*Parasitic Gaps and Locality Theory: Some Results**

M. RITA MANZINI

Since being discussed in Taraldsen (1981), Chomsky (1982) and Engdahl (1983), parasitic gaps have consistently surfaced in the literature as evidence in the debate between various theories of locality. They are central to connectedness theories, of the type in Kayne (1983); and the debate remains open between these and barriers-based theories, of the type in Chomsky (1986). The purpose of this paper is to provide a preliminary discussion of parasitic gaps in terms of the Locality theory proposed in Manzini (1992).

Sections 1 and 2 introduce connectedness and barriers-based accounts and some of their problems. Section 3 introduces Locality theory. In section 4, I show that parasitic gaps are licenced only by argument A'-dependencies, or exactly the class of dependencies that displays less stringent locality behaviors than all others. In technical terms, parasitic gap dependencies must be address-based. In section 5, I argue that for head dependencies, the BC notion of barrier, in the sense of Chomsky (1986), is a theorem of the Minimality notion and of c-command. This result is relevant for address-based dependencies, since they characteristically consist of an operator and a variable linked by a sequence of heads. I argue that if the BC definition of barrier is abandoned as an independent construct, parasitic gaps can be accounted for by simply allowing for forking dependencies in the grammar. Finally in section 6, I present some problems that the results in the previous sections open.

*This paper represents a preliminary report of ongoing work on parasitic gaps and locality. I am grateful to M. Brody for discussing the relevant issues with me. I presented this and related material at the 1992 Incontro di Grammatica Generativa in Ferrara, the 1992 GLOW Conference in Lisbon, and as part of a cycle of lectures at the Université de Paris 8. I am grateful to A. Rouveret for inviting me to Paris and to the various audiences involved for discussion.

1 Connectedness Theories

The crucial construct of connectedness theories is the notion of g-projection. In (1)-(2) I modify Kayne's (1983) definition along lines suggested notably by Longobardi (1984; 1985), Bennis and Hoekstra (1984), Koster (1986). In essence the ordering constraint on g-projections in (2) is Kayne's (1983); the subsequent literature adds the government constraint proper, where government is understood as head-government. Given the notion of g-projection in (1)-(2), a g-projection set of a category α is defined as in (3), essentially as the set of all g-projections of ß, where ß is a governor for α. The Connectedness Condition is then formulated in terms of g-projection sets as in (4). Roughly speaking, given a set of empty categories $\alpha_1, ..., \alpha_n$ with antecedent ß in a tree T, under (4) ß together with the g-projection sets of $\alpha_1, ..., \alpha_n$ must constitute a subtree of T:

(1) Y is a g-projection of X iff
 a. Y is a projection of X or of a g-projection of X; or
 b. Y immediately dominates W and Z, where Z is a g-projection of X, and W canonically governs Z

(2) W canonically governs Z iff W governs Z and
 V governs NP to its right in the language and W precedes Z; or
 V governs NP to its left in the language and Z precedes W

(3) G_α is the g-projection set of a category α, where ß governs α iff
 a. all g-projection of ß belongs to G_α
 b. α belongs to G_α and
 b' if α' dominates α and and does not dominate ß, α' belongs to G_α

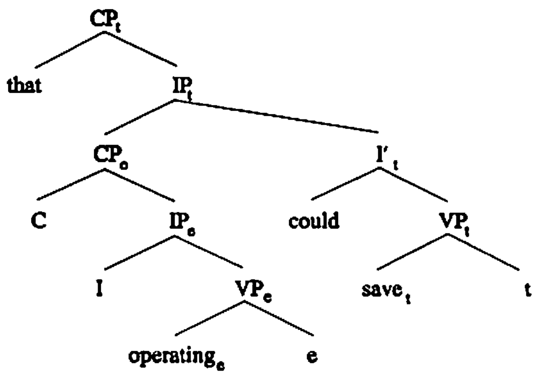
(4) *Connectedness Condition*
 Let $\alpha_1, ..., \alpha_n$ be a maximal set of empty categories in a tree T such that for some ß, all α_i are bound by ß.
 Then ß together with the g-projection sets of all α_i must constitute a subtree of T.

Kayne (1983) argues that Connectedness correctly predicts the contrast between examples involving subject islands of the type in (5a) and (5b); I equivalently will consider the contrast between (6a) and (6b). The structure for (6a) is as in (7), where the nodes bearing an *e* subscript belong to the g-projection set of the parasitic gap *e*, the nodes bearing a *t* subscript to the g-projection set of the main gap *t*. In (7) the g-projection set of *e* does not by itself form a subtree with the higher Spec of CP; however, the g-projection set of *t* does, and so does the union of the *e* and *t* g-projection sets. By contrast, consider (8), the relevant structure for (6b). In (8) the g-projection set of *e* stops at the embedded CP, as in (7), but contrary to (7), it forms a subtree distinct from the subtree formed by the g-projection set of *t*. Thus intuitively in (8), contrary to (7), *t* cannot rescue *e* from its island violation:

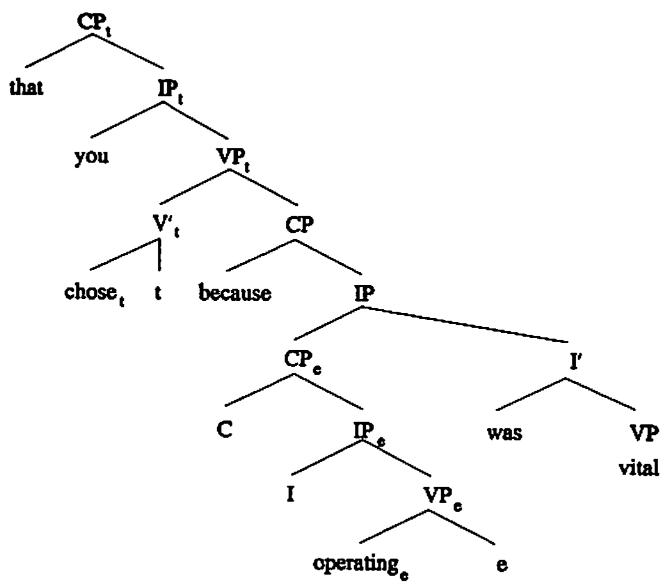
(5) (a) A man that [friends of e] admire t
(b) *A man that you admire t [because [friends of e] become famous]

(6) (a) A patient that [operating e immediately] could save t
(b) *A patient that you chose t [because [operating e immediately] was vital]

(7)



(8)

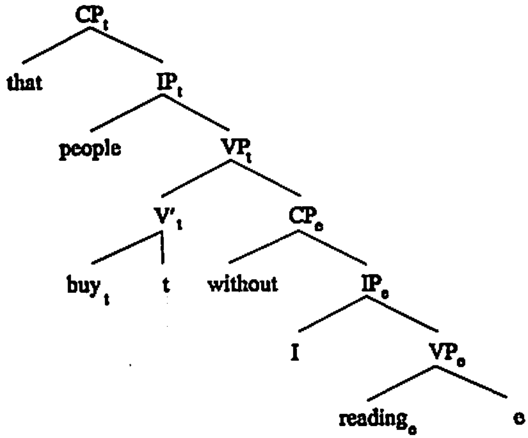


Notice that in order to predict the contrast in (7) vs. (8) it is in fact sufficient, as in Kayne (1983), to define g-projections in terms of canonical ordering. Thus the g-projection of *e* stops at the subject CP in both (7) and (8) because this is on a left branch, and English is a right-branching language. The reason for complicating Kayne's (1983) definition of g-projection to include a government requirement is found in contrasts of the type in (9)-(10). In the structure for (9) provided in (11) the g-projection of *e* does not extend beyond the adjunct CP, because this is not governed by a head; but the presence of the g-projection of *t* allows for the satisfaction of Connectedness. Crucially, the illformedness of (10) is also predicted under the structure in (12). In (12) the g-projection of *e* stops at the adjunct CP again; but the g-projection of *t* is not sufficient to rescue it, because the two do not form a subtree. Hence Connectedness is violated. In terms of canonical ordering, on the other hand, the g-projection of *e* extends beyond the adjunct CP in (11) in that this is on a canonical right branch; but for the same reason it extends beyond both of the adjunct CP's in (12), incorrectly predicting wellformedness:

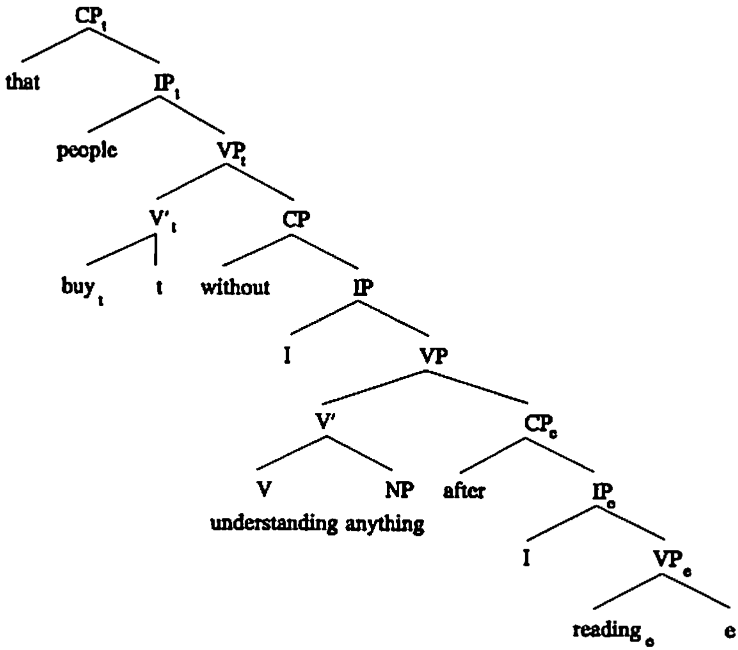
(9) A book that people buy t [without reading e]

(10) *A book that people buy t [without understanding anything [after reading e]]

(11)



(12)



Thus summing up so far, the theory in (1)-(4) is adequate to account for all of the data in (5)-(12). On the other hand notice that (1)-(4) represents at best half of locality theory. In particular locality phenomena depending not on tree geometry, such as subject and adjunct islands, as in (8) and (12) respectively, but on antecedent intervention, such as wh-islands, as in (13), or SSC effects, as in (14), or HMC effects, as in (15), appear to fall entirely outside the scope of Connectedness. Thus assuming the empirical adequacy of Connectedness, the problem of how it integrates with whatever subtheory predicts (13)-(15) still remains open:

(13) *How do you wonder [what to repair t t]
(14) *Mary seems that it is likely [t to leave]
(15) *Have Mary will [t finished by tonight]

2 Barriers Theories

Consider Chomsky's (1986) theory. This is largely motivated by the attempt at defining a (partially) unified theory for all locality phenomena. According to Chomsky (1986), a maximal projection which dominates a given category can be a barrier for it, inherently or by inheritance. It is a barrier inherently if it is not L-marked, i.e. a BC; it is a barrier by inheritance if it is the first maximal projection that dominates a BC. IP however has an exceptional status. It can be a barrier by inheritance, and barrierhood can be inherited from it; but it cannot be an inherent barrier. The relevant definitions read as in (16)-(17). Under Subjacency no category can move across more than one barrier, in other words no more than one barrier for the position moved from can exclude the position moved into, as in (18). The interplay of dominance and exclusion in (16)-(18) has the effect of making any adjunction to a maximal projection into an escape hatch. However adjunction to two maximal projections, CP and NP, is blocked by a constraint against adjunction to an argument; furthermore IP has exceptional status also with respect to adjunctions, since adjunctions to it are not possible, as in (19):

(16) ß is a B(locking) C(ategory) for α iff
ß is an XP, ß dominates α and ß is not L-marked

(17) ß is a barrier for α iff
(i) ß (other than IP) is a BC for α, or
(ii)ß is the first XP that dominates a BC for α

(18) *Subjacency*
If α is a trace, there is antecedent ß for α such that at most one barrier for α excludes ß

(19) Arguments and IP's cannot be adjoined to

The theory in (16)-(19) is sufficient to account for the ungrammatical occurences of parasitic gaps. Consider for instance (6b) or (10). Under Chomsky's (1986) theory they are associated with partial representations of the type in (20) and (21) respectively, where we assume that movement has successfully taken place from the parasitic gap position *e* to the embedded Spec of CP, *e'*:

(20) [$_{CP}$ e'' because [$_{IP}$ [$_{CP}$ e' operating e] [$_{I'}$ was vital]]]

(21) [$_{CP}$ e'' without [$_{IP}$ [$_{I'}$ understanding anything] [$_{CP}$ e' after reading e]]]

In (20) adjunction to both CP and the immediately superordinate IP is impossible by (19); thus the first possible landing site from *e'* is the Spec of the higher CP, *e"*, as indicated. This means that two barriers are crossed, namely the embedded CP, which is a BC in that it is not L-marked, and IP which is a barrier by inheritance; hence Subjacency is violated. Consider then (21), where under Chomsky's (1986) theory the adjunct CP must be attached under IP. In (21) the lower CP is a BC and hence a barrier for *e'*, and so is IP by inheritance. Since neither CP not IP can be adjoined to, movement from *e'* to the next possible landing site in the Spec of the higher CP, *e"*, crosses at least two barriers, namely the embedded CP and IP; hence Subjacency is violated.

As I have anticipated, Chomsky's (1986) theory provides a (partially) unified account of subject and adjunct islands, as in (20)-(21), and of phenomena such as wh-islands, SSC effects, and HMC effects, as in (13)-(15). In particular, government is defined to hold between a trace and its antecedent just in case no barriers intervene between them, as in (22); the antecedent-government clause of the ECP then requires all non-theta-govered traces and all A-traces to be antecedent-governed, as in (23). Thus Subjacency and the antecedent government clause of the ECP differ only in that an ECP violation is defined in terms of just one barrier crossed, while a Subjacency violation is defined in terms of two or more barriers crossed:

(22) ß governs α iff there is no barrier for α that excludes ß

(23) *ECP: Antecedent-Government Clause*
If α is a non-theta-governed trace or an A-trace, there is an antecedent ß for α such that ß governs α

Objects, which are theta-governed, need not satisfy the antecedent government clause of the ECP; correspondingly wh-island effects are not found, or found in a weakened form, when an argument is extracted across a wh-island, as in (24). Adjuncts, on the other hand, that do not satisfy theta-government, give rise to wh-island violations of the type in (13). The relevant structure for such violations is provided in (25). t_i in (25) represents the base-generated position of the adjunct. The embedded IP is of course a BC since it is not L-marked; the embedded CP is then a barrier for t_i by inheritance. Thus in order to reach the next possible landing site, the VP-adjoined position t'_i, the CP barrier must be crossed, which produces a violation of the antecedent-government clause of the ECP:

(24) What do you wonder [how to repair t t]
(25) *How$_i$ do you [$_{VP}$ t'$_i$ wonder [$_{CP}$ what$_j$ [$_{IP}$ to repair t$_j$ t$_i$]

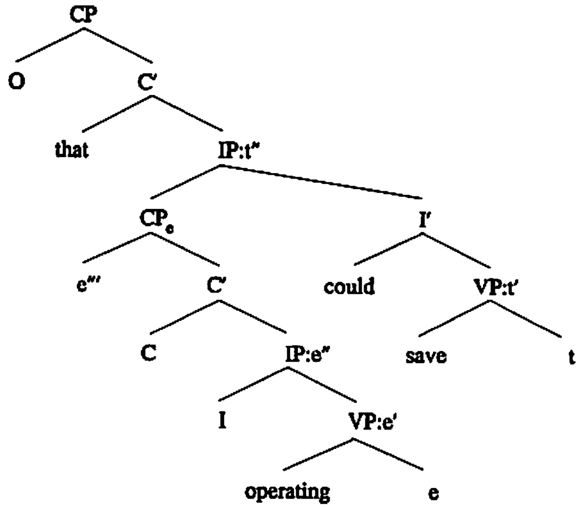
Unfortunately, Chomsky's (1986) theory has serious empirical problems with parasitic gap examples. The fundamental intuition that it seeks to capture is that examples like (6a) or (9) are wellformed in that movement can take place from the parasitic gap to an empty operator position governed by, or subjacent to the main gap (chain). Consider then the structures associated with (6a) and (9), as in (26) and (27). Quite simply, in (26) adjunction of O$_e$ to the subject CP would need to take place in order for government/ subjacency to hold with any member of the t-chain; but if so, the basis for predicting the impossibility of extractions from subject islands would disappear altogether. Similarly, adjunction of O$_e$ to the adjunct CP would need to take place in (27); but this would nullify all predictions concerning adjunct islands:

(26) [$_{CP}$ O$_t$ that [$_{IP}$ [$_{CP}$ O$_e$ operating e] [$_{I'}$ would save t]]]
(27) [$_{CP}$ O$_t$ that [$_{IP}$ people [$_{I'}$ buy t] [$_{CP}$ O$_e$ without reading e]]]

Now, an empirically adequate theory of parasitic gaps within Chomsky's (1986) general framework is presented in Frampton (1990). In Frampton (1990), (16)-(18) hold unaltered, but (19) is crucially modified to allow for adjunction to IP, as in (28). At the same time the theory of adjunctions is strengthend by the introduction of the Head Government Condition on Adjunction in (29):

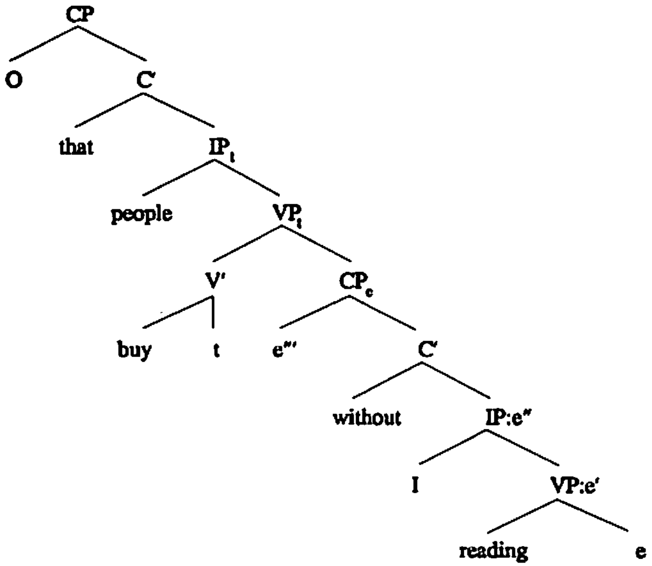
(28) Arguments cannot be adjoined to
(29) *Head Government Condition on Adjunction (HGCA)*
A wh-element can only be adjoined to a maximal projection XP from a position that is canonically governed by the head of XP

Consider first (6a). This is associated with a structure of the type in (30), where adjunction of t to XP is notated XP:t. *e'''* in the Spec of the embedded CP is now subjacent to the IP-adjoined *e''*, since the link between the two only crosses one barrier, the embedded CP itself. Subjacency is then sufficient to predict the wellformedness of (30). Notice on the other hand that adjunction of *t''* to IP in (30) is licenced under the HGCA in that the position moved from, *t'*, is canonically governed by the head of I. Crucially, then, movement from *e'''* directly to *t''* is impossible, precisely because the head of IP, the matrix I, does not canonically govern *e'''*. This means that subject islands are derived by the theory, and indeed the illformedness of (6b) is correctly predicted:



Consider next the relevant structure for (9), as in (31). Adjunction of *t'* to the higher VP is licenced by the fact that the head of VP, V, canonically governs the position moved from, *t*. *e'''* in the embedded Spec of VP is then subjacent to *t'*; indeed only one barrier, CP itself, intervenes between the two. Hence the parasitic gap is once more licenced. However movement from *e'''* directly to the VP adjoined position *t'* is blocked by the HGCA; this is the basis for predicting adjunct island violations, and the illformedness of (10) as a subcase of it:

(31)



Consider however wh-islands again, as in (13). Under Frampton's (1990) theory (13) is associated with the partial derivation in (32). Under the HGCA movement can take place from t_i to the IP-adjoined position t'_i and from there to the VP-adjoined position t''_i. Indeed it must be assumed that a V can govern an IP-adjoined position across an empty C head, in order to allow for wellformed extractions of arguments across wh-islands, as in (24). But at this point the theory of antecedent-government in (22)-(23) no longer derives wh-islands at all, since movement from t'_i to t''_i in (32) crosses no barriers. Thus the (partial) unification of Subjacency and antecedent-government achieved in Chomsky (1986) must be abandoned:

(32) How$_i$ do you $[_{VP}$ t''$_i$ $[_{VP}$ wonder $[_{CP}$ what $[_{IP}$ t'$_i$ $[_{IP}$ to repair t t$_i$]]]]]

In short, there is an apparent trade-off between (partial) unification of the theory of locality, of the type in Chomsky (1986), which leads to apparently unsurmountable difficulties with parasitic gaps, and empirically adequate theories of parasitic gaps, including (a suitably modified version of) Kayne's (1983) and Frampton's (1990), which are not integrated into any comparably unified theory of locality.

3 Locality Theory

In Manzini (1988; 1990; 1992), I argue that all locality principles can be unified into one. This is essentially the antecedent-government clause of Chomsky's (1986) ECP, as in (33). In (33), government is understood as in Chomsky (1986) as lack of any intervening barriers; furthermore government is taken to be satisfied by a dependency just in case it holds of each of its links:

(33) *Locality*
If A is a trace there is an antecedent B for A and a dependency (B, ..., A) that satisfies government

In Chomsky (1986), in turn, government is defined in terms of two notions of barrier. In particular, by Minimality an X' (but not I') is a barrier for everything it dominates except its head, X. I also assume Minimality, in the version in (34) under which XP is a barrier for everythig it dominates except its head and its Spec; essentially the same definition is arrived at in Sportiche (1989):

(34) B is a barrier for A if B is a maximal projection and B dominates A, unless A is the head or Spec of B

Otherwise, a maximal projection XP that dominates a given category can be a barrier for it inherently or by inheritance. XP (other than IP) is a barrier inherently if it is not L-marked, i.e. it is a BC; XP is a barrier by inheritance if it is the first maximal projection that dominates a BC. Of course, if all locality conditions are reduced to the government principle in (33) and no more than one barrier at a time is ever relevant, the inheritance clause can be eliminated. If the exception concerning IP is also eliminated, the second definition of barrier reduces to that of BC, as in (35):

(35) B is a barrier for A if B is a maximal projection, B dominates A and B is not L-marked

Let us then test the theory in (33)-(35) against the standard contrast between the extraction of adjuncts and of arguments out of wh-islands. Consider first adjunct extraction, as in (13). Suppose we adopt a simple theory of adjunctions under which nothing prevents adjunction to IP, though adjunction to arguments remains impossible. The relevant derivation for (13) is as in (36), where movement from t'_i to t''_i obviously violates government. Indeed CP is barrier for t'_i under (34) simply because the latter is neither its head nor its Spec; but CP excludes t''_i. This means that (33) is not satisfied, and (13) is predicted to be ungrammatical, as desired:

(36) How$_i$ do you [t''_i [wonder [what [t'_i [to repair t t_i]]]]]

Consider argument extraction across a wh-island, as in (24). Apparently, a wellformed dependency cannot ultimately be formed for the same reason as in (13). The relevant derivation is as in (37). In (37) movement from t'_i to t''_i crosses a CP barrier exactly as in (36), thus violating government and ultimately (33). Under this derivation, then, (24) is predicted to be ungrammatical, and the contrast between (13) and (24) remains unexplained:

(37) What$_i$ do you [$_{VP}$ t''_i [$_{VP}$ wonder [$_{CP}$ how [$_{IP}$ t'_i [$_{IP}$ to repair t_i t]]]]]

In Manzini (1988; 1990; 1992), on the other hand, I impute the grammaticality of (24) to the existence of an alternative derivation open to it. Under this derivation, the local links between an operator and its variable are supplied by heads, producing a structure of the type in (38), where elements in the relevant dependency are shown in **boldface**. In (38) the dependency crucially includes the link (*wonder*, C). This satisfies government across the wh-island, in that the only maximal projection that dominates C and excludes *wonder* is CP; and CP is not a barrier for its head C. Thus there is a derivation for (24) that satisfies (33), as desired:

(38) **What$_i$ do** [you **I** [**wonder** [how **C** [**to** [**repair t$_i$** t]]]]]

The contrast between (13) and (24) can then be imputed to the unavailability of a derivation of the type in (38) for (13) and similar examples. The reason, quite naturally, appears to be that dependencies of the type in (38) are licenced only by a privileged relation of the empty category to a head, obviously the V head *repair* in the case of (38).

Following Manzini (1992), we can formalize the type of dependency seen in (38) and the conditions that licence it in the following way. Assume that there are two types of indices in the grammar, categorial indices and addresses, the latter term borrowed from Vergnaud (1985). On the one hand, every lexical head is assigned a categorial index, which in the case of DP's can be identified with a referential index, in the standard sense of the term. On the other hand, a position made visible by a head is assigned an address; this, we can assume, consists of a pair of indices, corresponding to the categorial index of the head and of the visible position respectively. Formally, then, we will say that a categorial index and address are associated with A when the conditions in (39) and (40) respectively are satisfied:

(39) i is licenced as the categorial index of A iff A is lexical
(40) (j, i) is licenced as the address of A iff
$A = A_i$, and there is B, $B = B_j$, such that A is made visible by B

Suppose next that since dependencies can standardly be formed on the basis of categorial indices, and since addresses are just another type of index, dependencies are definable on the basis of coaddressing as well. A dependency, subsuming both categorial index and address-based ones, can be defined simply as a set of coindexed positions ordered by c-command, as in (41):

(41) $(A_1, ..., A_n)$ is a dependency iff
for all i, A_i c-commands and is coindexed with A_{i+1}

Finally, assume that apart from movement, there is a second fundamental mechanism through which coindexing takes place, namely percolation. Percolation is a standard device of all current theories, to the extent that it is accepted that indices percolate upwards and downwards from heads to maximal projections and viceversa. Generalizing this mechanism, an index can be assumed to percolate freely till it is blocked by an incompatible index.

Coindexing by percolation is in principle available for the formation of categorial index dependencies, but in practice always blocked. Typically in (37) the V' projection of the embedded V must bear the categorial index of V; thus the categorial index *i* of t_i cannot even percolate to V'. The same holds of the VP projection of V, thus blocking the percolation of the categorial index *i* of t_i in (36), if the adjunct is attached under VP; and so on. Consider however coaddressing. None of the maximal projections that dominate t_i in (38) has an address of its own, since none of them is visible in terms of Chomsky's (1981) theory, including CP. Thus the address of t_i can percolate to VP, IP, CP recursively and from there to their heads, V, I and C, so that all and only the terminal elements shown in **boldface** in (38) are in fact coaddressed.

The similarity of address-based dependencies to other non-chain derivations of the type in Kayne (1983), Gazdar et al. (1983), or more recently Hoekstra (1988), is at this point evident. However, one obvious alternative formalization is also suggested by Chomsky's (1986) notion of CHAIN. Under current assumptions, a V will typically enter a head-to-head chain with its associated functional categories, say I and C. It is well-known furthermore, for instance from studies of the subjunctive, that I itself can be dependent on a superordinate I, leading to the formation of extended CHAINs (C, I, V)*. But if an operator and a variable are linked to one another via such a CHAIN of heads, then the resulting CHAIN (O, (C, I, V)*, e) overlaps with an address-based dependency in the sense of (38). The contrast between (13) and (24) can then be predicted in terms of the availability or not of a CHAIN derivation; in particular a final link (V, e) in a CHAIN will be licenced just in case the head has a (visibility) relation to the empty category. For the time being I will disregard these issues, noting however that they may well have empirical, rather than purely notational import and therefore deserve careful investigation.

4 Parasitic Gaps Dependencies are Address-Based

Consider once more the crucial examples of parasitic gaps in (6) and (9)-(10). In (6a) a parasitic gap inside a subject island is wellformed; while in (6b) a parasitic gap inside a subject and an adjunct island is illformed. Similarly in (9) a parasitic gap inside an adjunct is wellformed; however in (10) a parasitic gap inside two adjuncts is illformed.

In all of the examples in (6) and (9)-(10) movement can in principle proceed through the formation of either a categorial index or an address-based dependencies, since an argument is involved. Furthermore, at least one derivation must be wellformed in (6a) and (9) since the examples are wellformed. The obvious question is whether it is the address-based derivation, or the categorial index derivation, or both that are wellformed. Now, notice that in all cases in which address-based dependencies are not licenced, and only categorial dependencies are, parasitic gaps are not licenced either. In particular, we have seen that under our definition of addressing, adjuncts are not addressed. Indeed the fact that they cannot form address-based dependencies explains their sensitivity to wh-islands, as in (13). In the same way, adjuncts cannot licence parasitic gaps, as in (42):

(42) *The way O he chose his tools t [before repairing the car e]

In the grammar that I propose, two other major classes of traces pattern with A'-traces of adjuncts in not being addressed, namely A-traces and traces of head-movement. Indeed the fact that they cannot form address-based dependencies explains their locality behavior under my theory, as in (14) and (15) respectively. Consider first (14). If the trace is not addressed, it can only form a categorial index dependency, which can include the most embedded and the matrix subject positions, but not the intermediate subject position, which is independently filled, as in (43). If so, a number of barriers are crossed in the intermediate clause, for instance the intermediate CP, and the ungrammaticality of (14) is predicted, as desired:

(43) Mary$_i$ seems [$_{CP}$ that it is likely [$_{IP}$ t$_i$ to leave]]

Similarly, consider (15). Again, the V head moved in (15) is not addressed. It follows that movement can proceed only via categorial coindexing from the V position to the C position, and across the I position, which is independently filled, as in (44). If so, a barrier is crossed, namely IP, and the ungrammaticality of (15) is predicted, as desired:

(44) [$_{CP}$ Have$_i$ [$_{IP}$ Mary will [$_{VP}$ t$_i$ finished by tonight]

Crucially, A-movement and head-movement also pattern with A'-movement of adjuncts in not licencing parasitic gaps, as exemplified for A-movement in (45) with passive and for head-movement in (46) with (Italian) cliticization. Thus the parallelism between impossibility of address-based derivations and impossibility of parasitic gaps appears to hold generally. The conclusion then emerges that parasitic gap dependencies cannot be categorial-index based; hence they must be address-based:

(45) *Mary has been rejected t without seeing e
(46) *Gianni l'ha respinta t senza vedere e
John her+has rejected without seeing

The question of course is why such a conclusion should hold. Notice however that no matter what the answer turns out to be, the formulation of this question already sets my theory apart from others. In other theories, (45)-(46) are imputed to the fact that parasitic gaps are licenced by A'-movement, but not by A-movement; (42) is imputed to the fact that parasitic gaps are pro's at some level of representation, as originally proposed in Chomsky (1982) and more recently in Cinque (1991). To these partial generalizations I tentatively substitute a single one, in terms of address-based vs. categorial index dependencies.

In more factual terms, the generalization is that all and only those elements that can undergo long movement in the sense of Rizzi (1990) or Cinque (1991), are able to licence parasitic gaps. We thus expect 'non-referential' wh-phrases, in the sense of Rizzi (1990) or Cinque (1991) not to be able to licence parasitic gaps either. This appears to be true in the case, say, of measure phrases, though not in that of *how many* phrases in general, as in (47)-(48):

(47) *How many kilos does she weigh t without admitting to weigh e
(48) How many books has she reviewed t without reading e

Since both types of phrases are argued to be non-referential, or non-referential under their normal reading, in Rizzi (1990) and Cinque (1991) respectively, the contrast in (47)-(48) could be construed as problematic for the factual generalization I am proposing.

In reality, the only data concerning *how many* phrases of the type in (48) found in Cinque (1991) concern their interaction with other quantifiers; in the absence of such interactions, no 'non-referential' behavior is apparent, as exemplified in the contrast between (49) and (50), both involving inner islands in the sense of Rizzi (1990):

(49) *How many kilos doesn't she weigh
(50) How many books hasn't she read

Thus since no interaction with wh- or other operators is involved in the parasitic gaps examples, we expect *how many* phrases in general to be able to licence them. In other words, the factual generalization still holds.

If this generalization about the distribution of parasitic gaps is correct, another interesting prediction is made, potentially separating my theory from theories of parasitic gaps as pro's. This concerns PP's. Taking Italian, a non P-stranding language for ease of exemplification, it is clear that (selected) PP's can be extracted from weak-islands. If so, they are expected to be able to licence parasitic gaps under my generalization, while Cinque (1991) explicitly excludes them from the range of categories that can be realized as pro. In my own intuition, not only simple dative PP's, but also complex locative PP's display the relevant contrast between direct extraction and parasitic gap inside an island, as in (51)-(54). Thus the generalization I am proposing appears to be again supported by the data:

(51) *A che medico sei stata curata senza dare un soldo
To which doctor have you been cured without giving a penny
(52) A che medico to sei rivolta senza dare un soldo
To which doctor did you turn without giving a penny
(53) *Davanti a che altare sei andato in chiesa senza mettere una candela
Before which altar did you go to church without putting a candle
(54) Davanti a che altare hai pregato senza mettere una candela
Before which altar did you pray without putting a candle

Finally, the discussion that precedes bears on a treatment of parasitic gaps not considered so far, under which parasitic gaps are related to ATB extractions from coordinate structures, as in Ross (1967), Pesetsky (1982), Huybreghts and van Riemsdijk (1984), Haik (1985), Williams (1986). While A'-movement of adjuncts, A-movement and head-movement do not licence parasitic gaps, they licence ATB extractions, as in (55)-(57). Thus ATB extractions appear to genuinely differ from parasitic gaps in being licenced under a categorial index derivation:

(55) The way O_i I selected my tools t_i and I repaired the car t_i
(56) $Mary_i$ seems t_i to have come, but does not seem t_i to have left
(57) $Will_i$ Mary t_i leave and Peter t_i come

Another major difference between parasitic gaps and coordinate structures is that A'-movement of both adjuncts and arguments is not licenced unless it takes place from both conjuncts in a coordinate construction; while in parasitic gap constructions the parasitic gap can alternate with an overt pronoun. This is illustrated with argument A'-movement in (58)-(60):

(58) The book that Mary bought t and Peter read t
(59) *The book that Mary bought t and Peter read it
(60) The book that Mary bought t without reading it

In fact, however, coordinate structures parallel parasitic gaps when they involve A-movement or head-movement; A-movement and head-movement are free to take place in only one of the conjuncts, as illustrated in (61)-(62):

(61) Mary seems t to have lost, and is sad

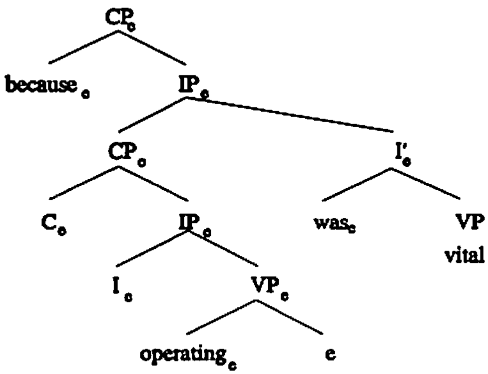
(62) Maria vuole salutarci t e evitare Lucia
Mary wants to greet+us and to avoid Lucy

In essence, assuming that operators such as *and* impose a parallelism requirement on the elements they conjoin, this requirement appears to involve that neither conjunct, or alternatively both, are introduced by a (wh-)operator, as in (58)-(59). This seems to be at the root of the ATB phenomenon. The fact that ATB extractions can involve both adjuncts and arguments, as opposed to parasitic gap extractions, is then the one difference between the two that remains to be explained.

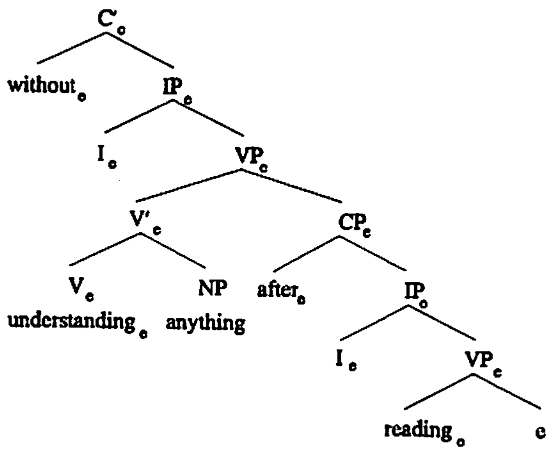
5 Parasitic Gaps: Address-based Derivations

So far, I have shown that in Locality theory terms, parasitic gap dependencies must be address-based. If this conclusion is correct, the question of course arises why it must hold. For the time being, however I will diregard this question and I will proceed to consider the relevant, address-based derivations for the crucial examples of parasitic gaps in (6) and (9)-(10). The structures for the illformed subject and adjunct island violations embedded in (6b) and (10) are as in (63) and (64) respectively, where I have indicated elements coaddressed with *e* simply by a subscripted *e* index:

(63)



(64)



The violations in (63)-(64) are in fact correctly predicted by my theory. Consider the embedded C in (63). Under the definition of barrier in (35), there is no c-commanding and coaddressed position in the tree that governs it, since CP is not L-marked and therefore is a barrier for all positions it dominates. Similarly, consider the embedded C, *after*, in (64). Again the embedded CP is a barrier for it under the definition in (35), since it is not L-marked. This means that there is no element in the tree in (64) that can both form a dependency with and govern *after*, and Locality is violated.

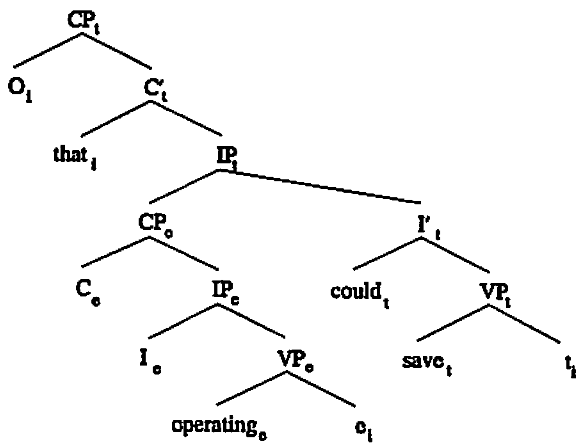
Now, while the notion of barrier in (35) is indeed sufficient to predict the ungrammaticality of (63)-(64), it is by no means necessary. Consider (63) again. By Minimality, as in (34), CP is not a barrier for its head C. By the percolation mechanism that we have postulated, furthermore, the address of *e* can percolate up to the matrix IP and from there to its head I. Since the only maximal projection that intervenes between C and I is CP and CP is not a barrier for C, I governs C. However, under a strict notion of c-command, as in Reinhart (1976), I does not c-command C (nor viceversa), hence no (I, C) link of a dependency can be formed. The next element in the tree that c-commands the embedded C is the matrix C, *because*. But under (34) once more, the matrix IP is a barrier for the embedded C, since the latter is neither its head nor one of its daughters. Thus *because* does not govern the embedded C and (63) is ultimately predicted to be illformed.

Very much the same can be repeated for (64). In (64) the embedded CP is not a barrier for its head C under (34). Thus C is governed by the matrix V, which can furthemore be coaddressed with it by percolation. But this is irrelevant since dependency formation is blocked by lack of c-command. The next element in the tree that c-commands the embedded C and is coaddressed with it is the matrix I; but the matrix I does not govern the embedded C in that VP intervenes between them, and VP is of course a barrier for C under (34), in that C is neither the head nor a daughter of VP. Thus the illformedness of (64) follows.

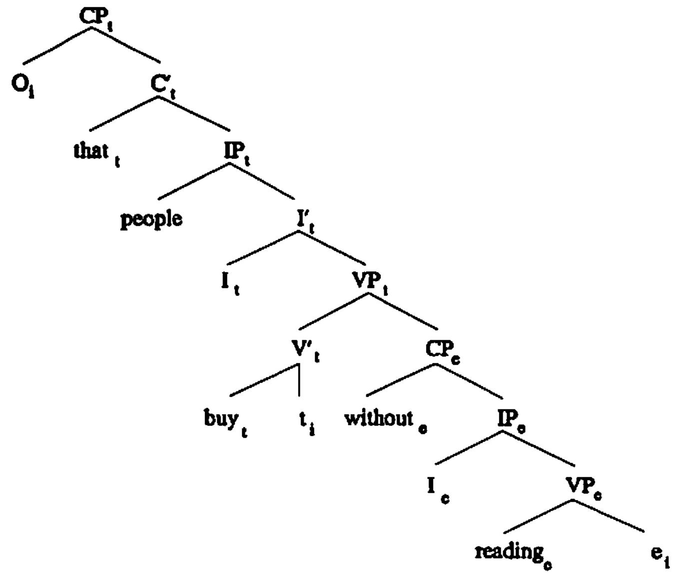
In short, the Minimality notion of barrier in (34) is sufficient to predict subject and adjunct island violations with address-based dependencies. In other words, for address-based dependencies the BC notion of barrier in (35) is a theorem of Minimality and c-command. Since of course the elimination of the definition of barrier in (35) also leaves the derivation of wellformed examples involving extraction out of wh-islands unchanged, as in (13), we can conclude that at least for address-based dependencies the definition of barrier reduces to Minimality, as in (34).

At this point, the crucial wellformed examples of parasitic gaps in (6a) and (9) can be taken into consideration. Under the address-based derivation, these are associated with structures of the type in (65) and (66) respectively. Following the same notation as before, an *e* subscript indicates the elements coaddressed with the parasitic gap *e*, and a *t* subscript the elements coaddressed with the main gap *t*:

(65)


CP_t
O_i
C'_t
$that_i$
IP_t
CP_e
I'_t
C_e
IP_e
$could_t$
VP_t
I_e
VP_e
$save_t$
t_i
$operating_e$
e_i

(66)



Consider first a grammar that includes both the definitions of barrier in (34) and in (35). In (65) the t dependency (O, *that*, *could*, *save*, t) obviously crosses no barriers of either type, and thus satisfies government. As for *e*, it can form a dependency that satisfies government as far as the embedded C. But a link between the embedded C and the matrix I violates government, since the embedded CP is a barrier by (35). Similarly in (66), the t dependency (0, *that*, I, *like*, t) obviously satisfies government. But the e dependency can satisfy government only as far as the embedded C, because the embedded CP is a barrier for everything it dominates under (35). Thus under (35), (65)-(66) are predicted to be just as illformed as (63)-(64).

Consider then a grammar that includes only the definition of barrier in (34). In (65) the t dependency (O, *that*, *could*, *save*, t) obviously satisfies government. As for *e*, it can form a dependency that satisfies government as far as the embedded C. But a link between the embedded C and the matrix I violates c-command, though not government; while a link between the embedded C and the matrix C, *that*, satisfies c-command, but not government. If the central question in parasitic gaps is what enables the t dependency to rescue the e dependency in a configuration like (65), the obvious observation is that while the matrix I cannot be part of the e dependency because of the lack of c-command, it is already part of the t dependency. Thus suppose that the e and t dependencies can be composed. If so, the embedded C is actually governed in the resulting dependency, by the matrix I, and Locality is satisfied.

Very much the same can be repeated for (66). The t dependency (O, *that*, I, *buy*, t) satisfies government. The e dependency cannot, because there is no element in the tree that both c-commands and governs the embedded C, *without*. However there is an element in the t dependency that governs C, namely the matrix V, *buy*. If the t and e dependencies are composed, then, Locality is satisfied for all members of the resulting dependency, as desired.

The observations that precede can be straightforwardly embedded within my theory by defining some mechanism to compose dependencies. In particular, we can say that if two dependencies have the same head, their union is also a dependency, as in (67). Concretely, the union of the t and e dependencies in (65)-(66) is also a dependency, since *t* and *e* are bound by the same operator O_i:

(67) If $A = (\alpha_1, ..., \alpha_n)$ and $B = (\beta_1, ..., \beta_m)$ are dependencies, and $\alpha_1 = \beta_1$, then the union of A and B is a dependency

In turn, the union under (67) of the e and t dependencies in (65)-(66) satisfies government, if satisfaction of government is technically defined as in (68). If so in particular, government is satisfied at the crucial juncture between the embedded C and the matrix C in (65) or the embedded C and the matrix V in (66). On the other hand, under (67)-(68) the derivation of the ungrammaticality of (63)-(64) remains unchanged. Indeed no member of the t dependency is present in the subtrees in (63)-(64), thus making a rescue of the e dependency impossible:

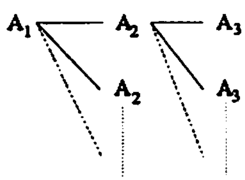
(68) A dependency $(A_1, ..., A_n)$ satisfies government iff for all $i > 1$, for some j, A_i is governed by A_j

In summary, at least as far as address-based derivations are concerned, a version of Locality theory that includes only the Minimality definition of barrier in (34) not only explains the ungrammaticality of extractions from subject and adjunct islands, but also predicts the grammaticality of parasitic gaps within the same islands, provided the chain-composition mechanism in (67) and the definition of government in (68) are adopted. By contrast, a version of Locality theory that includes both of the definitions of barrier in (34) and (35) is redundant with respect to extractions from subject and adjunct islands and does not predict any grammaticality contrast with parasitic gaps within the same islands. Thus we appear to be in the ideal situation where simplicity and empirical arguments agree in choosing one theory over another.

On the other hand the sequence composition mechanism in (67) could itself be objected to on grounds of simplicity. Notice however that a theory under which dependencies are defined in terms of c-command and coindexing as in (41) and the composition od dependencies is allowed on the basis of the identity of the heads of the dependencies, as in (67), is empirically equivalent to a system where dependencies are defined so as to be allowed to fork. Such a definition is provided in Brody (1992) and is reproduced here for present purposes in (69):

(69)

C =



is a dependency iff for each element A_i, i not=1, in C there is an A_{i-1} such that A_{i-1} c-commands and is coindexed with A_i

Brody (1992) also provides definitions for the notions of head and root of a sequence, given that forking is allowed as in (69). These definitions read as in (70) and (71):

(70) A_i is the head of the dependency iff there is no A_j in C such that A_j c-commands A_i

(71) A_i is the root of the dependency iff there is no A_j in C such that A_i c-commands A_j

Of course every dependency will have just one head, but it will be a property of a forking dependency that it will have more than one root. In the examples considered so far there are in fact two roots, corresponding to the real and the parasitic gap.

6 Parasitic Gaps: Categorial Index Derivations

As I have argued, parasitic gap dependencies must be address-based, and cannot be categorial index based. Thus, as a matter of fact, parasitic gaps are licenced by A'-movement of arguments, but not by any other kind of movement. However the question why this is the case has been left open in section 4. Similarly, in section 5 I have argued that the BC notion of barrier can be abandoned for address-based dependencies, and in fact it must be if parasitic gaps are to be accounted for. But this leaves the question open whether the elimination of (35) in favor of (34) is possible with categorial index dependencies as well. As I will show, the two questions turn out to be related.

Consider the canonical examples of parasitic gaps in (6) and (9)-(10) once again. Since arguments can form categorial index dependencies as well as address-based ones, a categorial index derivation is potentially available for (6) and (9)-(10). In the case of of (6b) and (10), we know that this derivation must be blocked, because the sentences are ungrammatical. Consider then the categorial index derivation for (6b), as in (72). Assuming that movement from t to t' in the Spec of the embedded CP is legitimate, the next landing site for movement is t" in IP-adjoined position. From there movement takes place directly to the Spec of the matrix CP. The crucial question then is whether the theory excludes movement from t' to t". The answer is positive under the definition of barrier in (35). Indeed the embedded CP is a barrier for t' in that it is not L-marked. Since in turn the embedded CP excludes t", Locality is violated. Consider however the Minimality notion of barrier in (34). Under it, it is obvious that t" governs t'. Indeed the only maximal projection that intervenes between them, CP, is not a barrier for t' under Minimality, because t' is the Spec of CP. Hence (72) is ultimately predicted to be possible, incorrectly:

(72) ... because $[_{IP}$ t"$_i$ $[_{IP}$ $[_{CP}$t'$_i$ operating on t$_i$] $[_{I'}$ was ...

Very much the same can be repeated for (10), with the partial derivation in (73). Under the definition of barrier in (35), the embedded CP is a barrier for its Spec, t', in that it is not L-marked. Thus movement from t' to t" violates Locality. However under Minimality, as in (34), the embedded CP is not a barrier for its Spec t'; thus movement to t" is allowed under government, and (73) is incorrectly predicted to be possible:

(73) ... without $[_{VP}$ t"$_i$ $[_{VP}$ $[_{V'}$ understanding anything] $[_{CP}$t'$_i$ after reading t$_i$]]]

Thus it appears that while the notion of locality for address-based dependencies reduces to Minimality, the notion of locality for categorial index dependencies must include (35) as well.

Similarly, consider the categorial index derivation for the wellformed (6a) and (9). Of course, if the categorial index derivation is blocked in (6a) and (9), the systematic impossibility of parasitic gaps with adjuncts, which can only form categorial index dependencies, is also explained. But the illformedness of (6a) and (9) under the categorial index derivation is again predicted under the definition of barrier in (35), and not under Minimality. In particular, (6a) and (9) are associated with structures of the type in (74) and (75), where the parasitic gap sequence has reached the embedded Spec of CP position e'. Under Minimality, t' governs e' in (74)-(75), and Locality is satisfied. Under (35), on the other hand, t' does not govern e', because a barrier, CP, intervenes between them, and Locality is not satisfied:

(74) that [$_{IP}$ t' [$_{IP}$ [$_{CP}$ e' operating e] [$_{I'}$ could save t]]]

(75) that I [$_{VP}$ t'[$_{VP}$ [$_{V'}$ bought t] [$_{CP}$ e' without reading e]]]

Thus (35) appears to be needed to constrain categorial index sequences. Indeed under (35) the categorial index derivation of the illformed (6b) and (10) can be blocked, as well as the categorial index derivation of the wellformed (6a) and (9), and ultimately the licencing of parasitic gaps by adjuncts. On the other hand, if the present line of reasoning is correct (35) cannot hold for address-based dependencies, given precisely the wellformed cases of parasitic gaps in (6a) and (9). If so, an additional disjunction must apparently be introduced into Locality theory, in the form of partially different conditions on categorial index and address-based dependencies. This clearly problematic result invalidates neither the generalization argued for in section 4, nor the theorem derived in section 5, nor the link between the two established in this section; hence I will let it stand for the time being. The disjunction however will ultimately have to be eliminated, if the unification problems that beset all locality theories at present are to be overcome.

References

Bennis, H. and T. Hoekstra (1984), Gaps and Parasitic Gaps, *The Linguistic Review* 4, 29-87.

Brody, M. (1992), *Lexico-Logical Form*, ms, UCL.

Chomsky, N. (1981), *Lectures on Government and Binding*, Foris Publications, Dordrecht.

Chomsky, N. (1982), *Some Concepts and Consequences of the Theory of Government and Binding*, MIT Press, Cambridge, Mass.

Chomsky, N. (1986), *Barriers*, MIT Press, Cambridge, Mass.

Cinque, G. (1991), *Types of A'-Dependencies*, MIT Press, Cambridge, Mass.

Engdahl, E. (1983), Parasitic Gaps, *Linguistics and Philosophy* 6, 5-34.

Frampton, J. (1990), Parasitic Gaps and the Theory of Wh-Chains, *Linguistic Inquiry* 21, 49-78

Gazdar, G., (1983), Foot Features and Parasitic Gaps, in W. de Geest and Y. Putseys (eds.), *Sentential Complementation*, Foris Publications, Dordrecht.

Haik, I. (1985), *The Syntax of Operators*, Doctoral Dissertation, MIT

Hoekstra, T. (1988), Parasitic Gaps: A Unified or Composed Chain, in J. Gueron and J.-Y. Pollock (eds.), *Linguistique comparee et theorie du liage*, CNRS, Paris.

Huybregts, R. and H. van Riemsdijk (1984), Parasitic Gaps and ATB, in *Proceedings of NELS 15*.

Kayne, R. (1983), Connectedness, *Linguistic Inquiry* 14, 223-249 and in Kayne (1984).

Kayne, R. (1984), *Connectedness and Binary Branching*, Foris Publications, Dordrecht.

Koster, J. (1986), *Domains and Dynasties*, Foris Publications, Dordrecht.

Longobardi, G. (1984), Connectedness and Island Constraints, in J. Gueron, H. Obenauer and J.-Y. Pollock (eds.), *Grammatical Representation*, Foris, Dordrecht.

Longobardi, G. (1985), Connectedness, Scope, and C-Command, *Linguistic Inquiry* 16, 163-192.

Manzini, M. R. (1988), Constituent Structure and Locality in A. Cardinaletti, G. Cinque and G. Giusti (eds.), *Constituent Structure: Papers from the XI GLOW Conference, Annali di Ca' Foscari* and Foris Publications, Dordrecht, 157-201.

Manzini, M. R. (1990), A New Formalization for Locality Theory, in J. Mascaro and M. Nespor, *Grammar in Progress: GLOW Studies for H. van Riemsdijk*, Foris Publications, Dordrecht.

Manzini, M. R. (1992), *Locality: A Theory and Some of its Empirical Consequences*, MIT Press, Cambridge, Mass.
Pesetsky, D. (1982), *Paths and Categories*, Doctoral Dissertation, MIT.
Reinhart, T. (1976), *The Syntactic Domain of Anaphora*, Doctoral Dissertation, MIT.
Rizzi, L. (1990), *Relativized Minimality*, MIT Press, Cambridge, Mass.
Ross, J. R. (1967), *Constraints on Variables in Syntax*, Doctoral Dissertation, MIT.
Sportiche, D. (1989), *Conditions on Silent Categories*, ms, UCLA.

Taraldsen, K. T. (1981), On the Theoretical Interpretation of a Class of Marked Extractions, in A. Belletti, L. Brandi and L. Rizzi (eds.), *Theory of Markedness in Generative Grammar. Proceedings of the 1979 GLOW Conference*, Scuola Normal Superiore, Pisa.
Vergnaud, J.-R. (1985), *Dependences et niveaux de representation en syntaxe*, Benjamins, Amsterdam.
Williams, E. (1986), The ATB Theory of Parasitic Gaps, *The Linguistic Review* 6.3.