Some Empirical Issues in the Grammar of Extraction

Robert D. Levine
Ohio State University

Ivan A. Sag
Stanford University

Abstract

This paper compares transformation-based and constraint-based treatments of unbounded filler-gap dependencies, the latter specifically as articulated in terms of HPSG, and argues, contrary to the commonly made allegations of 'notational variance', that there is purely empirical evidence that is consistent with only the constraint-based account. Recent proposals to deal with parasitic gaps in terms of null pronominals and 'empty operators' are unable to account for the phenomenon of 'symbiotic' gaps, the apparent case mismatches found in parasitic gap constructions, or (in general) for the well-known 'across-the-board' effects within coordinate structures.
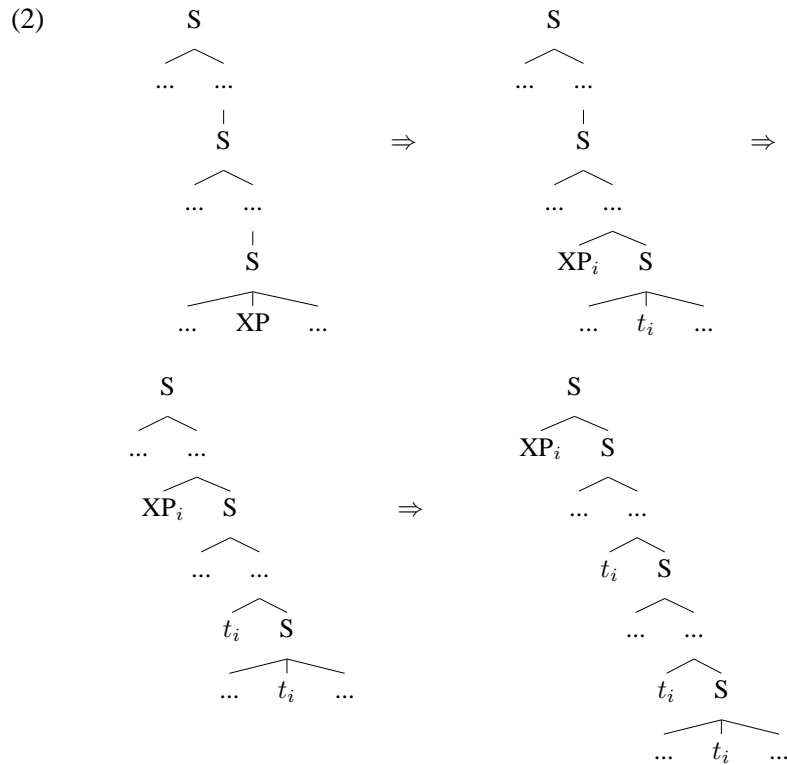
1   Filler/Gap Constructions: Two Approaches

Historically, filler/gap constructions (or unbounded dependency constructions – UDCs) such as those in (1) have been approached two ways:

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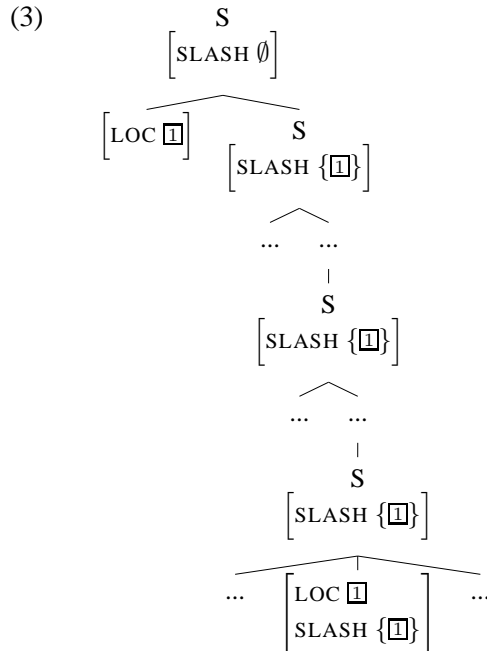
Transformational approaches posit a sequence of representations in which the filler is initially in the position notated by the underline in (1), which is then relocated, possibly via a series of movement steps, to its final position on the left of the highest clause. Schematically, the derivational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

(2)

S
... ...
S
... ...
S
... XP ...

$\Rightarrow$

S
... ...
S
... ...
XP_i S
... t_i ...

$\Rightarrow$

S
... ...
XP_i S
... ...
t_i S
... t_i ...

$\Rightarrow$

S
XP_i S
... ...
t_i S
... ...
t_i S
... t_i ...

The bottom-up derivations found in current work within the Minimalist Program are similar in relevant respects. They differ primarily as to where the higher structure is introduced within a derivation.

There are two crucial aspects to the analysis depicted in (2): (i) the filler is the same object at the end of the derivation as the in-situ category at the beginning of the derivation, merely relocated by movement, and (ii) a series of intermediate traces is left at each of the positions occupied by the trace in transit in addition to the trace demarcating its original position prior to movement. Compare this picture to the HPSG connectivity mechanism linking fillers and gaps given in (3):

(3)

S
[SLASH ∅]

[LOC 1]   S
           [SLASH {1}]

            ...   ...
              |
              S
              [SLASH {1}]

               ...   ...
                 |
                 S
                 [SLASH {1}]

          ...   [LOC 1   ]   ...
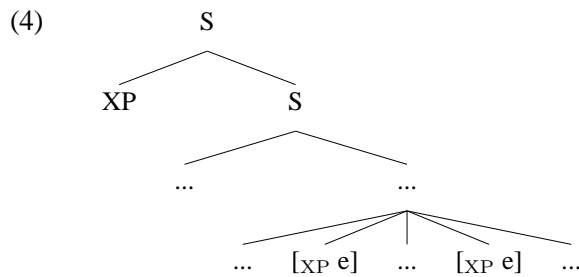   [SLASH {1}]

Casual comparison of (2) and (3) would suggest that these representations are essentially equivalent, *as long as you only look at single filler/gap constructions.* Indeed, Chomsky has insisted, over much of his career, on the empirical indistinguishability of monostratal representations with 'base generated gaps' with derivationally derived gaps as per (2). In *LGB*, for example, he not only asserts their 'virtual indistinguishability', arguing that the problem of choosing between them is 'a fairly marginal one', but makes the unsubstantiated (and factually incorrect) claim that all nonderivational theories of filler/gap linkages are 'transformational theories, whether one chooses to call them that or not)'. Over the past two decades, the notion seems to have circulated in certain circles that monostratal feature-linkage analyses of filler/gap constructions are nothing more than old wine in new, not very interesting bottles.

This is a charge that might be legitimately levelled at GB treatments of syntactic unaccusativity vis-à-vis the original Relational Grammar studies of that phenomenon. However, we argue that it has no merit in the comparison of (2) and (3). Not only are there clear framework-architectural differences between the approaches, it turns out that multiple gap constructions make very clear, on purely factual grounds, the inferiority of derivational approaches.

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):

(4)

```
                    S
                  /   \
               XP      S
                      /  \
                   ...    ...
                         / | \
                      ...[XP e] ...[XP e] ...
```

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 under the movement analysis. Therefore there appear to be only two possible choices:

- there is a single linkage mechanism between the filler and one of the gaps and a different 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 mechansim.

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 element which is something other than a trace. In the second case, movement is the sole linkage mechanism involved 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 a second moved element that must be 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.

Both variants, as well as various hybrids, exist in the literature. Sticking to very familiar examples, Chomsky 1982 manifests the first alternative and Chomsky 1986 the second. But the plausibility of such approaches is only as strong as the arguments for the asymmetry assumed. There are remarkably few of these, in fact.

2.1 The Kearney Paradigm

The primary argument in the literature, as far as we are aware, is 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:

> [e]xample [(5a)] is a normal parasitic gap construction, but [(5b)] is ungrammatical. It follows, then, that the *wh*-phrase in [(5a)], [(5b)] 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'.

Frampton (1990, p. 58) cites the same data in support of Chomsky's line of reasoning about the source of (5). While hardly transparent, that reasoning appears to be the following: if p-gap constructions were in fact instances of some kind of multiple-gap (i.e. symmetrical) phenomenon, then reconstruction of the filler should proceed symmetrically to yield identical effects in (5a) and (5b). In both cases, the result would be a representation in which an anaphor was compatible with its antecedent in one of the sites but not in the other. Hence, on the crucial assumption that the ill-formedness of (5b) arises from reconstruction of an anaphor into a gap site where only an incompatible antecedent is present, we would expect (5a)—where *which books about himself* is reconstructed to a site where *Mary* must antecede the anaphor—to be just as bad. But this is not what we find. Rather, the general pattern is that when the anaphor is compatible with a main clause subject antecedent, the result is good, and when it is not, the result is bad. Hence, the simplest conclusion is that the overt filler reconstructs only to the main clause gap site, which must then be its transformational point of origin.

But this conclusion is inconsistent with previously overlooked examples like (6):

(6) a. There were pictures of herself which, once Mary finally decided she liked __, John would have to put __ into circulation.

   b. There were pictures of himself which, once Mary finally decided she liked __, John would be able to put __ into circulation.

(6a,b) instantiate the 'fronted adverbial' p-gap construction discussed in general terms for the first time, to our knowlege, in Haegeman 1984. Examples like these demonstrate that binding patterns reveal nothing about the extraction site of the *wh*-phrase, even on Chomsky's own line of reasoning. No matter which gap is taken to be the 'true' gap in adverb fronting, the fact that both *John* and *Mary* are possible reflexive antecedents shows that the distinction between true and parasitic gap is irrelevant to the determination of anaphor binding.

These observations, incidentally, are exactly as predicted by the convergent binding theories of Pollard and Sag (1992) and Reinhart and Reuland (1993), who demonstrate that Principle A cannot be the basis for determining the antecedent of anaphors in 'picture noun' phrases. Clearly, extragrammatical factors such as point-of view, intervening potential controllers, and proximity play a significant role in defining the notion of prominence that determines well-formedness in cases

like those we have been looking at. The importance of proximity is underlined by further contrasts like the following, involving across-the-board extraction:

(7) Which pictures of himself/*herself did John approve of __ and Mary like __ enormously?

This observation about ATB extraction is not inconsistent with the assumption that picture noun reflexives are governed by extragrammatical factors, as argued at length by Pollard and Sag (1992, 1994). Under the assumptions made by Chomsky or by Frampton, however, these data make no sense whatsoever. The conclusion we come to then is that the Kearney paradigm has been badly misunderstood since it was first introduced into the literature as a justification for the putative 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. Example such as those in (8) are often offered as illustrations of this claim:

(8) a.*Jack, who$_i$ I heard about __$_i$ before you said __$_i$ would hire us... (Frampton 1990, p.68.)

b.*Someone who$_i$ John expected __$_i$ would be successful though believing __$_i$ is incompetent... (Chomsky 1982, p.55)

c.*The militant who they arrested __$_i$ after learning __$_i$ was 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 the examples in (9):

(9) a. [Which people]$_i$ did you invite __$_i$ without thinking __$_i$ would actually come.

b. Jack, who$_i$ even before you said __$_i$ would hire us I was favorably disposed towards __$_i$, is a prince among men.

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   Symbiotic Gaps

The foregoing discussion has shown that the chief published arguments for chain asymmetry in derivational theories of p-gap licensing are unsound. We now examine evidence that poses further difficulties for chain-asymmetric approaches to multiple gap constructions. Consider the data in (10), where *both* gaps seem to be within islands:
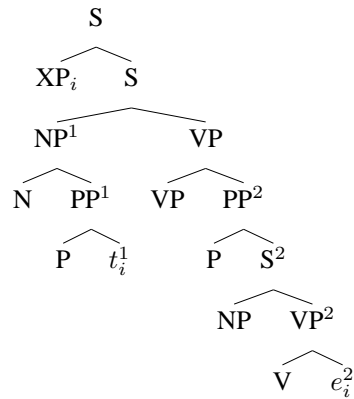
(10)   a.   What kinds of books do authors of __ argue about royalties after writing __?

   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 we 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. We reason as follows: First, under Chomsky's (1982) approach in *Concepts and Consequences* (see 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.

Next, on Kayne's 1983 'connectedness' approach, a parasitic gap can only establish a connection to a parasitic gap if the path from the parasitic gap to the true gap can be continuously mediated in terms of what Kayne calls the g-projection path. Longobardi (1984) showed that in order to work correctly, Kayne's definition of g-projection path had to be strengthened with a proper government requirement. 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 (10) presumably unlicensed, as charted in (11), where superscripts indicate g-projections:

(11)

```
                    S
                  /   \
               XP_i    S
                     /    \
                  NP^1     VP
                 /   \    /    \
                N    PP^1 VP    PP^2
                    /  \       /   \
                   P   t_i^1  P     S^2
                                  /    \
                                NP     VP^2
                                      /    \
                                     V     e_i^2
```

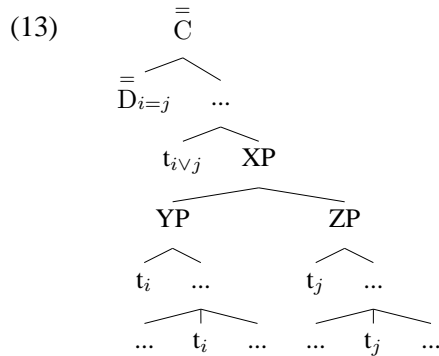But in Chomsky's (1986) account (the *Barriers* analysis), both the adjunct and the subject function both as barriers and as blocking categories, which ensures that the dominating maximal projections closest to them (VP and IP respectively) are barriers. On this analysis, the empty operator within the subject cannot remain in situ since it will receive no intepretation at LF. But it cannot move out of the NP(DP) since, by stipulation, it can neither adjoin to NP(DP) nor move to Spec of CP, since that would involve crossing two barriers. But even if it could move out of the NP(DP) to [Spec,CP], it would be separated from the empty operator heading the parasitic chain by the barriers CP and VP, both of which are (intended to be) barriers for the empty operator heading the parasitic chain. Therefore the approach in Chomsky 1986 makes the incorrect prediction that examples such as (10) are ill-formed, as shown in (12), where unoccupied Spec positions have been suppressed:

(12)

```
                       C̿
                     /    \
                  D̿_{i,j}   C̿
                             |
                             C̄
                           /    \
                          C      Ī̿
                              /      \
                            D̿         Ī
                          /   \      /   \
                        D     P̿     I     V̿
                            /   \        /    \
                           P    t_i    V̿      P̿
                                      /    \
                                     P     C̿
                                         /    \
                                       ...    ...
                                            /    \
                                          t_j    ...
```

244
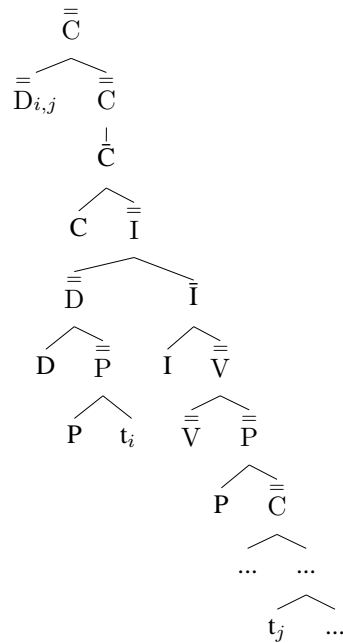
Finally, Frampton's (1990) treatment of parasitic gaps, a kind of hybrid of Kayne's connectedness with Chomsky's null operator treatment in *Barriers*, is in effect a derivational reconstruction of the multiple licensing of extractions pathways linked to a single filler. Everything we've said about Kayne carries over directly. We need only replace the notion 'g-projection' with 'trace-chain' and 'connectness' with 'inverted Y-path':

(13)

$$\bar{\bar{C}}$$

$$\bar{\bar{D}}_{i=j} \quad ...$$

$$t_{i \lor j} \quad XP$$

$$YP \qquad ZP$$

$$t_i \quad ... \qquad t_j \quad ...$$

$$... \quad t_i \quad ... \qquad ... \quad t_j \quad ...$$

And the same problem with connectedness in these cases carries over to Frampton's trace-based analogue. The upshot of all this is that no reasonably explicit P&P theory of p-gaps has even the beginnings of an adequate account of symbiotic gaps. In section 3.2 below, we propose a reassessment of the data in (10) and sketch an account in terms of Pollard and Sag's (1994) Subject Condition.

2.4 The Case Conflict Conundrum

Finally, consider examples such as (14):

(14) Robin is someone who$_i$ even good friends of __$_i$ believe __$_i$ 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 aysmmetry 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 ϕ features, so that we would have the situation in (15):

(15) *wh*$_i$ [Nom]... *O*$_i$ [Acc]...t$_i$ [Acc]....t$_i$ [Nom]

So the possibility of case mismatches seems to be predicted. This might appear to be a plus for the asymmetrical chain analysis.

245

But appearances are often deceptive. It turns out that none of the movement approaches we have 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 why then do we have the following data?

(16)   a.*Him$_i$, even friends of __$_i$ think __$_i$ likes power entirely too much.

     b. He$_i$, I very much DOUBT __wants to have anything to do with us.

     c. Robin is someone who(*m)$_i$ once I realized __$_i$ WOULD be coming to the party I made a special point of being nice to __$_i$.

The *Barriers* analysis gets these facts dead wrong: if the two chains are linked purely by Chain Composition in such as way that (14) is good, then certainly (16a) should be good, since the structure is literally identical to that of (15):

(17) *Him$_i$* [Nom]... O_i [Acc]...t$_i$ [Acc]....t$_i$ [Nom]

All that is different is that you can *see* the case on the filler, i.e. the pronoun *him* shows its case morphologically. On the other hand, (16c) is nothing more than the mirror image of (15):

(18) *whom$_i$* [Acc]... O_i [Nom]...t$_i$ [Nom]....t$_i$ [Acc]

Again, though it seems to be something of an urban legend that finite clause subject p-gaps are ungrammatical, there appears to be nothing ungrammatical about the case-neutral version of (16c), which presumably is structurally indistinguishable from (18). 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. Why is the same case good when it has no morphological realization, and bad when it does not? Alternatively, one could assume that Case identity between the two chains really *was* a condition on chain composition – in which case, one would incorrectly predict the badness of (14). This dilemma seems deeply problematic. Moreover, a variant of this double bind undercuts every one of the movement-based approaches we have considered, and various others as well.

2.5   Across-the-Board Extraction

Finally, let us now consider multiple gaps in coordinate structures. Critical examples here include the following:[1]

---

[1]We ignore here the issue of asymmetric conjunction and apparent counterexamples to the Coordinate Structure Constraint. For discussion and debate on the status of this constraint, see Postal 1998 (Chapter 3), Levine 2001, and Kehler 2002 (Chapter 5).

(19) a.*[Which dignitaries]$_i$ do you think [[Sandy photographed the castle] and [Chris visited __]]?

     b.*[Which dignitaries]$_i$ do you think [[Sandy photographed __] and [Chris visited the castle]]?

     c. [Which dignitaries]$_i$ do you think [[Sandy photographed __] and [Chris visited __]]?

(20) a.*[Which of her books]$_i$ did you find both [[a review of Gould] and [a reply to __]]?

     b.*[Which of her books]$_i$ did you find both [[a reply to __] and [a review of Gould's new book]]?

     c. [Which of her books]$_i$ did you read both [[a review of __] and [a reply to __]]?

These are of course the familiar data commonly referred to as Ross's (1967) Coordinate Structure Constraint and its 'across-the-board' exceptions.

As noted earlier, Gazdar et al. (1982) showed that a single mechanism linking fillers and gap sites in all relevant cases is in principle unavailable under the movement analysis. That is, it remains unclear how multiple gaps in across-the-board extraction structures are to be associated with a single filler. This objection has never been properly addressed in the transformational literature of the two decades that have transpired since the publication (in *Linguistic Inquiry*) of Gazdar et al.'s paper. We take this to be a testament to the correctness of Gazdar et al.'s conclusions.

3 A Feature-Based Analysis of Multiple Gaps

3.1 The Feature-Based Analysis of UDCs

The constraint-based phrase-structure theoretic analysis of parasitic gaps incorporates the fundamental insights about this phenomenon that begin with Gazdar 1981 – in particular, the observation that in the absence of any constraint to the contrary, a SLASH specification on a mother category can match a separate identical SLASH specification on each of any number of daughters. In Pollard and Sag 1994, this account of the origin of parasitic gaps is built into the formulation of the Nonlocal Feature Principle given in (21):[2]
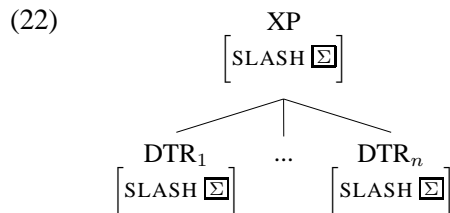
(21) **The Nonlocal Feature Principle (NLFP):**
     In any construction, the mother's SLASH value is the union of the daughters' SLASH values minus the BIND value of the head daughter.

---

[2]This formulation of the NLFP is restricted to the feature SLASH. Relative clauses have been treated in terms of the nonlocal feature REL (Pollard and Sag 1994; Sag 1997). Ginzburg and Sag (2000) treat interrogatives and exclamatives in terms of the nonlocal feature WH.

The value of BIND will be specified so that it is empty in general, but will contain an appropriate element v, just in case a given word (e.g. *tough*) or construction licenses the introduction of non-empty SLASH specification containing v. Above any such binding point, the set value of SLASH will not contain v. Thus BIND plays the role of a regulator, ensuring that nonlocal feature values only appear at the point where they are 'launched', and only propagate down below this point to the place in the structure where they are cashed out as a gap.[3]

Note, in particular, that as long as two daughters of a given category share identical SLASH values, that single SLASH value will also appear on the mother as the union of its daughters' specifications for SLASH, and the same structure can be extended to include any number of daughters:

(22)

$$
\begin{array}{c}
\text{XP} \\
\left[\text{SLASH } \boxed{\Sigma} \right]
\end{array}
$$

$$
\begin{array}{ccc}
\text{DTR}_1 & & \text{DTR}_n \\
\left[\text{SLASH } \boxed{\Sigma} \right] & \cdots & \left[\text{SLASH } \boxed{\Sigma} \right]
\end{array}
$$

Unlike earlier feature-based proposals, e.g. that of Gazdar et al. 1985, here there is no pressure on SLASH to follow a path from head to head, wherever else it may appear.[4] Hence the NLFP provides a unified account of individual gaps (on or off head paths) and multiple-gap constructions, where both head and nonhead paths bear identical SLASH features. Note further that this same mechanism will yield both of the following structures:[5]
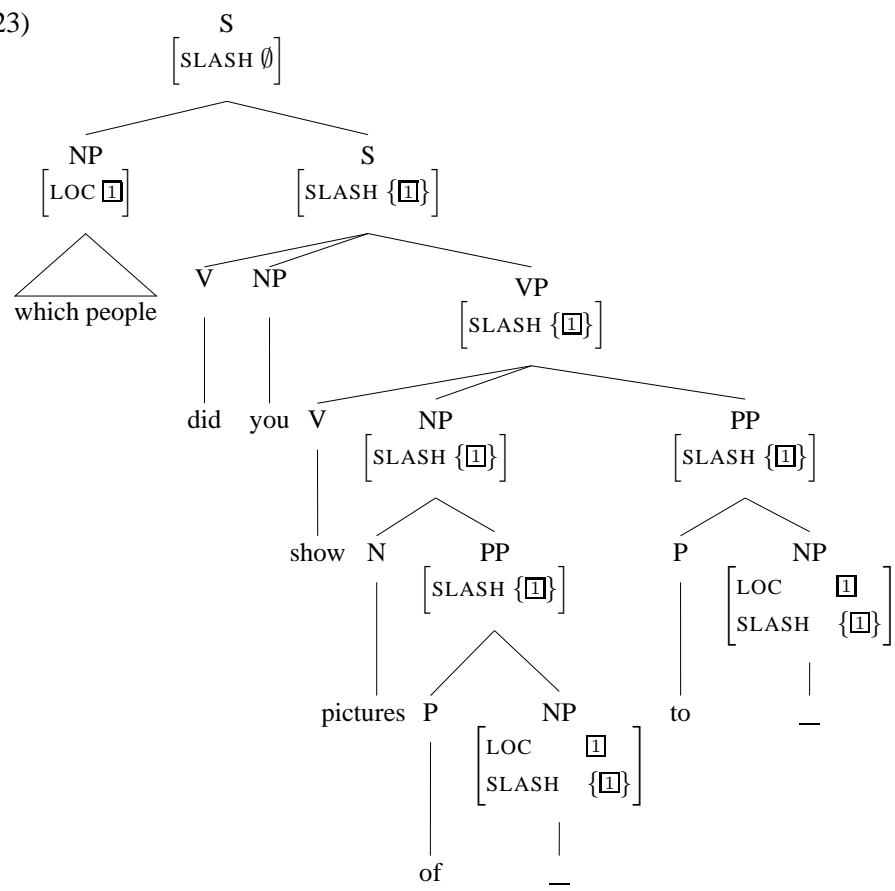
---

[3]In the case of SLASH. Other nonlocal features, such as WH or REL, will be cashed out as appropriate *wh*-words. Our BIND feature plays a role similar to that of Pollard and Sag's (1994) TO-BIND feature.
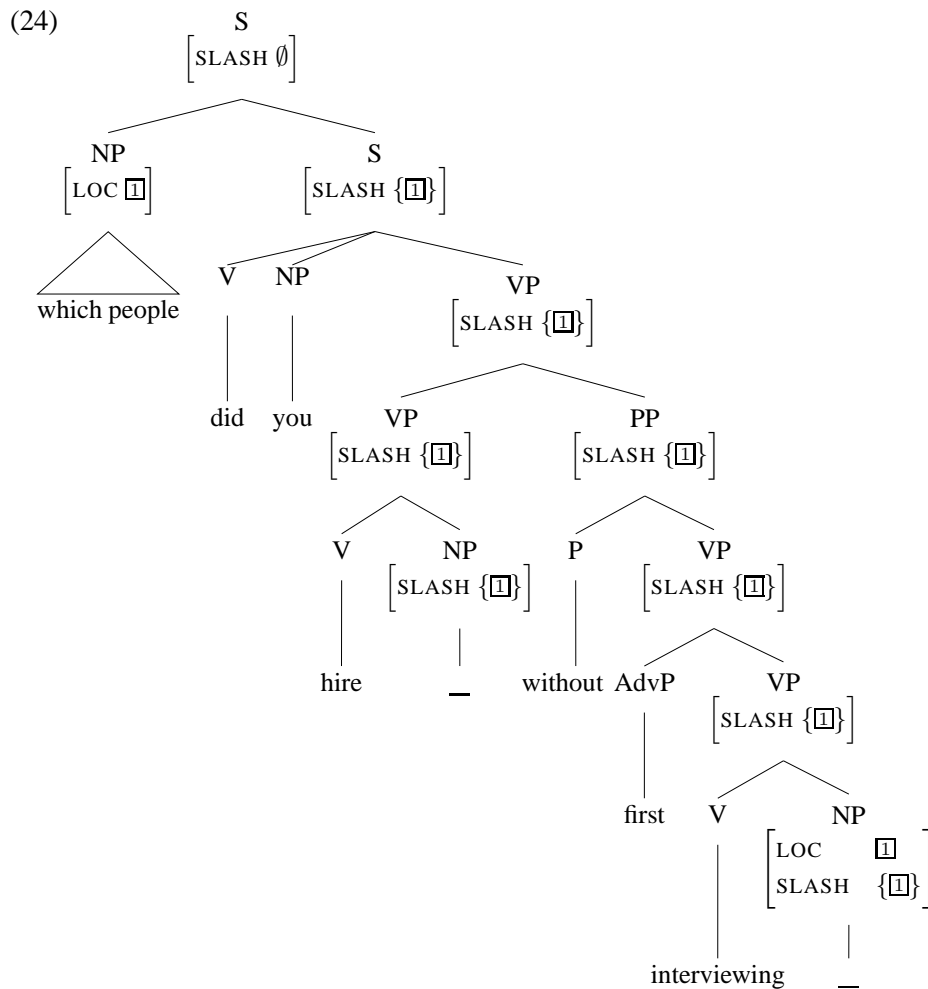
[4]But see the proposal of Ginzburg and Sag (2000).

[5]We appeal to binding theory to account for the deviance of examples like (i):

(i) *Who did they explain $\_\_i$ to $\_\_i$.

(23)

S
[SLASH ∅]

NP
[LOC ①]

which people

S
[SLASH {①}]

V — did

NP — you

VP
[SLASH {①}]

V — show

NP
[SLASH {①}]

N — pictures

PP
[SLASH {①}]

P — of

NP
[LOC ① / SLASH {①}]
___

PP
[SLASH {①}]

P — to

NP
[LOC ① / SLASH {①}]
___

(24)

S
[SLASH ∅]

NP
[LOC [1]]
　　which people

S
[SLASH {[1]}]

V — did
NP — you

VP
[SLASH {[1]}]

VP
[SLASH {[1]}]

V — hire
NP
[SLASH {[1]}]
　　___

PP
[SLASH {[1]}]

P — without

VP
[SLASH {[1]}]

AdvP — first

VP
[SLASH {[1]}]

V — interviewing

NP
[LOC [1]]
[SLASH {[1]}]
　　___

Very similar structures give rise to subject parasitic gaps, such as (25):

(25) Which of the candidates do you think my talking to $\_\_\_i$ would bother $\_\_\_i$?

3.2 A Reassessment of Symbiotic Gaps

Much of the literature on parasitic gaps in English has assumed that a gap within an adverbial phrase is on a par with one within a subject phrase in that both require the presence of another coindexed gap in order to be legitmate. Pollard and Sag (1994, Chapter 4) challange this assumption, citing examples like the following, where extraction out of adverbials is possible without the presence of any additional gap performina a 'licensing' function:

(26) a. That's the symphony that Schubert died [without fi nishing ___].

 b. Which room does Julius teach his class [in ___]?

c. Who did you go to Girona [in order to meet __]?

d. What kind of wagon did they used to ride to school [in __]?

e. How many of the book reports did the teacher smile [after reading __]?

f. This is the blanket that Rebecca refuses to sleep [without __].

But if these examples are well-formed (as they certainly seem to be), then we need to rethink the 'parasitic' nature of examples like (10b), repeated here as (27b):

(27) 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?

In short, what seems empirically motivated is an approach to island phenomena that appeals to independent, partly extragrammatical factors that will explain why extraction out of adverbial phrases are sometimes of reduced acceptability. More-over, one of the relevant factors is the presence of an overt direct object NP in the preceding VP. Controlling for this or other (only partly understood factors) restores full acceptability to the putative island-violating extractions:

(28) a. What kinds of books do authors of malicious pamphlets get sick after writing __?

b. What kinds of books do authors of malicious pamphlets congratulate each other after writing __?

c. Which of our books did the authors get fired after writing __?

'Parasitism', at least in the case of gaps within adverbial phrases, is an illusion. We may thus delimit the scope of our account of parasitic gaps to deal with ex-tractions out of subjects, which seems to be possible only if a gap appears in some subsequent constituent.[6]

---

[6]Some might argue further that extractions from subjects, even in the absence of a licensing 'primary' gap, are in principle grammatical:

(i) (?)There are certain topics that jokes about __ are completely unacceptable.

(ii) (?)There are certain dignitaries that my jokes about __ are always considered over the top.

(iii) (?)There are certain dignitaries that my talking to __ would be considered improper.

We want to emphasize that this assessment of the facts would simplify our grammar further, allowing even the constraint that we are about to introduce to be eliminated.

251

Pollard and Sag (1994, Chapter 4) propose a principle they call the Subject Condition:[7]

(29)   Subject Condition:

       The initial element of a lexical head's ARG-ST list may be slashed only if that list contains another slashed element.

This predicts the familiar contrast in (30):

(30)   a.*That was the rebel leader who rivals of __ assassinated the British consul.

      b. That was the rebel leader who rivals of __ assassinated __.

This is because only the ARG-ST list of the verb *assassinated* in (30b) satisfies (29). Similarly, the contrast between (31a) and (31b) is accounted for, as illustrated in (32):

(31)   a.*Who did my talking to __ bother Hilary?

      b. Who did my talking to __ bother __?

(32)   Partial lexical entry for *assassinate* or *bother*:

$$\left[\text{ARG-ST} \quad \left\langle NP_1 \, , NP_2 \right\rangle \right]$$

The Subject Condition ensures that NP_1 can have a nonempty SLASH value just in case NP_2 also does.

And this approach immediately extends to explain the contrast between (27a,c) if we incorporate the 'adverbs as complements' analysis that has been proposed on entirely independent grounds by numerous researchers, including Bouma et al. 2001 and Przepiórkowski 1999. On this analysis, the ARG-ST of verbs is extended to include certain adverbials that are selected by the verb as though they were a complement. We will assume that this includes *after*-phrases, which we treat as a kind of PP. This leads to an ARG-ST list like the following as one possibility for the verb *argue*:

(33)   Partial lexical entry for *argue* with extended ARG-ST list:

$$\left[\text{ARG-ST} \quad \left\langle NP \, , PP_1 \, , PP_2 \right\rangle \right]$$

The Subject Condition ensures that NP_1 can have a nonempty SLASH value just in case PP_1 or PP_2 also does. This accounts for the contrast between (27a,c), as well as correctly predicting the grammaticality of the following examples:

(34)   a. What kinds of books do authors of __ always argue about __ (after hours)?

      b. What kinds of books do authors of __ always argue about __ after finishing __?

3.3 The Case Conflict Conundrum Reconsidered

Since its inception, the phrase-structure theoretic approach to UDCs has assumed that the filler in parasitic gap constructions is linked by the same connectivity mechanism—the propagation of a SLASH feature—to all of the gaps that share its LOC specifications. And this assumption, taken together with an explicit proposal for the values of the feature CASE put forth by Levine et al. (2000), provides a resolution of the troublesome case conflict data discussed in the previous section (examples repeated here):

(35) Robin is someone who$_i$ even good friends of __$_i$ believe __$_i$ likes power entirely too much.

(36) a.*He$_i$ /*Him$_i$, even friends of __$_i$ think __$_i$ likes power entirely too much.

b.*Whom do even friends of __$_i$ think __$_i$ likes power entirely too much?

As Levine et al. show, the modeling assumptions of HPSG interact with lexical underspecification to predict exactly the observed contrasts. They assume that the case values form a semi-lattice structure like (37), where *p-nom* and *p-acc* stand for 'pure' nominative and accusative case, respectively:

(37)

$$\textit{case}$$

```
              case
            /      \
         nom        acc
        /    \      /
   p-nom  nom&acc  p-acc
```
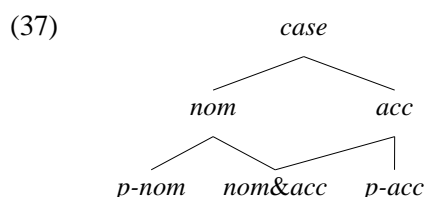
This assumes that there is a case value *nom&acc* that is compatible with both the constraints imposed by prepositions on their objects (that they be some subtype of *acc*) and those that finite verbs impose on their subjects (that they be some subtype of *nom*). Because a selector (verb, preposition, etc.) only bounds the CASE value of its argument(s) (rather than resolving it), the conflict in an example like (35) is only apparent. This is because various expressions, for example *who* and proper names, are lexically unspecified for case, and hence can be resolved to the *nom&acc* value in order to satisfy both selectional demands simultaneously. By contrast, the lexical entries for inflected nominals like *he*, *him*, and *whom* all include fully resolved case specifications: *p-nom*, *p-acc*, and *p-acc*, respectively. And since *p-nom* and *p-acc* are not only incompatible with each other, but also with the value *nom&acc*, there is no way to simultaneously satisfy the grammar's constraints in examples like (36a,b). The constraint-based approach to UDCs thus provides a satisfying solution to the vexed problem of case conflict in parasitic gaps which, as we have seen, has stymed transformational approaches to UDCs.

3.4 Across-the-Board Extraction

Finally, let us now reconsider coordinate structures. All analyses of coordination must posit some identity condition holding between the mother and the daughters (the conjuncts) of a coordinate structure. This is often assumed to be a requirement of category identity, though the precise resolution of examples like (38), first analyzed by Sag et al. (1985), remains as a challenge to most current accounts:

(38)   a. Kim is a Republican and proud of it.

     b. You can rely on our loyalty and that we will do everything in our power to protect you.

But any version of the identity condition is compatible with the constraint-based approach to extraction, as long as it includes the requirement that (in true conjoined structures) the SLASH value of the conjunct daughters must be identical. This requirement, taken together with the analysis of UDCs outlined above, provides an immediate account of the CSC/ATB contrasts considered earlier, repeated here:

(39)   a.*[Which dignitaries]$_i$ do you think [[Sandy photographed the castle] and [Chris visited __]]?

     b.*[Which dignitaries]$_i$ do you think [[Sandy photographed __] and [Chris visited the castle]]?

     c. [Which dignitaries]$_i$ do you think [[Sandy photographed __] and [Chris visited __]]?

(40)   a.*[Which of her books]$_i$ did you find both [[a review of Gould] and [a reply to __]]?

     b.*[Which of her books]$_i$ did you find both [[a reply to __] and [a review of Gould's new book]]?

     c. [Which of her books]$_i$ did you read both [[a review of __] and [a reply to __]]?

These contrasts are all straightforwardly derived from the the identity constraint on coordinate structures.[8]

---

[8]There is a further issue raised by the observation that gaps cannot be conjuncts:

(i) *[Which of her books]$_i$ did you find both [[a review of __] and [__]]?

(ii) *[Which of her books]$_i$ did you find [[__] and [a review of __]]?

(iii) *[Which rock legend]$_i$ would it be ridiculous to compare [[__] and [__]]? (cf. [Which rock legend]$_i$ would it be ridiculous to compare __ with himself$_i$?)

For further discussion, see Bouma et al. 2001 and Sag 2000, who account for such examples by eliminating *wh*-traces from their constraint-based analysis of UDCs.

4 Conclusion

We conclude with the following observations:

1. The HPSG theory of filler/gap UDCs takes the putative 'true' and the alleged 'parasitic' gaps to be completely on a par with one another. Hence the Kearney paradigm facts are predicted, given the binding theory of Pollard and Sag 1994 and processing constraints that are independently motivated by examples like (5) and (7).

2. The well-formedness of nominative subject p-gaps corresponds to the HPSG null hypothesis, and hence nothing further needs to be said about it.

3. 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 of Kluender, Kroch and others). As noted, the Pollard-Sag Subject Condition, taken together with the 'adverbs as complements' analysis, predicts the well-formedness of the symbiotic gap examples we have discussed.

4. The case mismatch facts fall out simply and directly from the case type hierarchy presented in Levine et al. 2000. Nothing further needs to be said.

5. The Coordinate Structure Constraint and its 'across-the-board' exceptions also fall out directly from the independently motivated identity constraints on coordinate structures within the HPSG analysis of extraction. Movement-based alternatives have yet to be reconciled with these long-standing problematic data.

In short, none of the difficulties we have noted, which have been significant deficiencies in movement-based approaches to p-gaps throughout all the variants we have examined, ever arises in HPSG. The conclusion seems inevitable: on general methodological, as well as purely empirical grounds, HPSG provides a superior account of parasitic and, more generally, multiple gap constructions.

References

Bouma, G., Malouf, R., and Sag, I. A.: 2001, Satisfying constraints on extraction and adjunction, *NLLT* 19, 1–65

Chomsky, N.: 1982, *Concepts and Consequences of the Theory of Government and Binding*, Cambridge, MA: MIT Press

Chomsky, N.: 1986, *Barriers*, Cambridge, MA: MIT Press

Cinque, G.: 1990, *Types of Ā–dependencies*, Cambridge, MA: MIT Press

Frampton, J.: 1990, Parasitic gaps and the theory of *wh*-chains, *Linguistic Inquiry* 21, 49–77

Gazdar, G.: 1981, Unbounded dependencies and coordinate structure, *Linguistic Inquiry* 12, 155–84

Gazdar, G., Klein, E., Pullum, G. K., and Sag, I. A.: 1985, *Generalized Phrase Structure Grammar*, Oxford: Basil Blackwell's and Cambridge, MA: Harvard University Press

Gazdar, G., Pullum, G. K., Sag, I. A., and Wasow, T.: 1982, Coordination and Transformational Grammar, *Linguistic Inquiry* 13.4, 663–677

Haegeman, L.: 1984, Parasitic gaps and adverbial clauses, *Journal of Linguistics* 20, 229–232

Kearney, K.: 1983, *Governing categories*, Unpublished ms. University of Connecticut, Storrs, CT.

Kehler, A.: 2002, *Coherence, Reference and the Theory of Grammar*, Stanford: Center for the Study of Language and Information

Levine, R. D.: 2001, The extraction riddle: just what are we missing?, *Journal of Linguistics* 37, 145–174

Levine, R. D., Hukari, T. E., and Calcagno, M.: 2000, Some overlooked parasitic gap constructions in English and their theoretical implications, in P. Culicover and P. M. Postal (eds.), *Parasitic Gaps*, pp 181–222, Cambridge, MA: MIT Press

Levine, R. D. and Sag, I. A.: 2003, *WH*-nonmovement, *Gengo Kenkyu: Journal of the Linguistic Society of Japan* 123, 171–219

Longobardi, G.: 1984, Connectedness and island constraints, in J. Gueron, H.-G. Obernauer, and J.-Y. Pollock (eds.), *Grammatical Representation*, pp 161–185, Dordrecht: Foris

Manning, C. and Sag, I. A.: 1998, Argument structure, valence, and binding, *Nordic Journal of Linguistics* 21, 107–144

Pollard, C. J. and Sag, I. A.: 1992, Anaphors in English and the scope of binding theory, *Linguistic Inquiry* 23.2

Pollard, C. J. and Sag, I. A.: 1994, *Head–Driven Phrase Structure Grammar*, Chicago: University of Chicago Press and Stanford: CSLI Publications

Postal, P. M.: 1998, *Three Investigations of Extraction*, Cambridge, MA: MIT Press

Przepiórkowski, A.: 1999, *Case Assignment and the Complement-Adjunct Dichotomy: A non-Configurational Constraint-Based Approach*, Ph.D. thesis, Universit"at Tübingen

Reinhart, T. and Reuland, E.: 1993, Reflexivity, *Linguistic Inquiry* 24, 657–720

Ross, J. R.: 1967, *Constraints on Variables in Syntax*, Ph.D. thesis, MIT, Published in 1986 as *Infinite Syntax*!. Norwood, N. J.: Ablex

Sag, I. A.: 1997, English relative clause constructions, *Journal of Linguistics* 33, 431–484

Sag, I. A.: 2000, Another argument against *wh*-trace, *Jorge Hankamer Webfest:* `http://ling.ucsc.edu/Jorge/`

Sag, I. A., Gazdar, G., Wasow, T., and Weisler, S.: 1985, Coordination and how to distinguish categories, *Natural Language and Linguistic Theory* 3, 117–171