0. Introduction

In this paper we will present the outlines of a theory in which all empty categories are treated on a par. Research of the past twenty years has made it perfectly clear that several distinct empty categories should be distinguished. There are different ways in which one might look at these differences. A very common way to approach empty categories is to take the differences as a consequence of the derivation. Such a theory separates base-generated empty categories from empty categories which are the result of movement. The former are base generated with the feature [+pronominal], PRO and pro, whereas traces are [-pronominal] empty categories that result from movement. We will call this the derivalational approach.

With the exception of Chomsky (1982), Chomsky adopts this view. In Chomsky (1981,1982) a different approach is suggested. Rather than making the distinctions between various types of empty categories in terms of different derivations, four types of empty categories are defined in terms of the different relations they enter into at S-structure. This approach is known as the functional approach to empty categories. As Brody (1984,1985) has argued, the specific functional definitions given by Chomsky (1982) are in fact redundant, as the interplay of the various modules of GB-theory picks out the right combinations of randomly assigned values for the features [+anaphoric] and [+pronominal], if certain assumptions concerning these modules are made. Nevertheless, the main stream of GB-theory has again abandoned this functional approach and has returned to a derivalational perspective again.

In this paper we shall explore a specific version of the functional approach, one which is strictly representational. According to this approach there is in fact only one type of empty category. The four traditional types are taken to be names for different ways in which this empty category is identified. We shall argue that the choice among these different identification strategies is determined by the particular configuration in which the empty category appears, and hence does not depend on any inherent property of the empty category, nor on the derivation that generated it.

There are at least three strategies of identification:\n
-AGR-coindexation which supplies phi-features
-chain-formation
-binding

In Bennis & Hoekstra (this volume) we argued that binding is relevant to identify empty categories which are the head of their chain. The focus of attention in this paper will be the chain formation algorithm. We shall adopt an algorithm in the spirit of Rizzi (1983), which builds chains at S-structure. We argue that this algorithm applies obligatorily whenever possible. In this way we distinguish between two instances of $\epsilon$, viz. those that may, and therefore must enter into a chain, and those which may not, and which are therefore the head of a (possibly single-membered) chain. Evidently, $\epsilon$ which is licensed by a chain is a trace in a derivalational approach, whereas $\epsilon$ which heads a chain is a pronominal under that approach.
We have argued in Bennis & Hoekstra (ibid.) that the distribution of PRO can be characterized in terms of the non-applicability of chain formation. From this it follows that there is no need to characterize PRO in terms of the requirement that PRO be ungoverned. There is no motivation in terms of its distribution, therefore, to inherently characterize PRO as [+pronominal, let alone as [+pronominal, +anaphoric]. We argued that PRO is best analysed as an anaphor with respect to its binding properties, i.e., that binding theory identifies PRO by binding it to an antecedent in a local domain.

In this article we shall argue that chains which relate empty categories to their antecedents are subject to a condition that subsumes the ECP and subjacency. This implies that we regard subjacency as a chain property, on a par with the ECP, rather than as a derivational property. Obviously, our representational perspective leaves us with no alternative option.

1. Chain Formation, ECP and Subjacency

In Bennis & Hoekstra (ibid.) we have introduced the Chain Condition as a condition which determines the distribution of PRO, or rather, the distribution of empty categories which are the head of their chain. The condition in (1) shows a striking similarity to the ECP as it is formulated in Chomsky (1986).

(1) Chain Formation (= B&H (28))
   a. Link an empty category to a c-commanding NP if possible
   b. CP breaks an A-chain
   c. Adjunct XP and DP break a chain

The condition on antecedent government is that two links in a chain cannot be separated by an intervening barrier. The categories that break a chain in (1) are also categories which constitute barriers for antecedent government. If a category is an adjunct, it is not L-marked. If a category is not L-marked, it is a Blocking Category. All Blocking Categories except IP are barriers (inherent barriers). If a category is L-marked, it is a barrier only if it immediately dominates a Blocking Category (barrier under inheritance). Take the following configurations:

(2)    a. anti ... V ... | CP adj e_i [IP e_i ... ]
    b. anti ... V ... | CP adj [IP e_i ... ]
    c. anti ... V ... | CP [IP e_i ... ]
    d. anti ... V ... | CP [IP e_i ... ]

In (2a) an empty category in [SPEC, CP] of an adjunct CP is coindexed with an antecedent outside CP. In this case CP is an inherent barrier because CP is not L-marked. Hence the empty category (ec) is the head of a chain, and both the antecedent and the chain headed by the ec must have an independent theta-role by virtue of the theta-criterion as a condition on chains. As we show in Bennis & Hoekstra (ibid.), the relation between the ec and the antecedent obeys the conditions typical of anaphors. This led us to postulate that ec's that are the head of a chain are identified through binding. In (2b) the CP is an inherent barrier for the same reason as in (2a). This CP is also a barrier under inheritance since IP is a Blocking Category intervening between the two links in a chain. Typically, infinitival adjunct clauses have PRO subjects, i.e., ec's that are the head of a chain. Again, their interpretation is not free, but also exhibits the locality and obligatoriness typical of A-binding. In (2c) the complement CP is not a barrier since CP is L-marked. This configuration, then, is typical of A-bar chains. As a complement to V, the CP in (2d) is not an inherent barrier, but it is a barrier for the ec inside IP under inheritance from IP. Again, such ec's can never be chain-licensed, i.e.
no NP-movement is possible across a CP. The contexts that break a chain, listed in (2), thus constitute barriers on the standard definition of barriers in Chomsky (1986). We can therefore replace (1) by (3).

(3) Chain Formation (preliminary version)
   a. link an empty category to a c-commanding NP if possible
   b. no barrier may intervene between two adjacent links in a chain

We now have a situation in which the barriers theory determines not only the restrictions on antecedent government, but also the distribution of PRO. When we look back at (2), it is clear that the situations in which a barrier intervenes between the antecedent and the empty category (2a,b,d) are precisely the configurations in which the empty category can be PRO.

Several questions suggest themselves, however. First, even if we reduce the ECP to antecedent government, i.e. to a property of chains, there is still the subjacency condition, which is not normally conceived of as a chain property. Hence, if we pursue a chain formation algorithm that applies at S-structure, we have to take a position on subjacency. This point is taken up in section 2, where we shall argue that the ECP and subjacency can in fact be integrated into a single wellformedness requirement on chains.

A second question relates to (2c), where the [SPEC,CP] of an L-marked CP allows a chain to pass through CP and extend into the matrix clause. Although it is clear that such chains are barred when they are headed by an A-position (i.e. because such a chain would be an "improper movement" chain), this is not true for chains headed by an A'-position. Several empirical problems arise in this case, which we shall discuss in section 3. In that section we shall also go into a discussion of other potential barriers, specifically VP. In Bennis & Hoekstra (ibid.) we argued that PRO in object position is excluded on the assumption that obligatory chain formation forms a chain which relates £ and the subject NP (cf.B&H (2Ob)-(20f),(21)). The resulting chain is then rejected as a consequence of the £-criterion, a condition on chains that requires them to have a unique £-role. If VP constitutes a barrier, however, chain formation would be blocked and no £-violation would ensue. Note that barrierhood of VP cannot be voided through adjunction, at least not for the purposes of forming a chain between the object and subject position, again because of the "improper movement" configuration that would result. In order to maintain the result that £c's in object position can only be chain licensed (put in traditional terms: cannot be PRO), we shall argue that VP does not constitute a barrier in these cases, so that our original explanation in terms of obligatory chain formation can be upheld. The non-barrier status of VP will turn out to have other desirable consequences as well.

2. Subjacency as a chain property

While the ECP is normally conceived of as a representational principle, subjacency is normally regarded as a property of derivations. To illustrate a typical line of argument in favour of this distinction, consider sentences of the type in (4).

(4) a. What did [you believe [John to have bought tij]?]
   b. Which picture did [John seem [£j to like tij most]]?

It is quite frequently argued that sentences of this type require a derivational conception of subjacency. The argument runs as follows: At S-structure the embedded clause must be of the category S (i.e. IP). In (4a) this is necessary, because otherwise Case assignment of John by believe is impossible; in (4b), because seem must be able to properly govern the subject
trace e in its complement. Hence, S'-deletion must have taken place prior to S-structure. Similarly, the bracket after did is an S-bracket. Hence, if S is a bounding node for subadjacency, the S-structure representations violate subadjacency, as the sentence initial WH-phrases are separated from their DS positions by two intervening S-brackets. But subadjacency need not be violated at any point in the derivation, if it is assumed that the complement starts out as a CP (S') at DS. WH-movement may then take place in two successive steps through the embedded COMP, each of which only crosses a single bounding node. Then, before reaching S-structure the embedded S'-node, which must be removed in order to allow case assignment in (4a) and head government in (4b), is removed by S'-deletion. Hence, subadjacency is a property of derivations.

Notice that this argument is not valid under the assumptions of Chomsky (1986). If A'-movement from a VP-internal position involves adjunction to VP, the only maximal projection separating the adjunction sites at the embedded and matrix VPs is the embedded IP, which is not a barrier, either because it is L-marked, or because of the assumption that IP never constitutes an inherent barrier. Hence, given the change in perspective we have to be careful with such lines of argumentation in favour of the derivational conception of subadjacency.

The independent empirical range of subadjacency is extremely limited. In Barriers Chomsky introduces the notion of n-subadjacency, where 0-subadjacency holds of a pair of positions that are not separated by any barriers. The subadjacency condition itself, then, requires that two positions related by an instance of move α are 1-subjacent to each other, i.e. are not separated by more than a single barrier.

The standard definition of ECP, on the other hand, contains a disjunction between two licensing conditions, viz. antecedent government and head government, as in (5):

(5) ECP
an empty category B must be properly governed, where α properly governs B iff α governs B and (i) or (ii)
(i) α is a zero level category (head government)
(ii) α and B are coindexed (antecedent government)

It is this disjunction that provides an empirical range for subadjacency. In Barriers the notion of government itself is formulated in terms of 0-subadjacency, i.e. α governs B iff there is no barrier for B that excludes α. The notion of head-government is normally taken to involve θ-marking by the head as well (cf. Stovell 1981). Hence, we can make a two-way distinction between chains, in terms of the absence vs. presence of head-government. The former type concerns adjunct-chains, i.e. chains that terminate in an adjunct (hence not-θ-marked), and head-chains, i.e. chains that terminate in a a zero-level category. These chains can only be licensed by the second clause of the ECP, i.e. by antecedent government. Hence, these chains are characterized by the requirement of zero-subadjacency. Obviously, the 1-subadjacency requirement imposed by the subadjacency condition is trivially met by every chain that meets the antecedent-government requirement of the ECP, and hence the subadjacency condition is empirically empty in these cases.

However, given that clauses (i) and (ii) of (5) are disjunctively formulated, no locality on chains is imposed by the ECP if (i) is satisfied. Hence, as far as the ECP is concerned, extraction from a head-governed position may cross an unlimited number of barriers. This prediction is obviously false. Hence, in order to make up for this lack of empirical precision of the ECP, subadjacency is introduced as a condition on transformational operations, which imposes the condition of 1-subadjacency on derivational steps. The distinction between head-
governed extraction sites and non-head governed extraction sites accounts for contrasts between such well-known pairs as those in (6):

(6) a. Which problem does John know \[ CP, how to solve t \]
b. *How does John know \[ CP, which problem to solve t \]

t is the trace of the WH-phrase in sentence-initial position. The distance which is legitimately bridged in (6a) by the complement which problem does not allow an adjunct such as how to be extracted. This complement/adjunct asymmetry follows from the ECP, under the assumption that CP is a barrier by inheritance from IP. The chain linking how to its extraction site therefore fails to meet the chain condition of the ECP [i.e. antecedent government], as CP will break the chain. This is inconsequential as far as the ECP is concerned for the chain in (6a), since the extraction site is already licit by the head government clause of the ECP.

Nevertheless, two barriers may not be crossed in a single step, as is shown by examples such as those in (7):

(7) a. *Which problem did you meet \[ NP, a man \] \[ CP, who solved t \] \[ CNPC \]
b. *Which problem did \[ IP, he die \] \[ IP, before \] \[ IP, solving t \] \[ subject island \]
c. *Which problem would \[ IP, he die \] \[ IP, before \] \[ IP, solving t \] \[ be impossible \] \[ subject island \]

In (7) the two labeled brackets constitute barriers, the leftmost of which are barriers under inheritance of the non-L-marked rightmost ones. Yet, in spite of these two barriers, the gap is head governed. Hence, the ungrammaticality is not a consequence of the ECP, at least not in its formulation in (5). Clearly then, the empirical range of subjacency is limited to head governed extractions. This is not sufficient motivation for construing it as different from a chain property. We therefore propose to collapse the ECP and subjacency into a single chain condition, as in (8):

(8) Chain Condition (final version)
a. link an empty category to a c-commanding NP if possible
   or maximally 1-subjacent if the foot of C is head governed
b. two successive links L₁ and L₂ of a chain C are 0-subjacent, 0-subjacent if the foot of C is head governed

Not only is subjacency taken as a condition on chains under this formulation, the crucial dependence of 1-subjacency and head-government is much more directly expressed.

3. Intermediate adjunction sites

Given the locality requirement on chains formulated in (8) we now turn to the case in (2c), where we had to conclude that the SPEC of an L-marked CP could be used as a link in a chain, thus providing an escape hatch for chains into the matrix domain. Clearly, this escape hatch is needed in order to allow long WH-movement in successive cyclic steps. The intermediary link cannot be used in all types of chains, however. Head-movement constructions cannot make use of this position, if we adhere to the structure preserving requirement, which states that zero-level categories may move to zero-level positions only. Reformulated in terms of a chain-formation algorithm, this requirement would boil down to saying that a chain whose head is a zero-level category may consist of zero-level categories only. A similar restriction could be imposed on chains headed by an argument position: such a chain may consist of argument positions only. We shall not at this point go into the question of why this requirement on A-chains would hold, but its effect is that CP will always break an argument chain, as it will always inherit barrierhood from IP.
For A'-chains, i.e., chains headed by an expression in A'-position, the possibility of using SPEC of L-marked CP as an intermediary link leads to incorrect predictions in conjunction with other assumptions, specifically the assumption that such chains may also contain A'-positions created through adjunction. The latter assumption is quite crucial in Chomsky's (1986) analysis of A'-movement. To illustrate, consider the analysis of (9).

(9) a. How do you think that John solved this problem?
   b. How do [IP you [VP t1 [ VP think [ cp t2 that [IP John [VP t3
      [VP solved this problem t4]]]]]]

$t_4$ is the extraction site of how. It is assumed that VP is a barrier, because VP is not L-marked by the non-lexical head I. This barrier can be circumvented by means of an adjunction strategy: how adjoins to the embedded VP, and is then dominated by only one segment of VP, which means that it is not included in VP. Hence, the barrierhood of VP is voided. The next step is movement to the embedded [SPEC,CP], which crosses only IP. Since IP is never an inherent barrier, no problem arises. Movement to matrix [SPEC,CP] in one jump is again not allowed, because this would cross the matrix VP barrier, as well as the matrix IP, which would be a barrier under inheritance from VP. Therefore, how adjoins again, this time to the matrix VP, as indicated by $t_1$. This step crosses no barriers, as CP is L-marked. In its adjunction site, $t_2$ is again not included in a barrier that excludes how, i.e., the outer segment of the matrix VP does not count as a barrier, and neither does matrix IP inherit barrierhood.

There are several problems with the assumption that VP constitutes a barrier (cf. Hoekstra 1989 for discussion). We limit ourselves here to a particular problem that arises if the derivation in (9) is adopted. The question is what excludes sentences of the type in (10c) and (11c):

(10) a. dat ik dacht dat hij er met jou t over correspondeerde
    that I thought that he there with you about corresponded
   b. (het probleem) waar ik dacht dat hij met jou t over correspondeerde
    the problem where I thought that he with you about corresponded
   c.*dat ik er dacht dat hij met jou t over correspondeerde
    that I there thought that he with you about corresponded

(11) a. il a dit que Jean a beaucoup lu t de livres
    he has said that J. has many read of books
   b. combien a-t-il dit que Jean a lu t de livres
    how many has he said that J. has read of books
   c.*il a beaucoup dit que Jean a lu t de livres
    he has many said that J. has read of books

(10a) and (11a) show that er and beaucoup may be moved leftward and adjoin to VP. (10b) and (11b) show that their [+WH] counterparts, waar and combien, may move up to the [SPEC,CP] of a matrix clause. If the analysis in (9b) is correct, these long movements would involve an intermediary adjunction onto the matrix VP. Under that analysis, the impossibility of successive cyclic movement of their non-WH counterparts to this adjunction site on the matrix VP is a mystery: why would a derivation be excluded which is a proper subpart of a different derivation which is legitimate?

From the perspective that we are taking, i.e., that chain formation is a S-structure phenomenon, the derivation in (9b) is rather suspect. As Chomsky (1986) notes in the context of his discussion of chain composition, the notion of adjunction makes sense only under an actual movement approach. If chains are formed at S-structure, it seems reasonable to assume that only positions which can be independently motivated are available at that level. The question then is: which positions are available?
Positions which are available without any question are argument positions, or selected positions in general, as their presence is determined by the projection principle. The same is true, although less forcefully so, for at least certain kinds of adverbial modifiers, the base position of which may be determined by some principle analogous to the projection principle for arguments. Apart from these, the non-argumental specifier position of CP is available by virtue of X-bar principles. The situation is radically different, however, for adjunction positions: if they do not host lexical material, there is no indication at S-structure as to their presence. This is to say that there is no ban on adjuncions, but given their invisibility at S-structure, they cannot function as intermediate adjunction sites, a natural consequence of the chain formation algorithm that we adopt.

If this conclusion is correct, the derivation in (9b), involving intermediate adjunctions to VP cannot be correct. If we reject this adjunction strategy, we have to assume that VP is not a barrier. There are several arguments that lead to this conclusion, but it would take us too far afield to discuss these in this paper (cf. Hoekstra 1989 for more extensive discussion). We limit ourselves here to a discussion how VP is not a barrier, at least not inherently. Notice that there is a difference between the relation between IP and VP on the one hand, and IP and other phrase types, such as IP-contained adjuncts and subjects, on the other, in that I entertain a specific selectional relation to V, but not to adjuncts or subjects. This relation is brought out by the process of V-to-I movement in many languages. This process is triggered by the tense connection that exists between I and V, a connection that we may think of as an instance of head-head agreement in the sense of Chomsky (1986:ch. 11). We maintain that by virtue of this relation IP and VP constitute a single multi-segmented projection. Hence, movement across VP does not in fact cross a maximal projection.

Let us return now to the problem in (10)-(11). Nothing in what we have said so far bars adjunction to VP. In fact, such an adjunction is required for the a-examples. What is excluded, however, is the formation of a chain headed by the element in matrix VP adjunction position which uses the adjunction site on the embedded VP as an intermediate link. But it should be possible to move the relevant element to the embedded [SPEC, CP] first, as this position is available at S-structure by virtue of X-bar principles. We should therefore either prevent these elements to use [SPEC, CP] as an intermediate position, or prevent movement from [SPEC, CP] to the adjunction position on the matrix VP. In either case, the only chains that could be constructed, are the chains consisting of er/becoup and their base positions. These chains would fail to meet the chain condition, as the embedded CP is a barrier separating the two links.

We would like to suggest that both solutions are valid, i.e. movement to [SPEC, CP] is restricted to [+WH]-elements (or rather operator elements), which prevents the [-WH]-elements er/becoup in (10)-(11) from moving into it; secondly, movement from [SPEC, CP] is limited to movement to the next [SPEC, CP], a restriction that follows naturally from the assumptions we have argued for in this section. From the hypothesis that only [+WH]-elements may move to [SPEC, CP], an element in an embedded [SPEC, CP] can only be [+WH]. Given that [+WH]-elements must move to [SPEC, CP] it is in fact impossible for such elements to adjoin to e.g. the matrix VP. as this would still require the further step to [SPEC, CP], while intermediate adjunction sites are unavailable under our approach. Hence we derive the COMP-to-COMP requirement from our S-structure approach to chains. Note that the fact that [SPEC, CP] cannot be used as a link in an A-chain also follows from this requirement.
4. Conclusion

In this paper we have provided an approach to empty categories in the tradition of Chomsky (1982). There are various ways in which \( e \) can be identified. We have discussed two identification strategies at some length: binding as a means of identification of \( e \) which is the head of a chain in Bennis & Hoekstra (ibid.), and chain formation for other instances. The former comes into play if the latter is excluded, i.e. \( e \) in its appearance of what is standardly taken to be PRO has a negatively defined distribution. Its distribution is not derived from the binding theory, however, but from the non-applicability of a chain formation algorithm, which at S-structure builds a chain for \( e \) where possible. In forthcoming work we shall provide further motivation for this conception of the distribution of PRO. We have focussed on the traditional arguments against an anaphoric interpretation of PRO, i.e. that PRO, but not anaphors can have a. no antecedent (PRO-arb); b. a long distance antecedent; and c. a split antecedent. We have argued that these arguments are invalid, so that nothing stands in the way to take the interpretation of PRO to be determined by the binding theory. On the positive side, we have shown that the non-existing theory of control does not have to duplicate most of the content of the binding theory.

The non-applicability of chain-formation is determined by the concept of barrier in the sense of Chomsky (1986). Given our S-structure orientation we were forced to abandon the derivational interpretation of subjacency in favour of a chain condition which integrates the ECP and subjacency. The integrated condition is conceptually superior to the traditional combination of subjacency and the ECP in that it expresses the crucial dependence between head-government and subjacency in a much more direct way. In section 3.2. we have shown that our S-structure orientation allows for a natural explanation of certain locality effects on A'-chains, which are quite puzzling from the point of view of the standard theory.

Notes

1. AGR-coindexation may be thought of as a specific instance of chain licensing or of binding, depending on one's assumptions concerning the nature of null subjects, a topic that goes beyond the concerns of the present paper. We shall also not go into null objects. For some comments, the reader is referred to Bennis & Hoekstra (ibid.).

2. In recent studies that accept the distinction between subjacency and ECP, it is often taken as an indication of the correctness of the approach that it allows one to express different degradations of acceptability, i.e. an ECP-violation "feels" stronger than a subjacency violation, while the combination of both an ECP and a subjacency violation yields still worse results. Although such distinctions tend to become rather subtle, they nevertheless seem to be real in some cases. Specifically, there is a rather clear distinction between non-head-governed extractions across a barrier and head-governed extraction across two barriers, the latter being less "dramatic" than the former, certainly if adjuncts are concerned. The gradual decline of acceptability in the latter cases may be taken to suggest a formulation which states that acceptability decreases with an increase of intervening barriers if the foot of a chain is head-governed.

3. Note that the mechanism of VP-adjunction to void the barrierhood of VP is in fact entirely superfluous in languages where V to I movement occurs, if we accept the argument that such raising leads to L-marking of VP. This conclusion would appear to be inescapable, given that it must be assumed that I theta-marks VP, because of the possibility of extracting VP across a WH-island, but also because of the mere possibility of this raising, which would otherwise yield an ECP violation itself. For languages without V to I
raising in syntax, such as English, the VP barrier must still be removed, perhaps as a result of head-head agreement, as Chomsky suggests, as otherwise NP-movement would be barred. Hence, our claim that VP does not constitute an inherent barrier ever can hardly be considered controversial. Notice that another motivation for taking VP to be a barrier might be constructed if IP could be shown to inherit barrierhood from VP. This seems to be impossible, at least insofar adjunct movement is concerned, as these could always be analyzed as involving the adjunction strategy, i.e. it would seem impossible to find relevant evidence. The only candidate to construct such an argument involves the possibility of V-raising directly to C. This impossibility could be argued to result from an ECP-violation as well as a subjacency violation, because IP would inherit barrierhood from VP. However, it seems quite reasonable to assume that the impossibility of this movement is a consequence of something like a constraint against dangling affixes.

References

BENNIS, H. & T. HOEKSTRA 1989 PRO and the Binding Theory, this volume

BRODY, M. 1984 On contextual definitions and the role of chains, in Linguistic Inquiry 15, 325-350
1985 On the complementary distribution of empty categories, in Linguistic Inquiry 16, 475-517

CHOMSKY, N. 1981 Lectures on Government and Binding, Foris Publications, Dordrecht


RIZZI, L. 1983 On Chain Formation, lecture GLOW, York

STOWELL, T. 1981 Origins of phrase structure, diss.MIT