

# The syntax of extraction: derivation or constraint satisfaction?

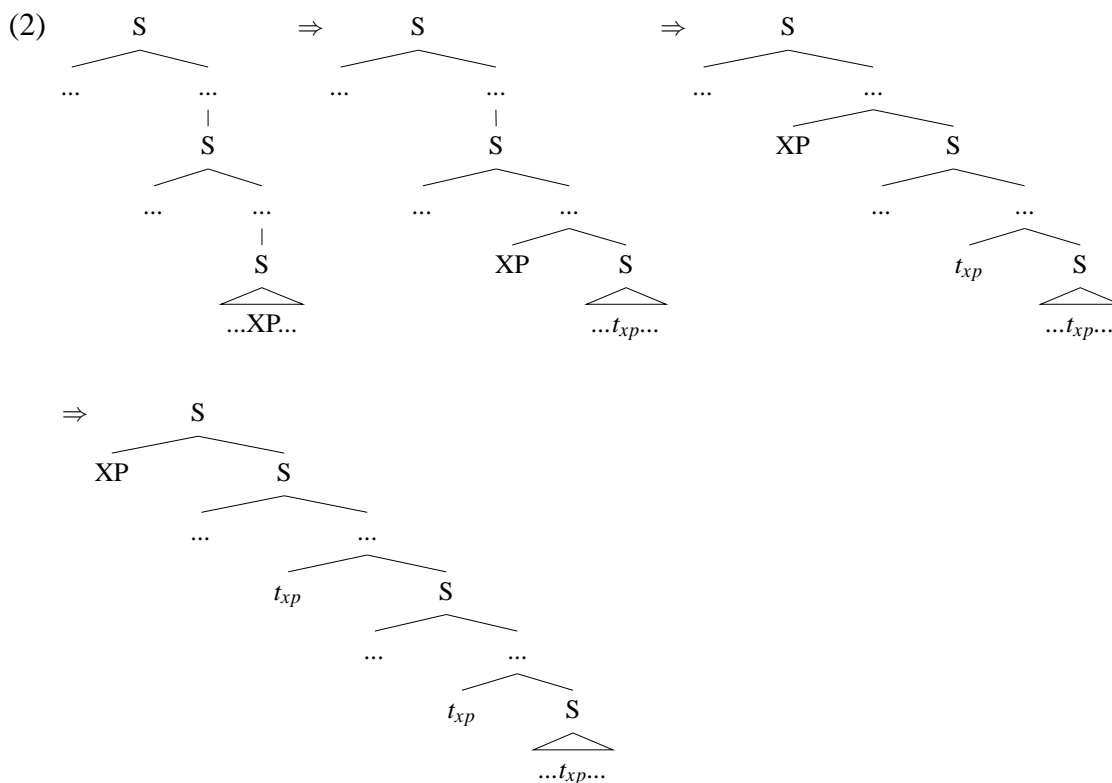
Robert Levine

## 1 Filler/gap constructions: two approaches

Historically, filler/gap constructions such as those in (1) have been approached two ways:<sup>1</sup>

- (1) a. THAT book, you should purchase \_\_\_.
- b. Which book does Leslie think you should purchase \_\_\_?
- c. This is the book which Leslie told me she thinks I should purchase \_\_\_.

Transformational approaches posit a sequence of representations in which the filler is initially in the position notated by the underline in (1), which is then relocated, possibly via a series of movement steps, to its final position on the left of the highest clause. Schematically, the derivational approach can be illustrated in (2):



<sup>1</sup>The material presented in this paper partially overlaps with the content of Levine and Sag (2003).

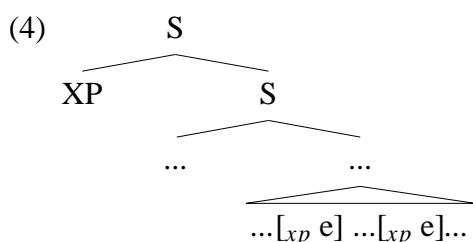


only are there clear framework-architectural differences between the approaches, but that the radical inadequacy of derivational approach to extraction on purely factual ground makes the monostratal position—which as I'll argue is the logically strongest position one can take on—distinctly preferable on uncontroversial methodological grounds.

The problem with framework comparison taking extraction phenomena as the basis is that *single filler/gap* linkages do not sharply distinguish between the predictions of derivational theories in which a constituent changes its structural position in the course of a derivation vs. those of monostratal theories in which nothing moves, but where instead a small set of very general constraints have to be satisfied at all structural positions. To drive a wedge between these two fundamentally different worldviews in an *empirically* pointed way, it is necessary to examine their relative utility in accounting for constructions in which a single filler is linked to two or more gap sites.

## 2 What multiple gap constructions tell us

The first point is straightforward: in a single filler/multiple gap construction, such as the parasitic gap phenomenon, the finale of the derivational picture looks not like that in (2), but rather like (4):



What is the relationship between the filler and the two gaps? There is no well-defined formal operation corresponding to movement of two distinct daughter constituents to a single phrase structure position, as emphasized by Gazdar et al. (1982). That is, a single linkage mechanism to the two gap sites is in principle unavailable given movement *simpliciter*. Only two possible choices are available:

- there is one kind of linkage mechanism between the filler and one of the gaps and a different kind of linkage mechanism between the filler and the other gap; or
- there is only a single kind of linkage mechanism available between fillers and gaps, and in multiple gap construction there are two separate instances of the same mechanism.

In the first case, there is an obvious asymmetry: one of the gaps must represent a trace of the filler, so that the other position must be occupied by a phonologically null something which is not a trace. In the second case, movement is the sole linkage mechanism in both cases, which entails that there is, in addition to the movement bringing the overt filler to its surface position, a second movement leaving the second trace—with the second moved element then necessarily, invisible. Here the asymmetry is between the movement chain linking the overt filler to the gap site, on the one hand, and that linking the null filler to the gap site, on the other.

Both variants, and various hybrids, exist in the literature; sticking to very familiar examples, Chomsky (1982) manifests the first alternative and Contreras (1984), later adopted in Chomsky (1986) the second, which has become known, in spite of the fact that Contreras first proposed it, as the *Barriers* approach. But the plausibility of such approaches is only as strong as the arguments for the asymmetry assumed. The following are the principle ones I am aware of:

- the so-called Kearney paradigm;
- the distribution of finite clause subject gaps;
- the supposed asymmetry of weak crossover;
- supposed (anti)pronominality effects and alleged categorial restrictions on parasitic extraction.

Each of these supposed phenomena has been used to motivate a radical dichotomy among filler/gap chains in at least certain multiple-gap constructions. Frampton (1990) cites this data in arguing for a distinction between true gaps and p-gaps, and Kiss (2001) repeats this evidence apparently under the impression that it constitutes support for an asymmetrical analysis of p-gap constructions.

But the data supporting this dichotomy turns out in every case to be illusory, in many cases straightforwardly clearly derivable from processing constraints that are, in the relevant respects, configurationally blind. When these are sorted out, none of the supposed evidence for chain asymmetry in multiple gap constructions turns out to be relevant.

There is, in fact, positive evidence *against* this asymmetrical characterization of multiple gap chains. I will argue later that in addition to parasitic gap constructions there is

- a *symbiotic* multiple gap construction which is intractable on derivational accounts predicated on an asymmetry between filler/gap linkages, and
- a pattern of apparent case-inconsistency between the gaps in multiple gap constructions which, though at first glance compatible with a Contreras/*Barriers*-style analysis of asymmetrical chaining, turns out to be badly mispredicted by such approaches.

## 2.1 The Kearney paradigm

The primary argument in the literature for chain asymmetry in p-gap constructions is, as far as I am aware, given in Chomsky (1986). Chomsky cites the following two examples, due to Kearney (1983):

- (5) a. Which books about himself did John file *t* [before Mary read *e*]?  
 b. \*Which books about herself did John file *t* [before Mary read *e*]?

Chomsky observes that

Example [(5)a] is a normal parasitic gap construction, but [(5)b] is ungrammatical. It follows, then, that the *wh*-phrase in [(5)a], [(5)b] is extracted from the position of *t*, not from the position of the parasitic gap *e*. As Taraldsen had originally assumed, the latter is truly ‘parasitic’.

While hardly transparent, the reasoning seems to be this: if p-gap constructions were instances of multiple, i.e., symmetrical, gap phenomena, reconstruction of the filler should proceed symmetrically to yield identical effects in both (5)a and b. In both cases, the result would be a structure with a reconstructed filler compatible with one antecedent *but not the other*, and so both should be bad. What we instead find is that the reconstruction is good when the anaphor is compatible with the subject of the clause containing the ‘true’ gap, but not otherwise. The simplest conclusion, the reasoning seems to be, is that the overt filler reconstructs only to the main clause gap site, which must then be its transformational point of origin.

But this frequently accepted conclusion, echoed for example in Frampton (1990) and Kiss (2001), is demonstrably incorrect. Consider first the contrast:

- (6) a. Which pictures of himself did John go to England without telling Mary to send *t* to the INS?  
 b. ??\*Which pictures of herself did John go to England without telling Mary to send *t* to the INS?

The pattern and the quality of the contrast here is exactly parallel to that in (5), but there is only one gap site, hence necessarily the ‘true’ gap site in both cases, and hence nothing remotely like the explanatory line Chomsky, Frampton, Kiss and others assume respectively for (5) here can apply in the case of (6). There is no structural difference whatever between the two examples in (6). The sole difference is that in (6), an incompatible potential antecedent is linearly closest to the reflexive within the filler.

This observation suggests that we might well be able to ameliorate the ill-formedness in (5) were we able to rearrange the structures involved somewhat so that the antecedent in the ill-formed case in (5)b were closer to the filler than the unmatchable potential antecedent *John*. And this is possible: as we have learned from Haegeman (1984), parasitic gap construction of the form VP/[ $\square$ ] XP/[ $\square$ ] have well attested analogues XP/[ $\square$ ] VP/[ $\square$ ]. Consider the result of replacing in (9) the right-adjunct version of the p-gap with the left-adjunct version:

- (7) a. There were pictures of  $\left\{ \begin{array}{l} \text{himself} \\ * \text{herself} \end{array} \right\}$  which John put into circulation once Mary approved of.  
 b. There were pictures of herself which once Mary finally decided she liked/approved of, John put into circulation.

And for completeness,

- (8) a. There are pictures of himself which John wants to put into immediate circulation, though they’ll take Mary a while to get used to.  
 b. There are pictures of herself which, though they’ll take Mary a while to get used to, John wants to put into immediate circulation.

The goodness of (7)b and (8)b immediately refutes Chomsky’s (rather implicit) account of the Kearney paradigm in (5). For when the parasitic-gap adjunct host precedes the main VP, the Kearney effect disappears completely. On Chomsky’s account, this would have to mean that for structurally unmotivated reasons, reconstruction into the parasitic gap was suddenly possible, in spite of the supposed fact that the overt filler has no direct syntactic relationship with this gap site. And if such reconstruction were possible, then what could possibly block it in the case of (5)?

The conclusion we come to then is that the Kearney paradigm has been badly misunderstood ever since its first introduction into the literature as a justification for the posited asymmetry of p-gap constructions, and in fact is at best irrelevant to the question.

## 2.2 Nominative subject p-gaps

A second argument for chain asymmetry is given in Chomsky (1982), Cinque (1990), Frampton (1990) and Postal (1998), based on the supposed ill-formedness of parasitic gaps in finite subject positions. Examples such as those in (9) are often offered as illustrations of this claim:

- (9) a. \*Jack, who<sub>i</sub> I heard about \_\_<sub>i</sub> before you said \_\_<sub>i</sub> would hire us... (Frampton (1990), p.68.)  
 b. \*Someone who<sub>i</sub> John expected \_\_<sub>i</sub> would be successful though believing \_\_<sub>i</sub> is incompetent...(Chomsky (1982), p.55)  
 c. \*The militant who they arrested \_\_<sub>i</sub> after learning \_\_<sub>i</sub> was \_\_<sub>i</sub> carrying a gun...

Since true gaps have no problem extracting from finite subject position, such examples, taken to be representative, have been important supporting evidence for the position that parasitic gaps really involve a different relation to overt fillers than true gaps do. But again, examination of a slightly wider range of data shows that whatever difficulty such examples pose for acceptability, they are very far from being representative of the general case. Consider examples (10):

- (10) a. [Which people]<sub>i</sub> did you invite \_\_<sub>i</sub> without thinking \_\_<sub>i</sub> would actually come?  
 b. Jack, who<sub>i</sub> even before you said \_\_<sub>i</sub> would hire us I was favorably disposed towards \_\_<sub>i</sub> , is a prince among men.  
 c. There go [the Endaby twins]<sub>i</sub> , who as soon as I realized \_\_<sub>i</sub> were on their way over to visit me I made immediate arrangements to avoid \_\_<sub>i</sub>.

There are so many good examples of such p-gaps that the claim that they are in general bad seems without any solid foundations.

### 2.3 Weak crossover

The next set of claims I want to address posits a distinction between true and p-gaps (and, in parallel fashion, between the first gap in an ATB extraction and a following gap as per Munn (2001)) based on the claim that there is a class of extractions which are not susceptible to weak crossover (WCO) effects, and that these include parasitic and non-leftmost coordinate structure gaps. The basic argument here is presented in Lasnik and Stowell (1991). They note the contrast between examples such as (11)a and b:

- (11) a. \*Who<sub>i</sub> did his<sub>i</sub> lawyers make a convincing case for t<sub>i</sub>?  
 b. Who<sub>i</sub> did the jury acquit after his<sub>i</sub> lawyers made a convincing case for t<sub>i</sub>?

Lasnik and Stowell argue that p-gaps line up with missing object constructions, clefts, non-restrictive relatives and other instances of what Postal has referred to as ‘B-extractions’, which he and Cinque, offering somewhat different variants of the analysis, have argued should be analyzed as instances of null resumptive pronouns. I don’t want to say too much about their argumentation, which consists of taking a reasonable generalization about extracted operators, generalizing it against all factual evidence to all  $\bar{A}$  fillers, then exempting the counterfactual cases by arguing that the gaps are not traces and hence their proposed WCO condition will not apply to them—an argument roughly like claiming that the observation that all primes greater than 2 are odd is too weak, that *all* prime numbers must be odd, and then arguing that the claim is justified if we decide that 2 is not a number. But the crucial point is that Lasnik and Stowell’s position—which is a centrepiece of Munn’s (Munn (2001)) analysis of multiple-gap constructions—turns out to be empirically completely unfounded. Parasitic gaps indeed manifest WCO effects, if the p-gap appears in a place where—to summarize altogether insufficiently—there is insufficient material introduced into the discourse to establish a credible antecedent for the pronoun. There are two such classes of p-gaps:

- subject p-gaps;
- left-fronted adjunct p-gaps

Consider the first kind of case, exemplified in (12):

(12) His<sub>i</sub> fan's ideas about Robin<sub>i</sub>'s work materially **improved** Robin<sub>i</sub>'s work.

As we would expect, such cases of cataphora are quite legal under all versions of binding theory. The pragmatics of ordinary discourse seem to require that a repetition of the NP *Robin's work* involve some kind of appropriate contrastive or emphatic stress, but this effect seems irrelevant to the syntax itself. Suppose now we attempt to form a subject p-gap on the basis of (12):

(13) ??\*[Whose<sub>i</sub> work]<sub>j</sub> did his<sub>i</sub> fan's ideas about e<sub>j</sub> materially improve t<sub>j</sub> ?

This sounds pretty dreadful—a classic, typical WCO effect, in fact. And as proof of this, note that we can ameliorate the effect by using the standard technique of increasing the denotational specificity of the extracted element, a point noted in Wasow's groundbreaking work on the WCO effect in the early 1970s:

(14) [[Which reknowned master of fiction]<sub>i</sub>'s work]<sub>j</sub> did his<sub>i</sub> fan's ideas about e<sub>j</sub> materially improve t<sub>j</sub> ?

Similarly, the other well-established technique of focalizing the pronoun works in these cases exactly as in standard WCO examples:

- (15) a. Who<sub>i</sub> did \*(even) his<sub>i</sub> mother complain about?  
 b. Who<sub>i</sub> did his<sub>i</sub> \*(own) mother complain about?  
 c. [Whose<sub>i</sub> work]<sub>j</sub> did even his<sub>i</sub> own fan's ideas about e<sub>j</sub> fail to materially improve t<sub>j</sub> ?

Next, consider the Haegeman variant of adjunct p-gaps in connection with the supposed WCO-immunity of p-gaps:

- (16) a. I've found that what<sub>i</sub> I can't understand t<sub>i</sub> until its<sub>i</sub> author explains e<sub>i</sub> to me is basically anything in post-WWII literature.  
 b. \*I've found that what<sub>i</sub> until its<sub>i</sub> author explains e<sub>i</sub> to me I can't understand t<sub>i</sub> is basically anything in post-WWII literature.  
 c. I've found that what<sub>i</sub> until my agent explains e<sub>i</sub> to me I can't understand t<sub>i</sub> is basically anything in post-WWII literature.

These examples are *exactly* what we predict on the assumption that p-gaps are subject to precisely the same WCO effects as any other kind of 'true' gap. They show conclusively that there is no chain asymmetry between parasitic gaps and the gaps that these are putatively parasitic on.<sup>2</sup>

<sup>2</sup>In any case, as exhaustively documented in Postal's unaccountably overlooked (1993b) discussion of weak(est) crossover, Lasnik & Stowell's claims about topicalization and other 'weakest' cases are factually untenable: 'in certain circumstances, extractions under at least topicalization, clefting and nonrestrictive relative clause formation do yield WCO effects, even when... the extracted phrases are not in any sense characterizable as "true quantifier phrases"'. (p. 546). Munn (2001) appears to wish to retain Lasnik & Stowell's argument at least insofar as it putatively applies to multiple gap constructions; he cites Postal's WCO paper approvingly, suggesting that he accepts Postal's demonstration of Lasnik and Stowell's empirical failings so far as single-gap constructions are concerned, but thinks that it still holds for multiple-gap constructions. The data cited above of course show that Lasnik & Stowell's claims about multiple gap constructions are as unsupported as those pertaining to the single-gap cases implicated in Postal's extensive counterevidence.

## 2.4 Alleged antipronominality

Finally, it has been claimed in Cinque (1990), Postal (1998) and Munn (2001) that p-gaps (along with other supposed B-extractions) do not tolerate extractions from sites that are resistant to the appearance of pronominal forms. This claim is extremely easy to falsify; so consider, e.g.,

(17) a. I'm a(n)  $\left\{ \begin{array}{l} \text{friend} \\ \text{brother} \\ \text{ally} \end{array} \right\}$  of Thorkill Skullsplitter, but Terry isn't a(n)  $\left\{ \begin{array}{l} \text{friend} \\ \text{brother} \\ \text{ally} \end{array} \right\}$  of his/\*him

a. There are [certain people]<sub>i</sub> you can't do business with   <sub>i</sub> unless you're a  $\left\{ \begin{array}{l} \text{friend} \\ \text{relative} \\ \text{ally} \end{array} \right\}$  of   <sub>i</sub>.

(18) \*There are certain people whose you can't do business with unless you're a relative of.

Many other examples are given in Levine et al. (2000) and Levine (2001). It seems very difficult to maintain the claim in question unless one appeals, as in Postal (1998), to the possibility that such environments rule out overt weak definite pronouns but allow covert instances of such pronouns—a position that Postal himself acknowledges has no independent support and amounts essentially to a diacritic invoked to neutralize real counterexamples.

## 3 Evidence against chain asymmetry

### 3.1 Symbiotic gaps

The foregoing discussion establishes the essentially negative point that the chief published arguments for chain asymmetry in derivational theories of p-gap licensing are entirely spurious. We now move to positive evidence that that chain-asymmetric approaches to multiple gap constructions are profoundly misconceived. Consider the data in (19), where *both* gaps seem to be within islands:

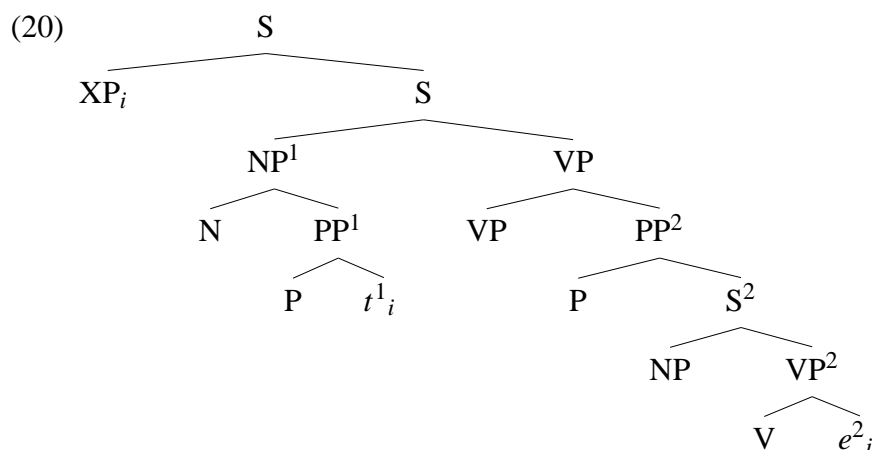
- (19) a. What kinds of books do authors of    argue about royalties after writing   ?
- b. ??What kinds of books do authors of malicious pamphlets argue about royalties after writing   ?
- c. \*What kinds of books do authors of    argue about royalties after writing malicious pamphlets?

If either gap is a 'true' gap, then the argument for chain asymmetry essentially disappears in the case of subject-gap/main VP gap or main VP gap/adjunct gap p-gap constructions—in which case multiple-chain analyses such as the *Barriers* analysis make no sense. The only defensible position seems to be to assume that subject and adjunct gap are *mutually* parasitic, or as I shall call them, symbiotic, i.e., depend on *each other* for licensing.

Can such constructions actually be licensed by movement approaches? The short answer is no. In particular:



- Under Chomsky's 1982 approach in *Concepts and Consequences* (and also Cinque (1990)), a parasitic gap starts out in DS as *pro*, and is subsequently coindexed with the filler linked to the 'true' gap site'; otherwise identification of *pro* is impossible (or the functionally determined equivalent reasoning). Island conditions apply to all variables, regardless of how they arise. But both gap sites are islands. Hence there is no legal extraction to establish a filler that can license the other gap.
- On Kayne's 1983 'connectedness' approach, a free gap can only establish a connection to a parasitic gap if the path from the parasitic gap to the true gap can be continuously xmediated in terms of what Kayne calls the g-projection path. Longobardi noted that in order to yield the correct results for p-gap constructions, it was necessary to ensure that each node in the projection path be properly governed. But on this assumption, it turns out however that the g-projections of the subject gap and the adjunct gap both terminate before a connected path can be established, leaving the legal examples in (19) presumably unlicensed, as charted in (20), where superscripts indicate g-projections.



- Chomsky's 1986 *Barriers* approach does not actually contain a particularly satisfactory account of gap parasitism, since in the end it posits a technical device—Chain Composition constrained by 0-subjacency—which is never reconciled with the existence of subject islandhood and subject parasitic gaps. It seems fairly clear from Chomsky's discussion in §7 that he regards IP as directly dominating the adjunct, in spite of the fact that examples such as *And go to England without signing these papers I will* make it clear that the adjunct is part of the VP. To ensure the parasitism of parasitic gaps, it is necessary on the *Barriers* analysis that the adjunct and the subject not only function as barriers themselves but, as blocking categories, ensure that the dominating maximal projections closest to them—which in the case of the subject is clearly IP but for the adjunct must be VP—are barriers for any movement originating within the subject and the PP respectively.

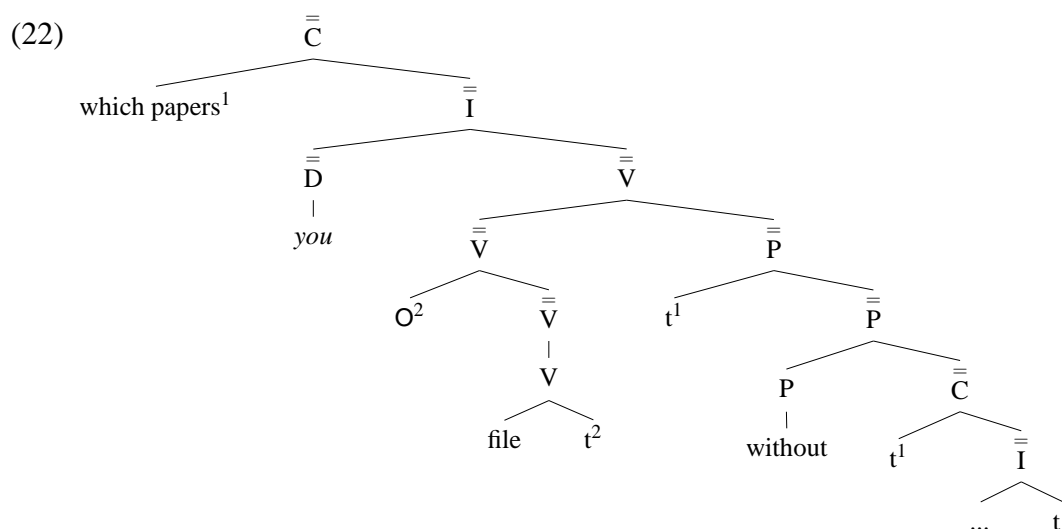
For symbiotic gaps there are only two possibilities: the *wh* phrase originates in the subject or in the adjunct island. If the subject is the source, then the stipulated prohibition on adjunction to DP entails a direct movement through IP (which itself also prohibits adjunction). But a movement from a barrier immediately under IP entails that IP counts as a barrier by inheritance for that movement. The result is of course a decisive subjacency violation ruling out (unsupported) subject extraction. But, while PPs are not L-marked and therefore count as barriers, adjunction to PPs is actually admitted in Chomsky's analysis (see *Barriers*, pp. 65–66). What then

blocks *wh* movement from the adjunct itself, and empty operator adjunction within the main VP? Clearly, Chomsky's discussion of the Kearney paradigm makes it clear that he does not envisage this possibility.

In fact, this possibility is clearly ruled out for parasitic gaps under Chomsky's notion of Chain Composition. Chain composition requires that the head of the parasitic chain be 0-subjacent to the lowest element in the true-gap chain. Consider (21):

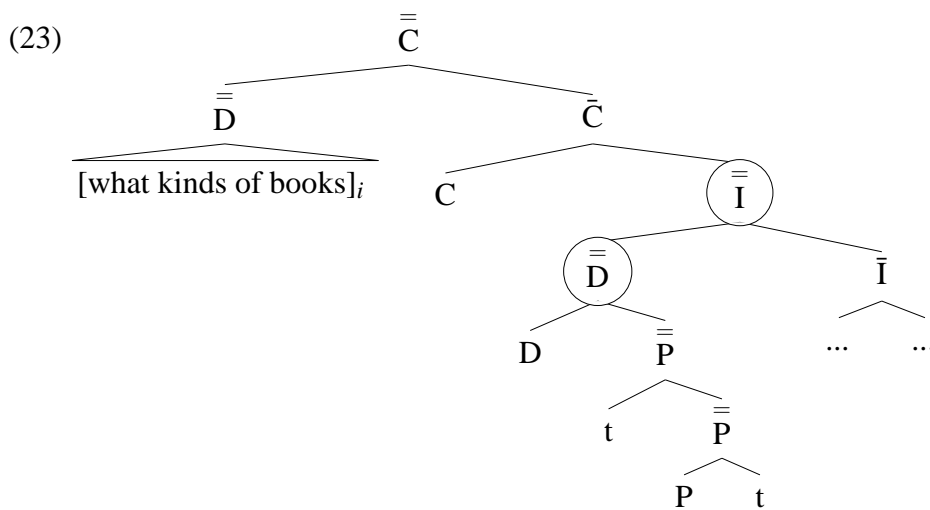
(21) Which papers did you file \_\_without reading \_\_?

If *wh* percolated from the *adjunct* clause to the highest [Spec,CP], and the empty operator main verb object adjoined to VP (from which it could go no higher) then we would have the situation depicted in (22), in which Chain Composition could not occur:



The head of the parasitic chain, the empty operator, is separated from the bottom trace of the true chain by the barrier PP. Hence the conditions on Chain Composition cannot be met.

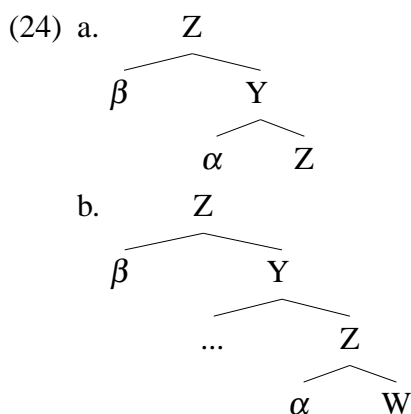
In exactly the same way, if in the case of (19)a we attempted to use the adjunction escape hatch to allow *wh* to move from the adjunct phrase to the highest Spec position, we would have the null operator phrase within the subject NP, unable to escape any further. This operator would be separated from anything outside the subject by the NP barrier, and still further by the main VP, adjunct PP and adjunct CP barriers from the tail of the true chain. Hence Chain Composition would be ruled out on this scenario. Nor could *wh* originate in the subject DP, since it would have to pass through two barriers on the way to matrix Spec, as shown in (23):



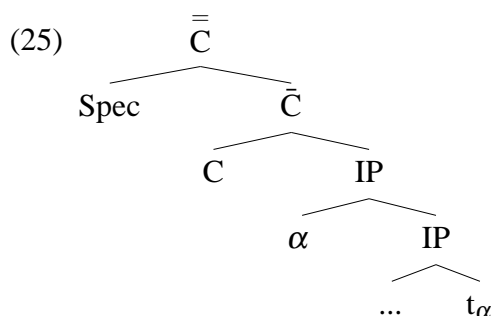
Therefore the approach in Chomsky (1986a) mispredicts that examples such as (19) are ill-formed.

Nonetheless, Chomsky's *Barriers* proposals for gap parasitism are seriously compromised by the fact that 0-subjacency is impossible for subject parasitic gap Chain Composition unless, as Chomsky himself observes, 'the empty operator [moves] out of subject position (which is a barrier) at S-structure, where Chain composition is licensed' (p.66). But where will it go? By stipulation it cannot adjoin to NP, which it must to render IP a non-barrier for Chain Composition. Chomsky never actually provides a solution, merely noting vaguely that 'several possibilities might be pursued, but the question remains obscure' (p.66). One of the motivations of Frampton's treatment of p-gap phenomena was clearly to propose a *Barriers*-framework alternative to Chomsky's technically highly problematic treatment of gap parasitism.

- Frampton's (1990) treatment of parasitic gaps, a kind of hybrid of Kayne's connectedness with Chomsky's null operator treatment in *Barriers* incorporating his *wh*-deletion analysis in 'On *wh* movement' (Chomsky (1977)), is in effect a derivational reconstruction of the multiple licensing of extractions pathways linked to a single filler. Frampton's strategy is to dissociate the formation of chains from the history of movement, and to allow *wh* NPs to spontaneously delete, leaving, in a manner never made particularly explicit, a trace behind which can then be gathered into a chain, as he puts it, by an SS-level operation which requires subjacency between all links. Adjunction is excluded to the 'potential argument categories' DP, PP and CP. In Frampton's system,  $\alpha$  is subjacent to  $\beta$  in (24)a but not (24)b:



where both Y and Z are barriers. Barrierhood is reserved either for non-L-marked categories or those which inherit barrierhood from blocking categories below them, along familiar lines. Frampton however allows adjunction to IP, yielding the configuration



Even though CP inherits barrierhood from IP, it will not be a barrier for a trace which is not properly dominated by the IP it inherits barrierhood from. Hence, if the CP in (25) is L-marked, it cannot be a barrier. However, adjunction is restricted by the crucial condition that it is possible only to a category whose head canonically governs the moved constituent. Thus, in (25),  $\alpha$  cannot be the DS subject of IP.

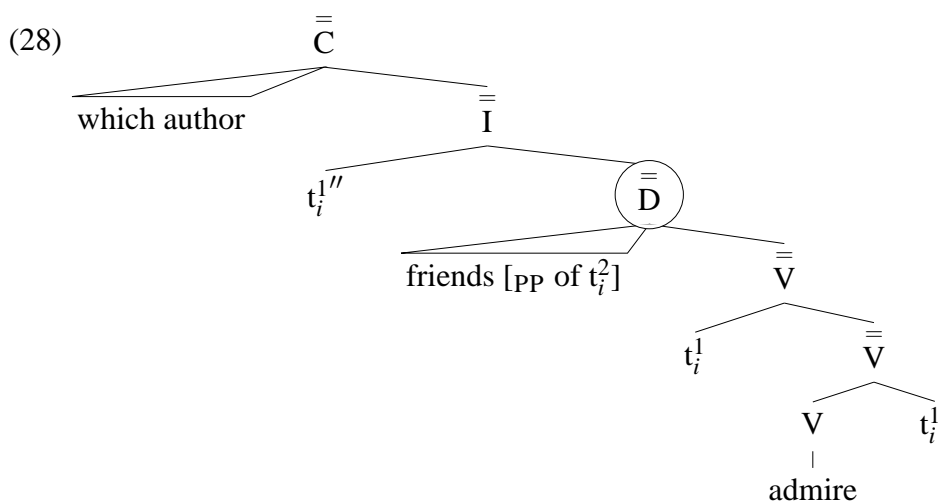
With these components in place, Frampton attempts to explain the pattern of judgments in (26)-(27):

- (26) a. \*Who do friends of hate Tolstoy?  
 b. \*Who did you praise Bill because you like?

- (27) a. Which author do friends of admire?  
 b. Who did you praise because you like?

In (26)a, *who* may not adjoin to NP, which is forbidden in principle, or to IP, since the subject is not canonically governed by Infl. Hence it must move directly to [Spec,CP]. But since both the non-L-marked subject NP and the barrier-by-inheritance CP intervene, subjacency is violated in this movement and the example is ill-formed. In (26)b, *who* can legally move as far as the [Spec,CP] of the clause. Since PP is not L-marked, it constitutes a barrier, to which adjunction is moreover ruled out by fiat. Hence the head of the VP containing the PP does not canonically govern *who* in its highest legal position within PP, and therefore *who* cannot adjoin to VP, which then constitutes a second barrier that may not be crossed without violating subjacency.

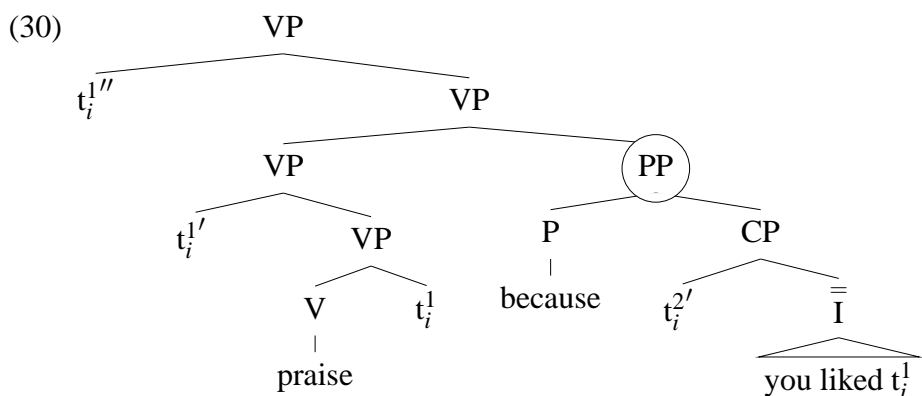
In order to motivate (27), Frampton no longer has to rely on the construction-specific stipulation of Chain Composition which Chomsky introduced in *Barriers*, and which, as noted, still faces the technical impasse of subject gap parasitism. In (27)a, the *wh* phrase object of *admire* legally adjoins to VP, then to IP and moves into [Spec,CP]. The subject of *admire* remains a barrier as in (26)a. But the prepositional object within this subject need not, in this case, move itself. It can spontaneously delete, i.e. become a trace, leaving us with the configuration in (28):



The two *wh* expressions are freely coindexed in DS. Superscripts identify the original ‘history of derivations’ for the two separate *wh* expressions. We have, according to Frampton’s constraints on chain formation, two legal chains

- (29) a. (*which author*,  $t_i^{1''}$ ,  $t_i^{1'}$ ,  $t_i^1$ )  
 b. (*which author*,  $t_i^{1''}$ ,  $t_i^2$ )

Crucially, there is a legal chain which includes the highest token of  $t^1$  and the sole element of the second ‘history of movement’  $t^2$ . Since the nominal head of the subject L-marks the *of* PP, the latter is not a barrier, and hence there is only a single DP barrier intervening between  $t^1$  and  $t^2$ . Thus the chain (29)b is legal. A similar derivation will license (27)b. Here there will be a legal chain reflecting a history of movement from the object of *to* to the *wh* phrase in [Spec,CP], and a second history of movement whose upper termination is in the Spec of the adjunct CP *you like*. If the *wh* phrase in the latter position morphs into a trace, we have for the critical part of the structure the tree in (30), with, again, the sole barrier circled:



Adjunction to PP is ruled out. Here again, however, there is a chain which can be formed linking  $t_i^{1''}$  to  $t_i^2$  legally, with only a single intervening PP barrier. This chain is parasitic on the ‘history of movement’ linking the *wh* filler to its DS site as the object of *like*.

And, as we might expect, the same problem with connectedness in these cases carries over to Frampton’s trace-based analogue. It is obvious that when one gap is within a subject DP and

the other within an adjunct, there is no ‘history of movement’ which can provide the materials for a chain that will include *either* gap. For the very reasons that Frampton has been at pains to explain, neither of the coindexed *wh* phrases that would have to be assumed in DS in (19)a can reach Spec of the matrix CP. As already explained in connection with (26)a and b., they will be prevented by subjacency from extracting from DP and adjunct PP respectively: adjunction to the PP subject by an internal *wh* phrase is ruled out in principle, and so is adjunction to PP, leaving neither barrier-internal operator able to establish a history of movement which can be utilized to license a gap corresponding to ‘spontaneous deletion’ of the other. Thus Frampton’s account falsely predicts that gap symbiosis does not occur in English syntax.

The upshot of all this is that no reasonably explicit P&P theory of p-gaps has anything that looks like even the beginning of an account of symbiotic gaps.

### 3.2 Case conflict and its resolution

Finally, consider examples such as (31).

- (31) Robin is someone who<sub>i</sub> even good friends of \_\_<sub>j</sub> believe \_\_<sub>j</sub> likes power entirely too much.

The filler here is linked to two gap sites, an accusative prepositional object and a nominative finite clause subject. Such mismatches seem to support the position that there is an asymmetry between the two chains that p-gap constructions comprise: if both gaps were linked to a single filler in precisely the same way, the latter would have to share case specifications with both gap sites. In contrast, a double chain analysis, for example, along *Barriers* lines, seems to fit the bill: there will be literal connectivity only along the true filler/gap pathway, while the null operator is linked to the true filler/gap pathway only anaphorically, sharing indices but no  $\phi$  features, so that we would have the situation in (32)

- (32) *wh*<sub>i</sub> [Nom]... O<sub>i</sub> [Acc]...t<sub>i</sub> [Acc]...t<sub>i</sub> [Nom]

So the possibility of case mismatches seems to be predicted. Surely this is a plus for the asymmetrical chain analysis?

In this case appearances are particularly deceiving, for it turns out that none of the movement approaches considered has a straightforward way of accounting for the fact that *such mismatches will occur only when the overt filler is morphologically neutral with respect to case marking*. On the *Barriers* approach, the true and parasitic gap are supposed to be case-independent of each other. So then why then do we have the following data?

- (33) a. \*Him<sub>i</sub> , even friends of \_\_<sub>j</sub> think \_\_<sub>j</sub> likes power entirely too much.  
 b. He I very much DOUBT \_\_ wants to have anything to do with us.  
 c. Robin is someone who(\*m)<sub>i</sub> once I realized \_\_<sub>j</sub> WOULD be coming to the party I made a special point of being nice to \_\_<sub>j</sub> .

The *Barriers* analysis gets these dead wrong: if the two chains are linked purely by Chain Composition in such a way that (31) is good, then certainly (33)a should be good, since the structure is literally identical to that of (32):

- (34) *Him*<sub>i</sub> [Nom]... O<sub>i</sub> [Acc]...t<sub>i</sub> [Acc]...t<sub>i</sub> [Nom]

All that is different is that the case on the filler is phonologically visible. On the other hand, (33)c is nothing more than the mirror image of (32):

(35) *whom*<sub>i</sub> [Acc]... *O*<sub>i</sub> [Nom]...*t*<sub>i</sub> [Nom]...*t*<sub>i</sub> [Acc]

Again, contrary to various urban legends about finite clause subject p-gaps being blocked, there is nothing in the least wrong with the case-neutral version of (33)c, which presumably is structurally absolutely indistinguishable from (35). What makes all the bad cases bad seems to be nothing more than the overt morphological form of the same case specification which supposedly corresponds to good examples when it is covert. The same case, that is, corresponds to a well-formed result when unmarked, but an ill-formed string when spelled out, bad, with no theoretically coherent account even vaguely suggested by the form of derivational approaches to movement phenomena. In response to this serious embarrassment, one might want to assume instead that Case identity between the two chains really was a condition on chain composition—in which case, of course, one would incorrectly predict the badness of (31).

## 4 Conclusion: the superiority of HPSG

We conclude with the following observations:

- The HPSG theory of p-gaps, which is in a sense the HPSG theory of filler/gap UDCs itself, takes the putative ‘true’ and the alleged ‘parasitic’ gaps to be on a complete par with each other. Hence the Kearney paradigm facts are just what we would expect, given the Pollard and Sag (1994) binding theory along with certain processing constraints that seem, in view of (6), to be necessary independently of multiple gap construction.
- The well-formedness of nominative subject p-gaps corresponds to the HPSG null hypothesis, and hence nothing further needs to be said about it.
- The HPSG theory of p-gaps, since it treats all gaps on a par, can treat symbiotic gaps exactly the same as parasitic gaps, assuming the general position on strong islands taken in Pollard and Sag (1994) (and strongly supported by the complementary work reported in Kluender (1998), Kroch (1989) and others sources). Note that the Pollard-Sag Subject Condition predicts the well-formedness of the symbiotic gap examples, since it imposes a language-particular restriction on English grammar that a gap within in an English subject must be matched by a gap in the VP of which the subject is a valent. This formulation correctly predicts the fact that the gaps need not correspond to the same filler, as documented in Hukari and Levine (1989):

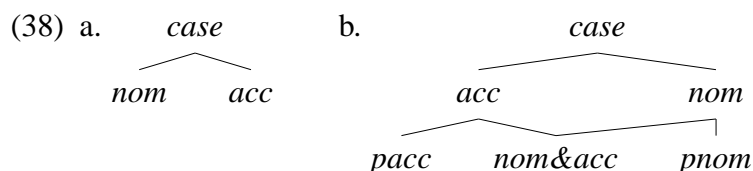
- (36) a. There are [certain heroes]<sub>i</sub> that that I find [long stories about *t*<sub>i</sub> ]<sub>j</sub> too boring to listen to *t*<sub>j</sub> .  
 b. There are [certain heroes]<sub>i</sub> that [long stories about *t*<sub>i</sub> ]<sub>j</sub> invariably prove to be too boring to listen to *t*<sub>j</sub> .  
 c. [Which heroes]<sub>i</sub> are [long stories about *t*<sub>i</sub> ]<sub>j</sub> bound to be too boring to listen to *t*<sub>j</sub> ?

Cf.

- (37) ??\*Which heroes are long stories about bound to be too dull to entertain Robin.

It is not in the least obvious how any of the P&P parasitic gap hypotheses are going to be able to license gaps in subjects that do not correspond to the supposedly ‘true’ gaps which in one way or another are going to have to license them, at least given the apparently universal assumption that English subject positions are syntactic islands.

- The case mismatch facts fall simply and directly out of the case type hierarchy presented in Levine et al. (2000). Briefly, we replace the case subhierarchy of English in (38)a with (38)b:



To implement this solution, Levine et al. (2000) propose the following case values for various English NPs.

(39)	{	he	[CASE <i>pnom</i> ]
		him	[CASE <i>pacc</i> ]
		whom	[CASE <i>pacc</i> ]
		who	[CASE <i>nom&amp;acc</i> ]
		Robin	[CASE <i>nom&amp;acc</i> ]
		t	[CASE <i>case</i> ]

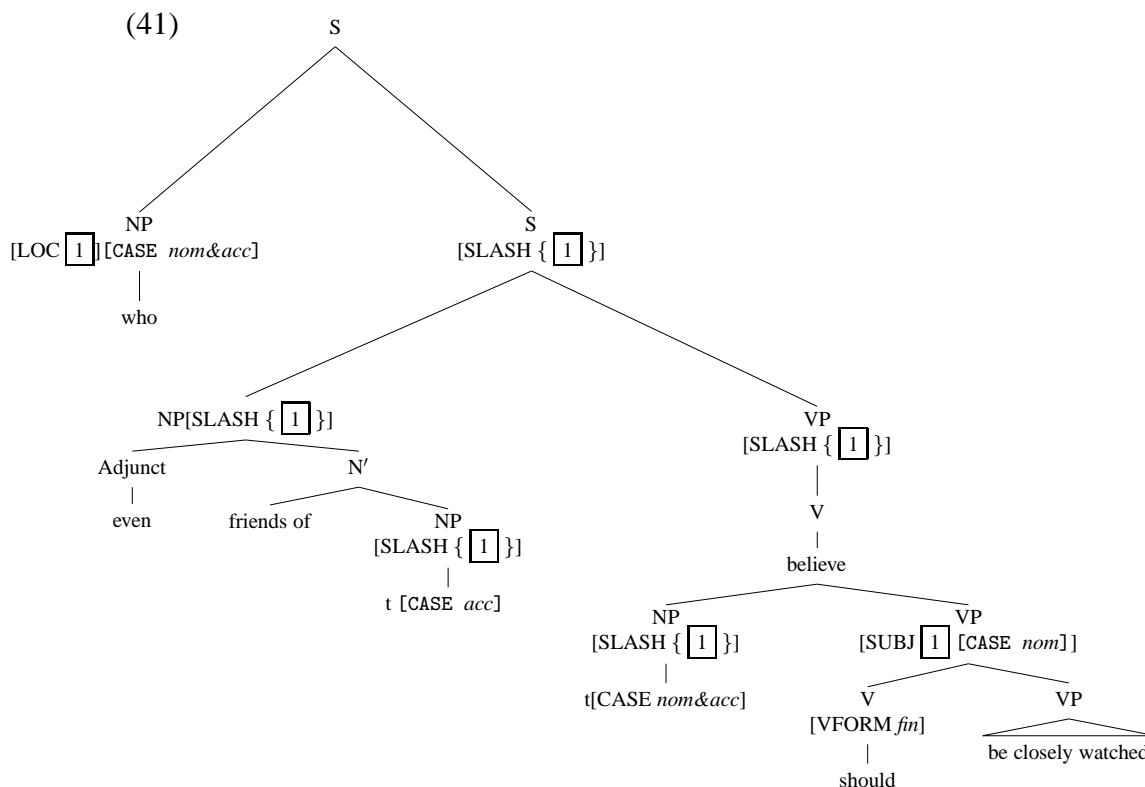
This solution gives us exactly what we want both in more typical examples (where there are constraints on case assignment, but no case mismatches) and in the previously problematic examples like (31 and (33). Consider the following examples:

- (40) a. Who(\*m) likes him?  
       b. Who(m) does he like?

Here, assuming the case theory we have outlined, the selectional properties of *likes* assign *nom* to the subject of this verb in the first example, and *acc* to its object in the second. However, on our proposal *nom* is really just an abbreviation for either *pnom* or *nom&acc*. The word *who* is, as specified in (39), *nom&acc*, and is compatible with the case assignment *pnom*, ensuring that the first example will be licensed when *who* is the filler. The word *whom*, however, is specified as a (pure) *acc* element, so it is not compatible with an *nom* case assignment, and therefore the first example is correctly ruled out when *whom* is the filler. Similarly, in (40)b, *acc* is just an abbreviation for *acc* or *nom&acc*, and *who* is *nom&acc*, compatible with the *acc* specification of the object trace, while *whom* is *pacc*, which is also compatible with *acc*. Both variants of this example are therefore correctly predicted to be good.

We are now able to address the case connectivity problem in parasitic gap examples like (31). On the account we have provided, this problem essentially disappears. The following tree illustrates the licensing of such sentences:





The short story here is that the morphsyntactic specifications included under the LOC description are token-identical to those of SLASH, which, looking at it from the ‘top’ of the representation, appears on any subset of the daughters of the clausal sister to the filler, and is shared between mother and daughter down to a point where it appears in the syntactic description of a category which also structure-shares that LOC specification. All components of the LOC description must be matched, including the case value. The case value assigned to the subject is *nom*, which subsumes the description *nom&acc*; the case value assigned to the object is *acc*, which also subsumes that description as per (38)b. Hence both subject and object can be token identical to the LOC value of the filler, each can be realized as an empty category, and the example is licensed.

As a last resort, one can imagine an effort to incorporate the HPSG analysis I’ve just sketched into something like the *Barriers* analysis. But it’s hard to imagine a natural, or even remotely plausible way this could be done while still allowing empty-operator-headed chains to involve genuine out-and-out case inconsistency as necessary, for example, in missing object constructions:

(42)  $He_i$  is [ $O_i$  tough to please  $t_i$ ]

The problem of course is that *he* is *pnom*, while the trace, and the empty operator that shares its  $\phi$  features, is either *pacc* or *nom&acc*. It thus appears that the relationship between chains in p-gap constructions would have to impose a condition of chain consistency which somehow gave rise to the effects alluded to above, while the relationship between the antecedent chain and the MO chain in MOCs would have to overlook such conflict. I am not aware of any means to ensure this outcome in a plausible, nonstipulative fashion.

One possibility would be to incorporate something like the case hierarchy already sketched into the Frampton account of multiple-gap constructions, notwithstanding the empirical failure of that account in the face of symbiotic gap constructions. One could imagine a parallel analysis to the HPSG account just sketched, with the filler of conjunctive case type, linked by two separate chains to gap

sites compatible with the type *nom&acc*. Nowhere in P&P Case theory, to my knowledge, has anything along these lines been proposed, but let's assume it to be possible. The fact is that the possibility of instantiating such a solution in Frampton's analysis should not be surprising, given the nature of that analysis—which is, fundamentally, an effort to implement a single filler/multiple gap analysis by an archaic theoretical technology forced to achieve the desired effect via separate establishment of multiple histories and then replacement of all but one of the fillers with traces—in a manner which is essentially arbitrary if it turns out, as I have argued, that there is no structural basis for positing an asymmetry between some 'true' history of derivation on the one hand and the remaining parasitic histories. To put it bluntly (but, I think, fairly), Frampton's analysis is an inevitably clumsy effort to replicate the feature-percolation model of UDCs in classical GPSG using a singularly unsuitable bit of machinery based on movement. Grafting a type-hierarchical solution from HPSG onto an awkward transformational simulation of a natural phrase-theoretic extraction treatment merely underscores the deficient nature of the movement analysis that requires many so many imported fixes.

In short, none of the phenomena I have surveyed in this paper—the Kearney paradigm, WCO effects, the distribution of finite clause gaps, pronominality, symbiotic gaps and case inconsistency—support any deviation from the strongest possible hypothesis about extraction, which is that there is a single mechanism linking a single filler to all gap sites, and that there is no asymmetry in any respect in the establishment of these multiple linkages. And this is an outcome which follows directly from the constraint-regulated feature percolation architecture of HPSG. It does not require an at best multistage, formally inexplicit *simulation* of a direct, symmetrical linkage; it expresses this linkage directly, as the null hypothesis—an hypothesis which I think the evidence shows is not just the strongest but also the most likely to be correct.

## References

- Chomsky, N. 1977. On *wh*-movement. In P. Culicover, T. Wasow, and A. Akmajian, eds., *Formal Syntax*, pages 71–132. Academic Press.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. 1982. *Concepts and Consequences of the Theory of Government and Binding*. Cambridge: MIT Press.
- Chomsky, N. 1986. *Barriers*. Cambridge, Mass.: MIT Press.
- Chomsky, N. 1986a. *Barriers*. Cambridge, Mass.: MIT Press.
- Chomsky, N. 1995. Bare phrase structure. In G. Webelhuth, ed., *Government-binding theory and the Minimalist Program*. Cambridge, Mass.: Blackwell's.
- Cinque, G. 1990. *Types of  $\bar{A}$ -Dependencies*. Cambridge, Mass.: MIT Press.
- Contreras, H. 1984. A note on parasitic gaps. *Linguistic Inquiry* 15:698–701.
- Frampton, J. 1990. Parasitic gaps and the theory of *wh*-chains. *Linguistic Inquiry* 21:49–77.
- Gazdar, G., G. K. Pullum, I. Sag, and T. Wasow. 1982. Coordination and transformational grammar. *Linguistic Inquiry* 13:663–676.

- Haegeman, L. 1984. Parasitic gaps and adverbial clauses. *Journal of Linguistics* 20:229–232.
- Hukari, T. E. and R. D. Levine. 1989. Category antirecursion: Paradoxical consequences of gap–within–filler constructions. In *Proceedings of the Seventh Annual West Coast Conference on Formal Linguistics*. Stanford: CSLI.
- Kearney, K. 1983. Governing categories. Unpublished ms. University of Connecticut, Storrs, CT.
- Kiss, K. 2001. Parasitic chains revisited. In P. Culicover and P. Postal, eds., *Parasitic Gaps*. Cambridge, Mass.: MIT Press.
- Kluender, R. 1998. On the distinction between strong and weak islands: a processing perspective. In *The Limits of Syntax*. New York: Academic Press.
- Kroch, A. 1989. Amount quantification, referentiality, and long *wh*–movement. Unpublished ms., Department of Linguistics, University of Pennsylvania.
- Lasnik, H. and T. Stowell. 1991. Weakest crossover. *Linguistic Inquiry* 22:687–720.
- Levine, R., T. E. Hukari, and M. Calcagno. 2000. Some overlooked parasitic gap constructions in English and their theoretical implications. In P. Culicover and P. M. Postal, eds., *Parasitic Gaps*, pages 181–222. Cambridge: MIT Press.
- Levine, R. D. 2001. The extraction riddle: just what are we missing? *Journal of Linguistics* 37:145–174.
- Levine, R. D. and Ivan A. Sag. 2003. *Wh*-Nonmovement. *Gengo Kenkyu: Journal of the Linguistic Society of Japan* 123: 171-219.
- Munn, A. 2001. Explaining parasitic gap restrictions. In P. Culicover and P. Postal, eds., *Parasitic Gaps*. Cambridge, Mass.: MIT Press.
- Pollard, C. J. and I. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Postal, P. M. 1993b. Remarks on weak crossover. *Linguistic Inquiry* 24:539–556.
- Postal, P. M. 1998. *Three Investigations of Extraction*. Cambridge, Mass.: MIT Press.

Rober Levine  
Ohio State University  
levine@ling.ohio-state.edu