_{FF[F]})$, consisting of the set of formal features $FF[F]$ and its trace; the other is $CH_{CAT} = (\alpha, t_\alpha)$, α a

²⁰ Or at least lexical insertion of an item with a strong feature.

²¹ Ochi’s concern is the locality of movement, in particular, the fact that only Relativized Minimality effects follow in any natural way from Attract F. Other island effects seem to make sense only from the point of view of the moving item, rather than the target. Ochi proposes that the feature chain, created by Attract F, is responsible for the Relativized Minimality effects whereas the pied-piping chain, created by Move α in order to remedy the defect in α created by the movement of the formal features out of α , is responsible for other island effects.

category carried along by generalized pied-piping and including at least the lexical item LI containing F. CH_{FF} is always constructed, CH_{CAT} only when required for convergence. . . . As noted, CH_{CAT} should be completely dispensable, were it not for the need to accommodate to the sensorimotor apparatus. (1995a:265)

Note that this seems to assume the second of the two possibilities Chomsky mentioned in the prior passage, that is, that failure of pied-piping causes a violation specifically at PF.²² Chomsky goes on to observe that even overt movement might be possible without pied-piping under certain circumstances, if no phonological requirement is violated.

Just how broadly considerations of PF convergence might extend is unclear, pending better understanding of morphology and the internal structure of phrases. Note that such considerations could permit raising without pied-piping even overtly, depending on morphological structure . . . (1995a:264)

Consider now how the ellipsis phenomena examined above might be reanalyzed in terms of this theory. Recall my analysis of pseudogapping in terms of the PF crash theory of strong features. Assuming the split VP hypothesis, in a nonelliptical transitive sentence, for example, the object raises to [Spec, Agr_O] and the lexical V raises to the higher shell V position in order that a strong feature of the lexical V will be checked. If the V does not raise, a PF crash will ensue, but only if the offending item exists at that level. Under a deletion account of ellipsis, ellipsis provides another way to salvage the derivation. When the lower VP is deleted without the V having raised, a PF crash is avoided and the result is acceptable pseudogapping.

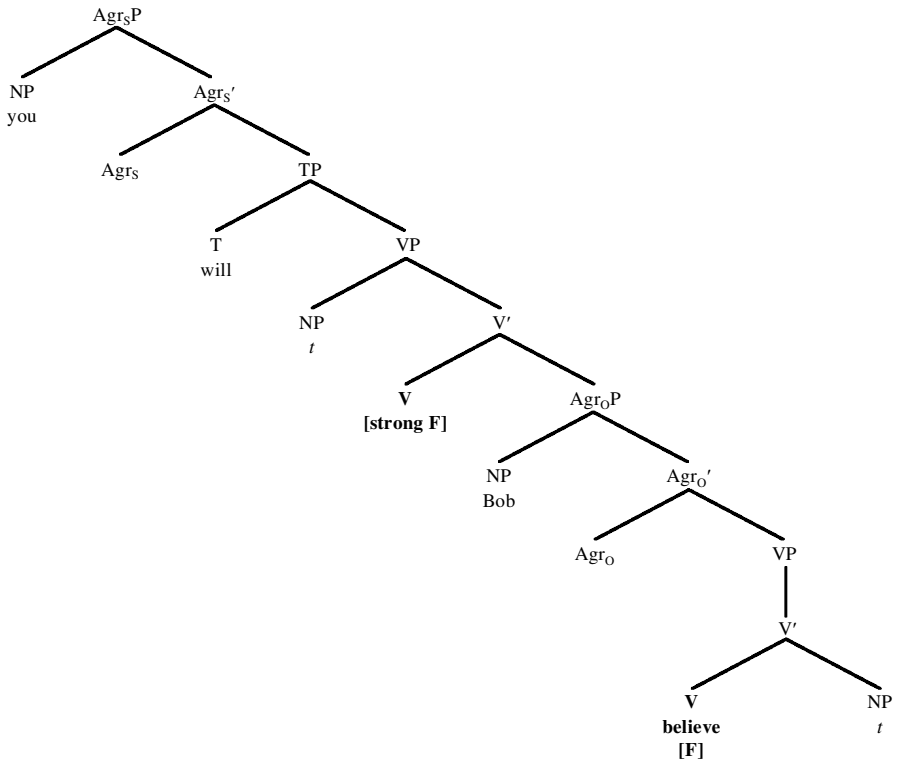
The alternative account preserves the idea of deletion averting a PF crash, but the potential crash now has another cause. The feature driving overt V-raising could be a strong feature of the *higher* V. Once the matching feature of the lower lexical V is ‘‘attracted’’ out of the lower V, the lower V becomes defective. A PF crash will be avoided if either pied-piping or deletion of a category containing the lower V (VP-deletion = pseudogapping in the relevant instances) takes place. This is illustrated in (30).²³ Thus, even under the virus theory there is a way to capture the saving effect of ellipsis in the pseudogapping construction.

Sluicing can be reanalyzed in parallel fashion. Suppose that in accord with the virus theory the strong feature driving overt I-raising in matrix interrogatives resides in C (the usual, and arguably more natural, assumption, at any rate), rather than in I. In a normal matrix interrogative, then, the matching feature of I raises overtly to check the strong feature of C. This leaves behind a phonologically defective I, which will cause a PF crash unless either pied-piping (i.e., overt raising of I) or deletion of a category containing that I (sluicing) takes place. This is illustrated in (31).

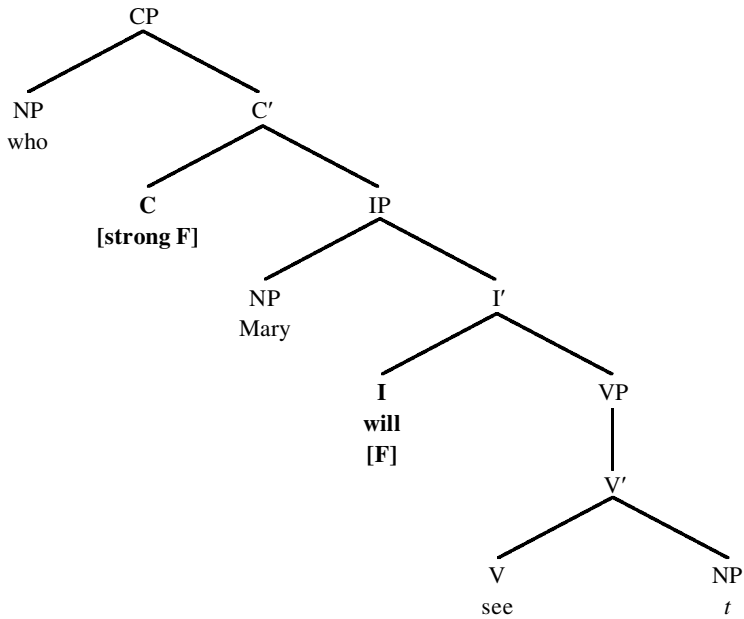
²² As Zeljko Bošković (personal communication) observes, ‘‘globality’’ is thus still present in the pied-piping process, just as it was (implicitly or explicitly) in the PF and LF crash theories.

²³ The entire tree is shown in (30) just for expository purposes. In the actual derivation, the strong feature of the higher V would attract the corresponding feature of *believe* immediately upon the introduction of the former into the phrase marker, in accord with the virus theory.

(30)



(31)



Thus, the essence of the PF account of the ellipsis facts based on the PF crash theory of strong features evidently can be captured under the virus theory as well, a rather surprising result—and perhaps a welcome one if Chomsky’s conceptual arguments for the virus theory are accepted. Before concluding, though, I will briefly mention one last, and rather well-known, argument that strong features reside in some moving categories and that the basic premise of the virus theory is therefore incorrect. There is a great deal of literature, going back to Toman 1982 and Rudin 1982, 1988, discussing the phenomenon of multiple *wh*-movement in the Slavic languages. Bošković (1997) presents a treatment of Serbo-Croatian multiple *wh*-movement in terms directly relevant to the present discussion. He argues that in Serbo-Croatian, *wh*-phrases have a strong focus feature and that that is why they all have to move overtly.

- (32) a. Ko šta gdje kupuje?
 who what where buys
 ‘Who buys what where?’
 b. *Ko kupuje šta gdje?
 c. *Ko šta kupuje gdje?
 d. *Ko gdje kupuje šta?

Whether he is right about the precise identity of the feature will not be of concern here. But the fact that all of the *wh*-phrases must move overtly strongly suggests Bošković’s basic conclusion, that the strong feature driving the movement resides in the *wh*-phrases themselves. If, instead, it resided in the head to which they move, why wouldn’t the movement of just one of them suffice (as in English)?

To summarize, I set out to compare three theories of strong features. Contrary to expectation, existing conceptual arguments based on the computational complexity of look-ahead and on the problem of LF insertion of a head with a strong feature turned out to be inconclusive. An examination of two ellipsis paradigms provided new potential evidence. I suggested a new generalization: that movement *or* ellipsis can rescue a derivation with a strong feature. In light of this, I argued that PF crash is relevant, either directly, as in Chomsky’s PF crash theory of strong features (Chomsky 1993), or indirectly, as in the virus theory augmented by the multiple-chain theory of pied-piping (especially as interpreted by Ochi (1998)). The existence of constructions, such as Slavic multiple interrogation, where multiple items apparently must move to the same position, tips the balance in favor of the PF crash theory, given present (admittedly limited) understanding.

References

- Bošković, Željko. 1997. Superiority and economy of derivation: Multiple *wh*-fronting. Paper presented at the Sixteenth West Coast Conference on Formal Linguistics, University of Washington, Seattle. [To appear in the proceedings.]
- Bošković, Željko. To appear a. Second position cliticization: Syntax and/or phonology. In *Clitics in Europe*, ed. Frits Beukema and Marcel den Dikken. Amsterdam: John Benjamins.
- Bošković, Željko. To appear b. Sometimes in Spec CP, sometimes in situ. In *Step by step: Essays on minimalism in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka. Cambridge, Mass.: MIT Press.

- Campos, Héctor, and Paula Kempchinsky, eds. 1995. *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero*. Washington, D.C.: Georgetown University Press.
- Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In *Principles and parameters in comparative grammar*, ed. Robert Freidin, 417–454. Cambridge, Mass.: MIT Press. [Also published in Chomsky 1995b, 129–166.]
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 1–52. Cambridge, Mass.: MIT Press. [Also published in Chomsky 1995b, 167–217.]
- Chomsky, Noam. 1994. Bare phrase structure. (MIT Occasional Papers in Linguistics 5.) MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass. [Also published in Campos and Kempchinsky 1995, 51–109, and Webelhuth 1995, 383–439.]
- Chomsky, Noam. 1995a. Categories and transformations. In Chomsky 1995b, 219–394.
- Chomsky, Noam. 1995b. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam, and Howard Lasnik. 1993. The theory of principles and parameters. In *Syntax: An international handbook of contemporary research*, vol. 1, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 506–569. Berlin: Walter de Gruyter. [Also published in Chomsky 1995b, 13–127.]
- Chung, Sandra, William Ladusaw, and James McCloskey. 1995. Sluicing and Logical Form. *Natural Language Semantics* 3:1–44.
- Dikken, Marcel den. 1995. Binding, expletives, and levels. *Linguistic Inquiry* 26:347–354.
- Haik, Isabelle. 1987. Bound VPs that need to be. *Linguistics and Philosophy* 10:503–530.
- Hornstein, Norbert. 1994. An argument for minimalism. *Linguistic Inquiry* 25:455–480.
- Jayaseelan, Karattuparambil A. 1990. Incomplete VP deletion and gapping. *Linguistic Analysis* 20:64–81.
- Johnson, Kyle. 1991. Object positions. *Natural Language & Linguistic Theory* 9:577–636.
- Kennedy, Christopher. 1997. Antecedent-contained deletion and the syntax of quantification. *Linguistic Inquiry* 28:662–688.
- Kennedy, Christopher, and Jason Merchant. 1997. Attributive comparatives and bound ellipsis. Ms., Linguistics Research Center, University of California at Santa Cruz.
- Koizumi, Masatoshi. 1993. Object agreement phrases and the split VP hypothesis. In *MIT working papers in linguistics 18: Papers on Case and agreement I*, 99–148. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.
- Koizumi, Masatoshi. 1995. Phrase structure in minimalist syntax. Doctoral dissertation, MIT, Cambridge, Mass.
- Lasnik, Howard. 1972. Analyses of negation in English. Doctoral dissertation, MIT, Cambridge, Mass.
- Lasnik, Howard. 1993. Lectures on minimalist syntax. (University of Connecticut Occasional Papers in Linguistics 1.) University of Connecticut, Storrs. [Also published in Lasnik 1999, 25–73.]
- Lasnik, Howard. 1995a. Last Resort. In *Minimalism and linguistic theory*, ed. Shosuke Haraguchi and Michio Funaki, 1–32. Tokyo: Hituzi Syobo. [Also published in Lasnik 1999, 120–150.]
- Lasnik, Howard. 1995b. Last Resort and Attract F. In *Proceedings of the Sixth Annual Meeting of the Formal Linguistics Society of Mid-America*, 62–81. Indiana University Linguistics Club, Bloomington.
- Lasnik, Howard. 1995c. A note on pseudogapping. In *MIT working papers in linguistics 27: Papers on minimalist syntax*, 143–163. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass. [Also published in Lasnik 1999, 151–174.]
- Lasnik, Howard. In press. Pseudogapping puzzles. In *Fragments: Studies in ellipsis*, ed. Elabbas Benmamoun, Hiroto Hoshi, and Shalom Lappin. Oxford: Oxford University Press.
- Lasnik, Howard. 1999. *Minimalist analysis*. Oxford: Blackwell.
- Lasnik, Howard, and Mamoru Saito. 1991. On the subject of infinitives. In *CLS 27. Part I, The General Session*, 324–343. Chicago Linguistic Society, University of Chicago, Chicago, Ill. [Also published in Lasnik 1999, 7–24.]

- Levin, Nancy. 1978. Some identity-of-sense deletions puzzle me. Do they you. In *Papers from the Fourteenth Regional Meeting, Chicago Linguistic Society*, 229–240. Chicago Linguistic Society, University of Chicago, Chicago, Ill.
- Levin, Nancy. 1979. Main verb ellipsis in spoken English. Doctoral dissertation, Ohio State University, Columbus. [Published 1986 by Garland, New York.]
- Lobeck, Anne. 1990. Functional heads as proper governors. In *Proceedings of NELS 20*, 348–362. GLSA, University of Massachusetts, Amherst.
- May, Robert. 1985. *Logical Form: Its structure and derivation*. Cambridge, Mass.: MIT Press.
- Nemoto, Naoko. 1993. Chains and Case positions: A study from scrambling in Japanese. Doctoral dissertation, University of Connecticut, Storrs.
- Ochi, Masao. 1998. Move or Attract? In *Proceedings of the Sixteenth West Coast Conference on Formal Linguistics*. Stanford, Calif.: CSLI Publications. [Distributed by Cambridge University Press.]
- Postal, Paul M. 1974. *On raising: One rule of English grammar and its theoretical implications*. Cambridge, Mass.: MIT Press.
- Romero, Maribel. 1997. Recoverability conditions for sluicing. In *Empirical issues in formal syntax and semantics*, ed. Francis Corblin et al. Bern: Peter Lang.
- Ross, John Robert. 1969. Guess who? In *Papers from the Fifth Regional Meeting, Chicago Linguistic Society*, 252–286. Chicago Linguistic Society, University of Chicago, Chicago, Ill.
- Rudin, Catherine. 1982. Who what to whom said: An argument from Bulgarian against successive cyclicity. In *Papers from the Seventeenth Regional Meeting, Chicago Linguistic Society*, 353–360. Chicago Linguistic Society, University of Chicago, Chicago, Ill.
- Rudin, Catherine. 1988. On multiple questions and multiple *wh*-fronting. *Natural Language & Linguistic Theory* 6:445–501.
- Sag, Ivan. 1976. Deletion and Logical Form. Doctoral dissertation, MIT, Cambridge, Mass.
- Saito, Mamoru, and Keiko Murasugi. 1990. N'-deletion in Japanese. In *University of Connecticut working papers in linguistics* 3, 87–107. Department of Linguistics, University of Connecticut, Storrs.
- Sobin, Nicholas. 1997. Agreement, default rules, and grammatical viruses. *Linguistic Inquiry* 28:318–343.
- Stump, Gregory. 1977. Pseudogapping. Ms., Ohio State University, Columbus.
- Tancredi, Chris. 1992. Deletion, deaccenting, and presupposition. Doctoral dissertation, MIT, Cambridge, Mass.
- Toman, Jindřich. 1982. Aspects of multiple *wh*-movement in Polish and Czech. In *Levels of syntactic representation*, ed. Robert May and Jan Koster, 293–302. Dordrecht: Foris.
- Vanden Wyngaerd, Guido. 1989. Object shift as an A-movement rule. In *MIT working papers in linguistics 11: Student Conference in Linguistics*, 256–271. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, Mass.
- Wasow, Thomas. 1972. Anaphoric relations in English. Doctoral dissertation, MIT, Cambridge, Mass. [Revised version published as Wasow 1979.]
- Wasow, Thomas. 1979. *Anaphora in generative grammar*. Ghent: E. Story-Scientia.
- Webelhuth, Gert, ed. 1995. *Government and Binding Theory and the Minimalist Program*. Oxford: Blackwell.

Department of Linguistics
University of Connecticut
Storrs, Connecticut 06269–1145
 lasnik@uconnvm.uconn.edu

Copyright of Linguistic Inquiry is the property of MIT Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.