Abstract

Starting from the observation that the standard concepts used for analyses of the syntax of A-positions (EPP, abstract Case) are simply stipulations, this paper proposes that the effects of the EPP and of Case Theory can be derived from an independently motivated component of the grammar, namely the theory of syntactic categories. More precisely, it is argued the EPP and Case phenomena can be analyzed in terms of a model of categorial bonding in which categories interact in order to compensate for deficiencies in their feature matrices.

Reference

1

Categorial Features as the Source of EPP and Abstract Case Phenomena

ERIC HAEBERLI

1 Introduction

Within the generative literature, the Extended Projection Principle (EPP) and the theory of abstract Case have played a prominent role for accounting for properties of the syntax of A-positions. Yet, it has never been clear what the motivation for the existence of these theoretical concepts is. Within the Government and Binding framework (GB, see e.g. Chomsky 1981, 1986), the EPP is a principle which states that every clause needs a subject and more precisely that [Spec, IP] has to be filled. However, it has remained unclear why such a requirement should exist. As for Case Theory, its main aspect within the GB framework is the Case Filter which requires that every overt NP must be assigned abstract Case (Chomsky 1981, Rouveret and Vergnaud 1980). Although the Case Filter has desirable consequences for

1 Earlier versions of the material discussed in this paper were presented at the NELS 28 poster session (Toronto, October 1997), the University of Geneva (November 1997), the 8th Colloquium on Generative Grammar (Palmela, April 1998), University College London (June 1998) and the Workshop on the Effects of Morphological Case (Utrecht, August 1998). I would like to thank the audiences at these presentations for their remarks and suggestions. Special thanks go to Liliane Haegeman, Luigi Rizzi and Ian Roberts for their comments on chapter 2 of my (1999) doctoral dissertation on which this paper is based. Finally, I am also indebted to Ellen Brandner for her valuable suggestions. Needless to say that all remaining errors are my own responsibility.
the analysis of the distribution of overt NPs, it is simply a stipulated principle which does not seem to be derivable in any way.

The same problems that have occurred within the GB framework with respect to the EPP and Case Theory reappear in a different form within the Minimalist Program (MP). Within MP, syntactic processes are driven by the presence of features which are uninterpretable for interface interpretation and therefore have to be made invisible in the course of a derivation. EPP-effects have been related to an uninterpretable D-feature (Chomsky 1995) or to an uninterpretable EPP-feature (Chomsky 1999, 2000), and phenomena related to abstract Case have been analyzed in terms of uninterpretable Case features (Chomsky 1995, 1999, 2000). What is common to all these features is that there is no genuine motivation for their existence. They do not seem to play any role at the interfaces and they therefore simply seem to be generated in order to be deleted again. Thus, the Minimalist versions of the EPP and Case Theory are equally stipulative and unsatisfactory as their predecessors within the GB framework.

In this paper, I propose that phenomena that have been related to the EPP and abstract Case can be derived from an independent component of the grammar, namely the theory of syntactic categories. The proposal will be based on the analysis of categories in terms of feature matrices (cf. Chomsky 1970 and much subsequent work) and on the assumption that the categorial feature matrices contained within a clause interact in order to be licensed. As a consequence, the EPP and the theory of abstract Case can be eliminated as components of the grammar. The paper is organized as follows. In section 2, the main proposal is introduced on the basis of the phenomenon of object movement which is argued to be the effect of a licensing requirement on categorial feature matrices. This proposal is extended to EPP phenomena in section 3. Section 4 deals with some consequences and makes certain basic theoretical points more precise. On the basis of these refinements, additional issues related to the EPP and Case Theory are explored in section 5. Section 6 concludes the paper.

## 2 Object Movement and Categories – Towards Deriving Abstract Case

The empirical starting point for my discussion is a well-known phenomenon which can be found in several languages (examples from Icelandic, see

---

2 In languages with a morphological case system, Case features could of course be argued to play a role at least at PF. However, in languages without case morphology (e.g. Chinese, many Creoles), Case would indeed be a feature which plays no role whatsoever at the interfaces.

(1) a. Hann las (bækurnar) ekki (bækurnar)  
   He read (books-the) not (books-the) 
   'He didn't read the books'

b. Hann las (?*bækur) ekki (bækur)  
   He read (books) not (books)

c. Ég las (þrjár bækur) ekki (þrjár bækur)  
   I read (three books) not (three books)

(1a) shows that a definite nominal object can either follow or precede negation in Icelandic. Given the traditional assumption that negation occurs in a VP-peripheral position, it has generally been proposed that the order object-negation is derived through object movement out of the VP to the left of the adjunct ("object shift", "scrambling"). But there are certain restrictions on movement of nominal objects out of the VP (see the references cited above). In a neutral context, the bare plural object in (1b) cannot precede negation. As Diesing (1996:67f.) points out, the only way to make this order grammatical is by forcing a generic (i.e. non-existent) interpretation for the object, for example by stressing the verb ('he doesn't read books, he only buys them'). Thus, the ungrammatical order in (1b) concerns the existential reading of the object. The same kind of restriction is illustrated in (1c). Although object movement in (1c) is unproblematic, movement affects the interpretation of the object. Whereas the object in a position following negation can have an existential reading, the existential reading is lost after object movement and only a specific (partitive) interpretation is possible for the object when it precedes negation in (1c).

The conclusion that has often been drawn on the basis of data like those shown in (1) is that object movement past a VP-peripheral adjunct is restricted to specific objects and that objects with an existential or non-specific interpretation cannot move out of the VP (see e.g. Cecchetto 1994, Diesing 1992, 1996, Enç 1991, Moltmann 1991, Sportiche 1996). However, Laka (1993) argues that the crucial factor determining movement in cases like (1) is not a semantic notion like specificity as such, but rather the categorical status of nominal arguments. Based on data from Basque, Laka pro-

3 The term object shift has generally been used for object movement in the Scandinavian languages whereas the term scrambling has generally been used for the West Germanic languages. However, I will follow much recent work (see e.g. Bobaljik 1995, Bobaljik and Jonas 1996, Zwart 1997) in treating object movement in these two language groups in a uniform way.
poses that objects remaining in their base position are NPs whereas objects which can move out of the VP are DPs. The observation that in many languages only objects with a specific interpretation can move out of the VP can then be captured under the assumption that in these grammars the semantic feature [+specific] is related to D (Laka 1993:162; see also Chomsky 1995:342 for relating specificity to D).

I will adopt Laka's distinction between NPs and DPs here and I will thus assume that object NPs remain in their VP-internal base position (see 1b and 1c with the existential interpretation) whereas DPs move out of the VP, at least at some stage in a derivation (see 1a/c). The same proposal has been made independently by Philippi (1997:68ff.) (see also Chomsky 1995, Frampton 1995 on the NP/DP distinction in expletive constructions). The question that arises then is why the distributional properties of nominal elements are closely related to their categorial NP/DP status.

Descriptively, the contrast between NPs and DPs can be expressed in a simple way: A constituent headed by a lexical head (N) stays within a projection headed by another lexical head (V) whereas a constituent headed by a functional head (D) moves to a projection headed by another functional head (T). This observation alone does not provide an explanation for the difference in syntactic behavior yet, but I propose that it can be derived in a principled way by developing the theory of syntactic categories which goes back to Chomsky (1970) and which has been adopted in much subsequent work. Chomsky (1970:208) argues that syntactic categories should be analyzed as combinations of features rather than as theoretical primitives. This proposal is made more explicit in later work of Chomsky's (Chomsky 1974). Chomsky proposes that the main categories can be defined on the basis of the features V and N and by attributing positive and negative values to them. The feature system Chomsky proposes is summarized in (2).

(2) a. verb: [-N, +V] c. adjective: [+N, +V]
   b. noun: [+N, -V] d. preposition: [-N, -V]

One way of linking (2) to the issue raised by (1) would be to say that feature matrices as shown in (2) are a source of feature checking within the Minimalist Program. Consider for example the feature matrix of a verb. A verb must be specified as being verbal and as not being nominal. We could assume then that not being nominal means that the nominal feature in the verbal feature matrix is uninterpretable and that it therefore has to be checked and deleted. Thus, the idea would be that categorial feature matrices basically start out with positive values (e.g. [+N, +V] for lexical categories) and that the adequate feature matrices for particular categories are es-
established through checking in the course of a derivation. A verb for example starts out with an uninterpretable N-feature that has to be checked or a noun has an uninterpretable V feature that has to be checked so that the adequate categorial status is established for the interfaces.

Such an approach (pursued in detail in Haeberli 1999, 2000) depends on two assumptions. First, when a categorial feature matrix is inserted in the derivation, its features must be specified as to whether they are interpretable or not. And secondly, the format of the clause structure must be examined at the interfaces with respect to the categorial features. In other words, there has to be an interface filter which determines whether the clause structure has been built adequately during the derivation or not. However, both of these assumptions could be avoided if the underlying motivation for categorial feature checking is conceived of in a slightly different way. Suppose that the presence of a negatively specified feature in a feature matrix is interpreted as a deficiency and that such a deficiency can be made up by an element which contains a positively specified version of the same feature occurring in a local configuration. Thus, all categories defined by one or more negative categorial features have to be licensed by the presence of an element bearing the corresponding positive feature(s).

Given this assumption concerning the licensing of categories, we now can account for the absence of NP-movement out of the VP in (1). When an NP is merged as the complement of a verb ([-N, +V]), the verb’s N-feature deficiency is compensated for by the local occurrence of the NP’s N-feature whereas the NP’s deficiency ( [+N, -V]) is compensated for by the V-feature the verb. The feature matrices of V and N are therefore licensed VP-internally. Thus, there is no trigger for object movement out of the VP past VP-peripheral elements and this movement therefore can be ruled out for reasons of economy. Hence for example the ungrammaticality of the order object-negation in (1b).

What about functional categories? Following Grimshaw (1991) and van Riemsdijk (1990), I will assume that functional categories are defined both by functional and lexical features. Thus, D is defined by N and a functional feature and T is defined by V and a functional feature. With respect to the functional feature, Grimshaw and van Riemsdijk propose that there is a single feature F which defines both D and T. Here, I will modify this proposal slightly and assume that the functional component of a functional categorial feature matrix mirrors the lexical level in the sense that it consists of one functional feature which is verbal and another one which is nominal. More precisely, I will assume that the functional component consists of the fea-

---

4 Thanks to Luigi Rizzi for suggesting an approach along these lines.
tures T and D. Given this feature system, D and finite T can be defined in terms of the following feature matrices:

\[ \text{(3) a. } D: [+D, -T; +N, -V] \quad \text{b. } T: [-D, +T; -N, +V] \]

Consider now the consequences of (3) for the analysis of the contrasts with respect to object movement illustrated in (1). Once a DP is merged as the complement of V, the object does not only contain [-V] but also [-T]. If, as proposed above, we assume that negatively specified features are deficiencies which have to be compensated, the object has to move out of the VP at some point in a derivation in order to establish a local configuration with [+T]. Let us assume that, as proposed by Chomsky (1995:350ff.), object movement targets a specifier in a multiple Spec configuration. In Chomsky’s system, this specifier is the outer [Spec, vP]. However, there do not seem to be any clear arguments which would identify the target of object movement as a Spec of vP. Instead we may assume that an object moves to TP and more precisely to the lower Spec of a multiple Spec configuration within TP given that a subject DP also has to establish a local configuration with [+T] (possibly before moving on to higher landing sites like AgrP or, in V2 languages, CP). Thus, in (1a/c) where an object DP precedes a VP-peripheral adjunct, the object DP has moved to the lower [Spec, TP] in order to license its [-T] feature.

At this point, we can return to the issues raised in the initial section of this paper, namely the theoretical problems raised by the EPP and Case Theory. The discussion of object movement is relevant in this context because object movement past VP-peripheral adjuncts has often been related to Case Theory in the literature (see e.g. Chomsky 1991, Vanden Wyngaard 1989 and much subsequent work). However, the contrasts between different

---

5 Alternatively, we could assume that the label for functional features is less specific, i.e. simply F as proposed by Grimshaw and van Riemsdijk. Hence, we would have two functional features FN and FV. However, for concreteness’ sake, I will use the features T and D.

Another point should be added here. Van Riemsdijk (1990) proposes that lexical categories are defined by a negatively specified functional feature (e.g. [+N, -V, -F]). However, given the analysis proposed here for the contrast in (1), we have to assume that an NP does not have to enter a local configuration with a T-feature and that N is therefore defined simply as [+N, -V] rather than as [-D, -T; +N, -V]. Hence, I propose that the basic building block of a categorical feature matrix consists of the lexical features and that the functional features get added only if they contribute a positive value to the feature matrix.

6 The basic idea mentioned in the text is similar to some extent to a proposal made by Cardinaletti and Starke (1999) for pronouns. In their system, certain elements have to move in order to compensate for structural deficiency. Here, the movement would be triggered by a deficiency in the feature matrix of a category.
types of nominal elements with respect to movement illustrated in (1) does not directly follow from such an approach.

Let us assume that object movement out of the VP is indeed a Case-driven movement. If we combine this assumption with the analysis outlined above according to which object DPs move out of the VP to compensate for a negatively specified T-feature, we may conclude that abstract Case features actually are not features of the type Nominative or Accusative but rather negatively specified verbal features ([-T] and [-V]) contained in the categorial feature matrices of nominal constituents.\(^7\) In terms of this analysis, abstract Case features like Nominative or Accusative can be dispensed with and, hence, also Case Theory as an independent component of the grammar.\(^8\) Thus, we do not have to stipulate that nominal constituents have to be assigned Case (GB) or that they have to check Case features (MP). Instead, abstract Case phenomena can be analyzed in terms of a theory of syntactic categories which is based on the assumption that categories are defined in terms of positive or negative values of features (N, V for lexical categories; D, T, N, V for functional categories) and that negative values are deficiencies that have to be compensated.

Note that this view of the interaction among categorial feature matrices is reminiscent of the model used in chemistry for explaining why atoms join together to form molecules in the way they do (cf. e.g. Gillespie et al. 1989:178ff., Gribbin 1999:74ff. for concise discussions). Atoms are in a chemically stable state if they have a certain number of electrons in their outer layer of electrons. Certain atoms do not have the appropriate number of electrons and they make up for this deficiency by forming molecules with one or more other atoms which also lack electrons. The atoms in a molecule share electrons so that they all have the correct number of electrons. This type of chemical bonding is referred to as “covalent bonding”. For example, a hydrogen atom contains just one electron but it needs two to be in a stable state. Carbon has four electrons in its outer electron layer but it needs eight. A carbon atom and four hydrogen atoms can then combine (CH\(_4\)) and thereby all obtain the right number of electrons by sharing electrons. Each of the hydrogen atoms contributes its electron to complete the outer electron

\(^7\) See also Emonds (1985:52ff.) for interpreting abstract Case as categorial features. In terms of an analysis which makes use of the concept of feature checking (cf. Haegerli 1999, 2000), the proposal made here is also very similar to one developed by Pesetsky and Torrego (2000) since abstract Case can be interpreted as uninterpretable T on D.

\(^8\) Here I am using the term Case Theory in the sense generally used in GB and Minimalist work, i.e. as referring to the theory of abstract Case. Theoretical issues do of course arise with respect to morphologically realized case in languages with a rich case morphology. Cf. Haegerli (2001) for some observations concerning the status of morphological case within the framework outlined here.
layer of the carbon atom (4+1+1+1+1=8). And the four electrons in the outer layer of the carbon atom are shared with the four hydrogen atoms so that each hydrogen atom seems to have the required two electrons (for each H: 1+1=2). Such a covalent bonding relation leads to a chemically stable state. This scenario is similar to the one outlined above where for example V and N help each other to be licensed by contributing a feature which is missing in the feature matrix of the other category. Given this analogy, I will use the term of "categorial bonding" here to refer to the relation between categorial feature matrices which establish a local configuration to compensate for their negatively specified features.

Before turning to additional consequences of this reinterpretation of Case Theory, I will briefly reconsider the second issue raised in the introduction, namely the status of the EPP.

3 Towards Deriving the EPP

Two standard examples illustrating the EPP are given under (4).

(4) a. *(There) is a cat in the office.  
b. *(It) seems that they left.

Given the obligatory occurrence of semantically empty elements such as there and it, the sentences in (4) suggest that there is a purely structural constraint requiring that the subject position be filled. But the source of this constraint has remained unclear.

In terms of the proposals made in the previous section, the obligatory presence of a subject in (4) can be accounted for. Due to the presence of the features [-D] and [-N] in the feature matrix of T (cf. 3b), T has to establish a categorial bonding relation with an element containing [+D] and [+N]. Omitting the expletives in (4) means that no such relation is possible because no element with positive nominal features is in the necessary local configuration with T. T therefore remains deficient, and it violates the licensing condition on categorial feature matrices proposed in section 2.

Once expletives are inserted in (4), the necessary categorial bonding relation can be established. Following basically Chomsky (1995:287), I assume that expletives are D categories, i.e. of the type [+D, -T; +N, -V] in terms of (3). Hence, insertion of the expletive in [Spec, TP] fills the categorial gaps in both T and D, and the requirements of both categorial feature
matrices are satisfied. Thus, the system outlined in the previous section in the context of Case Theory can easily be extended to EPP phenomena.  

4 Some General Consequences

Given the basic assumption made so far that categorial feature matrices interact to compensate for deficiencies, we have obtained a system in which the EPP and Case Theory turn out to be different manifestations of the same underlying phenomenon. What has been referred to as abstract Case is the situation where an element has to compensate for negative verbal features in its feature matrix, whereas the EPP is the effect of the opposite scenario, namely the one where a verbal element has to compensate for nominal features in its feature matrix. Thus, the EPP and Case Theory are simply two sides of the same coin. From a conceptual point of view, this is a desirable result since two apparently unrelated phenomena can be related to a common underlying source. And more generally, instead of having three different independent components of the grammar, i.e. the theory of syntactic categories, the EPP and the theory of abstract Case, we are left with one of these three components, namely the theory of syntactic categories to which a licensing condition on categorial feature matrices has been added. Furthermore, this proposal allows us to eliminate features which seem to be entirely uninterpretable (EPP in any language, abstract Case at least in languages without case morphology) and which therefore have a problematic status in frameworks pursuing Minimalist goals.

Before the consequences of the system outlined so far can be explored in more detail, some general theoretical issues have to be addressed.

4.1 Syntactic Categories

So far we have mainly considered the status of the categories N, D, V and T, and we have not said anything yet about some other major syntactic catego-

---

9 An additional point arises with respect to (4a), namely the status of the argument a cat. Given the proposal made in section 2 that certain nominal constituents are NPs rather than DPs, we may assume that a cat is an NP and that it is licensed within the VP projected by be. What has to be determined then is the structural position the indefinite determiner a occupies. One possibility would be to say that additional head positions above NP are available in the course of a derivation but that these positions are not defined in terms of functional features (cf. the notion of 'proxy category' discussed in section 4.1). Alternatively, as suggested by Ellen Brandner (p.c.), we may assume that the values of the functional features on D depend on the referential properties of D and that D/T on non-specific nominal constituents can remain unspecified. It could be argued then that unspecified features are not subject to the licensing conditions on categorial features proposed in section 2 and that it is therefore only N which has to establish a local configuration with a verbal element (i.e. with be in 4a).
ries. Two additional categories are included in Chomsky's original feature analysis shown in (2) above, namely A and P, and they are defined as [+N, +V] and [-N, -V] respectively. I will adopt the analysis of A here, but will not follow Chomsky's proposal that P is a [-N, -V] category. Chomsky's analysis of P would be incompatible with the proposals made in section 2 because prepositions license nominal constituents and should therefore be able to provide the T- and V-features that are missing in DPs. Hence, P has to be defined in a different way. Such a conclusion may not be undesirable however. Within the set of categories traditionally defined in terms of lexical features (N, V, A, P), P has always had a peculiar status because it also seems to have functional properties. Chomsky (1981:48) therefore only refers to "the first three" among N, V, A and P as lexical categories, thereby excluding prepositions. I will therefore propose here that P should be defined as a functional category. But before considering the status of P within the framework outlined here, let us briefly discuss another important functional category, namely C.

One option for defining C would be to add a feature C to the inventory of categorial features. However, a simpler possibility would be if C could be defined on the basis of the features which are already available. This is what I propose here. C is a functional category and it therefore has to be defined in terms of functional features. Given (3) above, only two combinations of the two functional features T and D have been used so far, namely those where one of the two features is negatively specified. Assuming that no categories are defined by two negative functional feature values (see fn. 5), one option remains, namely [+D, +T; ...]. We therefore could assume that it is this combination of functional features which defines C. That a T-feature partly defines C is in line with proposals according to which C plays a role for temporal interpretation (see e.g. den Besten 1983, Enç 1987, Guéron and Hoekstra 1988, Stowell 1981). As for the assumption that C bears a D-feature, it ties in with the assumption that the EPP is related to [-D] on T and the fact that clauses can satisfy the EPP (e.g. That John left is surprising; see section 5.1 for more details). A relation between C and D is also suggested by the fact that complementizers correspond to D-elements in several languages (see e.g. English that). As for the lexical features of C, I will assume that their status is parallel to the functional features and that they are therefore positively specified, too.\(^\text{10}\) Hence, C is defined as [+D, +T; +N, +V]. As we will see in section 5, this assumption has several desirable con-

\(^{10}\) Thus, the idea would be that prototypical cases of a functional category have lexical features whose values correspond to those of the functional features (cf. also the definitions of (finite) T and D discussed in example 3).
sequences for the analysis of phenomena related to the EPP and Case Theory.\footnote{The feature matrix \([+D, +T; +N, +V]\) will be relevant in particular for the analysis of the highest C-projection in embedded CPs. As for movement to CP (wh, negation, topicalization, focus), I will assume that it is not triggered by categorial features but by non-categorial features such as Wh, Neg, Top or Foc. This distinction could be a potential source for differences between licensing within the A-system and the licensing configurations which are required within the A'-system (see e.g. Haegeman 1995:232f., Laenzlinger 1998:25, Rizzi 1997:282).}

At this point, we can reconsider the status of P within the categorial feature system proposed here. If, as suggested earlier, P is not a genuine lexical category, it has to be defined in terms of functional features. But there are no additional functional feature combinations available any more. \([+D, -T; \ldots]\) defines D, \([-D, +T; \ldots]\) defines T, \([+D, +T; \ldots]\) defines C and finally \([-D, -T; \ldots]\) arguably does not occur (see fn. 5). However, the unavailability of a specific functional feature combination for P may not be surprising. As it has often been observed in the literature, C and P are closely related categories (see e.g. Dubinsky and Williams 1995). This similarity has led Emonds (1985) to propose that, in categorial terms, C and P should be treated as identical. I will adopt this proposal and assume that the highest category within a PP is defined like C (\([+D, +T; +N, +V]\)) (but see section 5.2.3 for a more detailed analysis of P). A desirable consequence of this proposal is that we can account for the licensing of nominal complements of P within the framework proposed here because P can provide the T- and V-features lacking in the feature matrix of DP.

The last issue that I would like to discuss here briefly is the status of the numerous additional functional categories proposed in the recent literature such as Agr, Neg, Asp or Mood in the clausal domain.\footnote{As for "little V" (v) in Chomsky's (1995) system, I will assume that it is simply an independent V head with its own \([-N, +V]\) matrix.} The feature system outlined so far only defines the categories N, V, A, D, T, C and P, and there do not seem to be any additional combinations of feature values which could be argued to characterize for example Agr, Neg, Asp or Mood. One way to integrate such elements into the system proposed here would be to propose that they start out as features on categories which are defined in terms of the inventory of categorial features assumed so far, i.e. for example on V or T in the case of functional categories in the clausal domain. As for the realization of specific projections, I will assume, along the lines of a proposal made by Nash and Rouveret (1997), that such projections are created through proxy categories, i.e. functional heads which have no features of their own and which are created only in the course of the syntactic derivation. Thus, the general idea would be that categorial feature matrices define the necessary "backbone" of the clause structure (e.g., as traditionally assumed, VP-IP-CP...}
in the clausal domain) and that additional projections, realized as proxy pro-
jections, have their origin as features on these basic categories.

4.2 Licensing of Categorial Feature Matrices

As discussed in section 2, the system proposed here is based on the assump-
tion that negative values in categorial feature matrices have to be compen-
sated for by positive values occurring in a local configuration (categorial
bonding) and that this requirement can be a trigger for movement. What has
to be made more precise then is how the licensing of categorial feature ma-
trices works in detail. The system I adopt here shares many basic properties
The following main assumptions will be made:

(i) *Locality*. The minimal assumption with respect to locality is that two
categorial feature matrices are in a local configuration if they occur right
next to each other in a tree structure. Such a relation can be captured by the
concept of immediate dominance. A node X *dominates* a node Y in a tree iff
there is a continuous set of branches going down the tree from X to Y and X
immediately dominates Y if no other node Z intervenes between X and Y.13
Thus, I propose that two categorial feature matrices are in a local configura-
tion if one of them immediately dominates the other one. If, in line with gen-
eral Minimalist assumptions with respect to phrase structure (see Chomsky
1995:241ff.), we assume furthermore that the features of a projection are
those of its head, the categorial features of a category like T can immedi-
ately dominate another categorial feature matrix at every level within the
projection of T. Thus, a local configuration with T can be created at the T-
level (head adjunction), at the T'-level (head-complement configuration) or
at the TP level (specifier-head).14

(ii) *Licensing*. Let us now consider what happens when a category with one
or more negatively specified categorial features is merged in the course of a
derivation. Such a category has to establish a categorial bonding relation
with another category carrying the corresponding positively specified fea-
tures. Adopting proposals made by Chomsky (1995:279ff.) in the context of
feature checking, I will assume here that the establishment of a categorial

---

13 Note that I am using the notion of dominance here simply in terms of nodes. The segment-
category distinction that is sometimes made in this context will therefore not be adopted here.
14 This view of locality does not exactly correspond to the definition of locality that has been
adopted for Minimalist feature checking (cf. Chomsky 1995). That definition generally does
not include head-complement configurations. But cf. e.g. Bobaljik (1995) for a Minimalist
system of feature checking which includes the head-complement relation. Cf. also Zwart
(1993:373, 1997:178ff.) for a similar proposal according to which licensing configurations are
sisterhood relations.
bonding relation is driven by *Attract*, i.e. that a negatively specified feature attracts an element containing a positively specified feature of the same type. Due to the c-command condition on movement (see e.g. Chomsky 1995:253), attraction is assumed to operate only downwards and not upwards. Given that the features of a projection are those of its head, the relevant structural relation for attraction can then be defined as *dominance*. Thus, a negatively specified feature [-F] attracts an element containing a positively specified feature [+F] if [-F] dominates [+F]. Furthermore, I will follow Chomsky (1995) in assuming that it is the closest adequate element that gets attracted. As for closeness, I will generally assume simply that "β is closer to the target K than α if β c-commands α" (Chomsky 1995:358).

If we consider the process of attraction in more detail, we can isolate three main steps. First, once a feature [-F] on a category X is introduced into a derivation, it searches for the closest [+F] in its domain of dominance with which it can establish a categorial bonding relation. Let us call this process *Search*. For reasons which will become clearer in section 5, let us also assume that once [-F] on X has identified another feature on the closest element Y for categorial bonding, no additional *Search* process can be initiated by the same [-F] any more. The way in which we may express this formally is through co-indexation of the two features. However, the use of an index here is simply meant to indicate that one feature has selected another one for a categorial bonding relation. The index therefore gets erased again once such a relation has been established. But before that, the index ensures that, once a [-F] bears an index, no additional features are considered for attraction by [-F] any more. Finally, I will assume that once *Search* has established a relation between two elements X and Y, the relation becomes mutual in the sense that the element Y whose feature [+F] has been selected by [-F] can then also select (and co-index) one or more positively specified features on X (e.g. [+G] on X for [-G] on Y, resulting in what Chomsky (1995:268, 282) calls "free riders" in the context of feature checking). But crucially, a categorial bonding relation is always initiated by a dominating feature.

The second step leading to categorial bonding is *Attract*, i.e. the element Y bearing [+F] gets attracted so that it enters a local configuration with [-F] on X. Given the structural relations observed earlier, *Attract* means that a relation of dominance is turned into one of immediate dominance.

And finally, as the third step, the categorial bonding relation between X and Y is established, i.e. positively specified features on Y compensate for the presence of negatively specified features on X and possibly vice versa. Let us refer to this process as *Bond*. Thus, categorial bonding can be analyzed as a sequence of steps *Search-Attract-Bond*. In section 5, I will illus-
trate and refine this system by reconsidering the way in which the EPP and abstract Case can be derived from categorial bonding.

4.3 The Model of Grammar

Before we return to the EPP and abstract Case, a more basic issue should be raised here briefly, namely the model of grammar on which the framework developed in this paper can be based. As in the previous section, I will adopt proposals made in the Minimalist literature. The two main models in the recent Minimalist literature are the Cyclic Spell Out Model developed by Chomsky (1999, 2000) and the Single Output Model proposed by Bobaljik (1995) or Groat and O'Neil (1996). It seems that in principle the approach adopted here is compatible with both models, but I will adopt the second one mainly for reasons of simplicity because it allows us to avoid potential complications related to the theoretical notion of “phase” introduced in the Cyclic Spell Out Model. In terms of a Single Output Model, a derivation produces a single syntactic structure as its output and this structure is then fed both to the PF interface and to the LF interface. A schematic representation of this model is shown in (5) (from Bobaljik 1995:349).

(5) Single Output Model

As for the distinction between overt and non-overt movement made within such a framework, it will not play an important role for the type of movement we are concerned with here, i.e. for movement establishing categorial bonding relations. I will assume that categorial features with negative values always trigger overt movement. Furthermore, I will follow much recent Minimalist work and assume that movement takes place as soon as possible, i.e. that it is driven by what basically corresponds to Pesetsky’s (1989)

---

15 Most Minimalist models are derivational models and, for the sake of convenience, I will also use derivational terminology here. However, categorial bonding could basically be formulated as a condition on representations (i.e. all categorial feature matrices have to be licensed in a certain syntactic environment), and much of what will be discussed below could probably be recast quite straightforwardly within a representational model of the grammar.
Earliness Principle.\textsuperscript{16} If a feature [+F] is available for attraction by [-F] and the element containing [+F] can get attracted, a categorial bonding relation is established immediately. Otherwise, one of the following two scenarios will arise. First, if a feature [+F] would be available but it cannot get attracted, the derivation proceeds and the categorial bonding relation has to be established later as a side effect of additional syntactic processes (see e.g. sections 5.1 and 5.2.5). As we will see, this option means that an element which gets attracted to a higher position can license features on its way to this position and hence that non-cyclic licensing is legitimate. The second possibility is that no feature [+F] is available in the domain of attraction of [-F]. Categorial bonding then has to be established either by merging an element carrying [+F] or through an additional syntactic process which leads to licensing of "free riders" (cf. point (ii) in section 4.2). However, both of these options are only used if necessary, i.e. if categorial bonding cannot be established immediately.

In this context, an additional issue should be raised briefly. In the previous paragraph I have been focusing on categorial licensing through categorial bonding and I have assumed that categorial bonding takes place as soon as possible during a derivation. However, we may formulate this assumption in more general terms. Consider for example "base generation" of arguments in their argument position. Although the system outlined here allows the creation of categorial bonding when an argument is merged in its base position (cf. the licensing of object NPs in section 2), we will see that merging an argument does not always go together with categorial bonding. Nevertheless, it would not be desirable to let Merge operate randomly during a derivation. I therefore would like to propose that the "as-soon-as-possible" condition does not only apply to categorial licensing but also to Merge which does not involve categorial bonding. Thus, an item is merged as soon as there is a possibility for it to be inserted in the structure. For example, as soon as a theta-role assigner is merged and a theta position therefore becomes available, the relevant argument is merged, too. Hence, syntactic operations in general occur as soon as possible during a derivation.

5 More on the EPP and Abstract Case

Having discussed some basic theoretical assumptions in sections 4.1 to 4.3, let us now return to some specific consequences of these assumptions for the categorial feature licensing framework outlined in sections 2 and 3.

\textsuperscript{16} See also Bobaljik (1995:350, fn. 6) for concluding that processes within a Single Output Model are driven by Earliness.
5.1 The EPP and CPs

Consider first subject CPs as in (6).

(6) \[\text{That Michael scored} \] is not surprising.

On the basis of sentences like (6) where no expletive is present in the subject position, we have to conclude that CPs can satisfy the EPP. Within the framework outlined here, this result follows from the definition of C as a 
\([+D, +T; +N, +V]\) category (cf. section 4.1), i.e. as a category which has the necessary 
\([+D]/[+N]\) features to license the occurrence of 
\([-D]/[-N]\) on T.

However, (6) is interesting for another reason. It has been shown (see e.g. Emonds 1976:127ff., Koster 1978, Stowell 1981:152ff.) that clause-initial CPs as in (6) actually do not occupy the canonical subject position but occur in a topic position. The topicalization analysis is based on the observation that subject CPs do not have the distribution of nominal subjects but rather the distribution of topics. For example, subject CPs (contrary to nominal subjects) cannot occur in preverbal position in clauses in which topicalization is impossible (7a/b; Stowell 1981:153). Similarly, subject CPs, just like topics, cannot occupy the post-auxiliary position in contexts of subject-auxiliary inversion (7c/d; Emonds 1976:131).

(7) a. * Although [with his sister], John was reluctant to travel t.
   b. * Although [that the house is empty] may depress you ...
   c. * Did [the geography course], Bill really want to take t?
   d. * Why did [that Mary liked old records] irritate him?

The data in (7) suggest that clause-initial subject CPs do not behave like nominal subjects but like topics, and that therefore a clause-initial subject CP occurs in a topic position. Hence, what triggers fronting of a subject CP is not the EPP but topicalization and the EPP simply seems to be met as a side effect of topicalization of the CP. This is an unexpected observation. If CPs are constituents which can satisfy the EPP, we would expect that they can move to \[\text{Spec, TP}\] to satisfy the EPP and that no additional trigger for movement would be necessary.

In order to account for this unexpected restriction on the distribution of clausal subjects, Stowell (1981) argues that the non-occurrence of clauses in the canonical subject position is due to the \textit{Case Resistance Principle} (CRP) which states that Case may not be assigned to a category bearing a Case-assigning feature. Yet, Stowell's analysis is problematic for several reasons. For example, there are cases which suggest that a Case assigner actually can
CATEGORIAL FEATURES, EPP AND ABSTRACT CASE

occur in a position where Case is assigned (e.g. PPs as complements of the Case assigners P or V: The baby crawled from under the table or The campaigners planned until Christmas in detail; Jaworska 1986:356). In addition, since I have argued in section 1 that the concept of abstract Case as such already has a problematic status due to its stipulative nature, it would be equally problematic to use a potential "resistance" against Case to account for certain syntactic effects. Finally, even if one adopted Case Theory, it would be unclear why the ability to assign Case should be incompatible with the ability to receive Case.

The framework proposed here provides the basis for a simple alternative analysis of the syntax of subject CPs. Consider the derivational steps that occur in the context shown in (6) when finite T is merged. Finite T has two negatively specified features ([−D] and [−N]) and T therefore searches for an element containing [+D] and [+N] (Search). Assuming that the subject CP has been merged lower in the structure (i.e. in its Theta-position), the result of Search is that the CP is selected for categorial bonding and that therefore the nominal features on T and on C get co-indexed (see section 4.2 point (ii)). This seems to be the right result given that, although the CP does not occur in the subject position in (6), it nevertheless must be the element which licenses the categorial feature matrix of T.

The next step in the categorial bonding process would be Attract. However, if attraction was possible, we would have to conclude that fronting of a subject CP is possible independently of topicalization, contrary to what the observations made above suggest. Let us assume that T actually cannot attract the subject CP. But why should DPs get attracted by T and not CPs? The obvious option for dealing with this contrast within the framework proposed here is that it is related to the categorial feature content of the two categories. CP and DP differ in one crucial respect. C is defined entirely in terms of positive categorial feature values, whereas D contains features with negative values ([−T], [−V]), i.e. features which have to occur in a local configuration with other categorial features. I therefore propose that it is this contrast which determines the movement properties for categorial bonding. Due to their own categorial deficiency, DPs have to be accessible to movement for the purposes of establishing categorial bonding relations. CPs however have a categorial feature matrix which does not require any interaction with other categorial features and CPs are therefore inert for the purposes of categorial bonding. This restriction is expressed in the following constraint along the lines of Stowell's Case Resistance Principle.
(8) Attraction Resistance

Categorial feature matrices which contain only positive values resist attraction for categorial bonding.

Let us now consider the continuation of the derivation of (6). Given (8), the subject CP cannot get attracted to TP purely for the purposes of categorial bonding. Hence, [-D] and [-N] on T still lack the required positively specified nominal features in a local configuration but they remain co-indexed with the nominal features on C. The derivation then proceeds. Suppose that, as pointed out above, (6) involves topicalization. Let us assume furthermore that topicalization involves movement of a topic to a topic projection in the CP domain (see e.g. Müller and Sternewald 1993, Rizzi 1997, Zwart 1993). Thus, in (6), a matrix C marked with a topic feature triggers attraction of the subject CP so that the relevant licensing requirement for a topic can be met. I therefore propose now that as a side effect of this attraction to C, the CP targets [Spec, TP] as an intermediate landing site and then moves on to [Spec, CP].  

Assuming that movement leaves a copy behind, the categorial licensing requirements of T can now be satisfied due to the presence of [+D] and [+N] on the copy of the CP in [Spec, TP]. Hence, although T is unable to attract the CP on its own, its feature matrix is licensed by CP once a higher attractor triggers CP movement. In more general terms, this proposal means that attraction by a certain feature can lead to licensing of categorial feature matrices which occur in the domain between the base position of the moved element and the target of movement, i.e. to non-cyclic creation of licensing positions.

Note finally that our comparison to molecular chemistry introduced in section 2 can be extended to the case discussed in this section. In section 2,
we observed that certain atoms lack one or more electrons in their outer electron layer and that covalent bonding allows them to make up for this deficiency and thus to obtain a chemically stable state. However, there are also atoms such as helium or neon which do not lack any electrons in their outer electron layer. Such atoms are already in a stable state without forming molecules with other atoms and, given this "self-sufficiency", they have the property of being reluctant to interact chemically with other elements. Thus, the status of CP is comparable to that of atoms like neon or helium. Since C has a feature matrix consisting only of positive values, i.e. a feature matrix which is licensed independently of the categorial features of other elements, C is inert and reluctant to interact with the feature matrices of other categories. Only independent processes can activate CP.

5.2 Abstract Case as Categorial Bonding

Having considered an additional aspect of deriving EPP effects from categorial feature matrices, let us now turn to the reanalysis of abstract Case in terms of categorial features. The central aim of Case Theory within the GB and the Minimalist framework is to account for the distribution of nominal constituents within the clause. A list of some phenomena in English that have been related to Case Theory is given in (9). The relevant nominal element in these examples is the DP in bold print.

(9) a. Michael likes Liverpool.
b. They can't imagine Liverpool without Michael.
c. * The manager is proud Michael.
d. * The description Michael was inadequate.
e. * Michael to leave Liverpool would be terrible.
f. They expected Michael to score again.
g. * They tried Michael to stay in Liverpool.
h. Michael was seen by the reporters.
i. Michael seems to be the best player.
j. The fans like Michael.

In the following sections, I will show how the phenomena in (9) can be accounted for in terms of the categorial licensing framework proposed here.

5.2.1 Subject of a Finite Clause (9a)

Within GB, (9a) has been accounted for under the assumption that finite I assigns Nominative and that I and the subject in its specifier are in the right
configuration (government) for Case assignment. Hence, the subject DP satisfies the Case Filter in (9a). A Minimalist version of this approach is that the subject DP moves to [Spec, TP] where it checks its Nominative feature. The result is that Nominative on T and the DP are deleted and no uninterpretable Case features remain at the interfaces.  

In terms of the approach proposed here, (9a) is derived as follows. First, V and the object DP are merged. Then this structure merges with the subject DP, i.e. the subject DP occurs in [Spec, VP]. Finally T and the VP are merged. At this point, [-D] and [-N] on T try to identify the closest element bearing [+D] and [+N] in their dominance domain (Search). The closest constituent which meets the requirements of T is the subject DP and it is therefore selected by T for categorial bonding. At the same time, the subject DP selects [+T] and [+V] on T as the features which can satisfy its own licensing requirements. The subject DP then gets attracted to [Spec, TP] and a categorial bonding relation is established. [+D]/[+N] on the DP satisfy the categorial requirements of T and [+T]/[+V] on T those of D. The feature matrices of the TP and of the subject DP are therefore licensed.

5.2.2 P and Case (9b)

Within the GB framework, prepositions like *without* in (9b) have the same status as verbs in the sense that they allow their complements to satisfy the Case Filter. Prepositions have not been dealt with in detail within the Minimalist literature, but we may assume that the complement of a preposition bears an Accusative Case feature and that this feature can be checked (possibly covertly) against the Accusative feature of P in an AgrO projection dominating the PP.

In section 4.1, I argued, following Emonds (1985), that P and C should be analyzed as related categories. I therefore proposed that the top projection within a PP is defined like C, i.e. as a [+D, +T; +N, +V] category. But consider now the structure that we would obtain under this assumption by merging the preposition *without* and the DP *Michael* in (9b).

---

20 For reasons of space, I will only mention early Minimalist analyses here and I will leave the most recent proposals aside (Chomsky 1999, 2000). In Chomsky's recent work, abstract Case on a nominal constituent can be deleted in a non-local configuration, and movement is triggered by EPP-features only (not by Case). But abstract Case still plays a central role for the analysis of movement since it has the function of "activating" a nominal element so that it can participate in some syntactic operation (Agree or Move; cf. Chomsky 1999:4, 2000:123).

21 For simplicity's sake, I do not distinguish different VP-shells for the two arguments here. But see fn. 12 above and the derivation in section 5.2.7.1 below.
In terms of the proposals made in section 4.2 (ii), the configuration in (10) cannot license the DP because I proposed there, in order to capture the c-command condition on movement, that \textit{Search} is only initiated by negatively specified features and that the \textit{Search} space is restricted to the domain of dominance. In (10), P has only features with positive values and it thus cannot initiate \textit{Search}. D does have negatively specified features but the features on P are not in its domain of dominance. Categorial bonding between P and the DP is therefore impossible and the categorial feature matrix of D is not licensed because D does not enter categorial bonding with [+T]/[+V] features.

Given the problem raised by (10), I propose that P is a complex category consisting of more than one head (see also Koopman 1997, van Riemsdijk 1990, Starke 1993) and that the second head is defined in terms of a different combination of feature values. More precisely, I will assume that the second P-head shares the functional features of the higher P-head but that it has the function of licensing a nominal complement. In order to be able to select a nominal element for categorial bonding, the lower P-head has to have a nominal feature with a negative value. Thus, the feature specification for the lower P-head is [+D, +T; -N, +V]. Due to its [-N] feature, the lower P-head selects its DP or NP complement for categorial bonding. The categorial feature matrix of the lower P-head is then licensed due to the presence of [+N] on its nominal complement. At the same time, the [-T]/[-V] features on nominal complements of P are licensed by [+T]/[+V] on P as a side effect of the categorial bonding relation established by P (“free riders”).

The assumption that P consists of two heads can be motivated as follows. The main argument that has led van Riemsdijk (1990) to propose that P is a complex category is the fact that in some languages we can find circumpositions. Thus, in German examples like those shown in (11), the PP seems to consist of two heads, one prepositional head and one postpositional head (van Riemsdijk 1990:233).

(11) \begin{tabular}{l}
\textit{auf mich zu} \hspace{1cm} (German) \\
on me to 'towards me'
\end{tabular}
In terms of an analysis of P as a complex category, the cases in (11) can simply be analyzed as an overt manifestation of the two heads which prepositions consist of.22

Similar evidence can be found in English. In English there are complex elements like because of or instead of which function as prepositions. These elements also can be argued to be overt manifestations of the double-headed P-structure. Because or instead occupy the higher head whereas of is the lower head which gives rise to categorial bonding (see also Starke 1993 for such phenomena in French and Italian). The licensing property of of will also turn out to be relevant for the analysis of (9c/d) in the next section.

5.2.3 A and N Do Not Assign Structural Case (9c/d)

A basic assumption made within the GB theory of abstract Case is that structural Case can only be assigned by [-N] categories (see e.g. Chomsky 1981:49, Stowell 1981:23). This proposal accounts for the fact that in English nominal elements cannot be licensed by the [+N] categories A ([+N, +V]) and N ([+N, -V]), as shown in (9c/d). However, since abstract Case is an independent property which is not related to categorial features, the restriction of structural Case assignment to [-N] categories seems accidental within the GB framework. In other words, there is no clear reason why [+N] categories should not have the capacity to assign structural Case. The Minimalist version of Case Theory does not lead to any improvement here. It remains unclear why A or N could not bear a structural Accusative feature against which the Case feature of a nominal constituent could be checked.

Within the framework proposed here, the fact that examples like (9c) or (9d) are ungrammatical follows from the categorial feature matrices of A and N. These two categories simply do not have the adequate feature matrices for establishing a categorial bonding relation with nominal complements. Consider first the simple case of an adjective taking an NP complement:

---

22 As for the exact structure of examples like (11), we could assume, following van Riemsdijk (1990), that the lower projection is head-initial and the higher one is head-final. Alternatively, in terms of a purely head-initial structure (see Kayne 1994), the lower projection would have to move to the specifier of the higher one. However, it would not be entirely clear what the trigger for this movement would be.
The situation in (12) is as in (10) above in the sense that we have a dominating feature matrix consisting only of positive feature values. The features of A therefore do not initiate the Search process which would be necessary for categorial bonding (cf. section 4.2). Furthermore, since the nominal complement does not dominate A, it cannot initiate categorial bonding with A. Hence, the [-V] feature on the NP complement of A cannot be licensed. The same conclusion would also hold for a DP complement as shown in (9c) since its V-feature would also remain unaffected by the presence of A.

As for the fact that nouns do not license the occurrence of other nominal elements as shown in (9d), the explanation within the framework proposed here is even simpler. Only positive values of a feature can compensate for the presence of a negative value on a feature matrix. Given that a noun and its nominal complement have the same type of deficiency (verbal features with negative values) they cannot license each other.

The device that saves sentences like those shown in (9c) and (9d) is of-insertion.

Within the GB framework, the examples in (13) are analyzed as involving inherent Case assignment. In addition to the structural requirement of government, the concept of inherent Case also makes reference to Theta role assignment. Thus, A or N assigns inherent Case under government to the nominal complement to which it assigns a Theta role. Of is then the overt manifestation of inherent Case. As for the reference to Theta role assignment, it is motivated by the fact that of-insertion is not an unrestricted device for licensing Case-less nominal constituents.

The adjective in (14) would be a potential inherent Case assigner but in (14) the DP Michael gets its Theta role from the verb score and not from the adjective. Hence, inherent Case assignment is not possible in (14).
An analysis along these lines can be maintained within the framework proposed here. Suppose that the categories A and N, which take complements but which cannot license them if they are nominal, have the capacity of attributing a dummy licenser to their complements. More precisely, this capacity could be related to the [+N] feature. The dummy licenser gets inserted if necessary and it is spelled out as *of*. Following the suggestion made in the previous section, we can assume that *of* is a [+D, +T; -N, +V] head given that it also can be found as the lower head of prepositions. Due to the presence of this head, the nominal elements in (13) can satisfy their licensing requirements imposed by their negative verbal features and the structures are therefore grammatical. As for the ungrammaticality of (14), it can be related to the fact that, once *of* is added on top of the complement of A, *Search* initiated by *of* identifies the infinitival clause (CP, i.e. [+D, +T; +N, +V]) as the closest element which could satisfy the licensing requirements of *of* and the presence of *of* would not have any effect on the subject of the infinitival clause.

In summary, I showed that the approach proposed here allows a principled explanation for the GB observation that A and N (i.e. [+N] categories) generally do not license the occurrence of nominal complements whereas [-N] categories do. In this respect, categorial bonding does not just produce the same results as the traditional Case Theory but it is superior to the latter because the relevant phenomena can be accounted for in a less stipulative way. Only [-N] categories license nominal complements because only these categories can initiate categorial bonding relations with nominal elements and thereby also allow the nominal elements to license their own negative features as "free riders".

5.2.4 No Overt Subjects in Infinitival Clauses (9e)

In the context shown in (9e), an overt subject cannot be licensed within an infinitival clause. The standard analysis of (9e) within the GB framework is that non-finite inflection is not able to assign abstract Case and that therefore subjects of infinitival clauses violate the Case filter. More recent proposals suggest that non-finite inflection assigns a special kind of Case, namely null Case (see e.g. Chomsky and Lasnik 1993), but that this kind of Case is not sufficient to license overt nominal constituents.

Both of these analyses depend on particular stipulations concerning the status of non-finite inflection. The proposal that I will make here may also be stipulative to some extent, but it simply makes use of one of the options which are available within the system outlined here. I propose that finite and non-finite T do not differ with respect to Case assignment but rather with
respect to their lexical categorial features. More precisely, non-finite inflection is a category of the type [-D, +T; -N, -V].

Let us consider what this proposal means for the analysis of the syntax of infinitival clauses. As suggested in section 4.2, the licensing of categorial feature matrices can be defined as a three-step process. First, an element with a negatively specified feature [-F] selects a candidate for attraction, i.e. the closest [+F]. Then, the element bearing [+F] is attracted, and finally the categorial bonding relation is established. So when non-finite T is merged with a VP it searches for [+D], [+N] and [+V] to compensate for [-D], [-N] and [-V]. However, in (9e), there is no constituent in the structure of the infinitival clause which could meet all these requirements of T at the same time. [-V] on non-finite T has to select [+V] on the VP for categorial bonding because the VP is the only element with a positively specified verbal feature in the dominance domain of T. As for the nominal features on T, they have to select the nominal features on the subject DP for categorial bonding since the subject DP is the closest element with [+D] and [+N] in its feature matrix. Hence, the result of Search for non-finite T is inconclusive. Different features on T need to be licensed by different elements. If we assume that attraction always attempts to attract a single element, we may conclude that in such a situation of ambiguity, i.e. in a situation where Search cannot clearly designate a single element for categorial bonding, the process leading to categorial bonding (i.e. Search-Attract-Bond) is interrupted. Given the proposal made in section 4.2 that Search co-indexes features selected for categorial bonding, the only result of the process initiated by non-finite T at this point is that [-D] and [-N] on T are co-indexed with the nominal features of the subject and [-V] on T is coindexed with [+V] of the VP.

The derivation then proceeds and the C-head of the non-finite clause is merged with the non-finite TP. But since C is defined as [+D, +T; +N, +V], its categorial features do not interact with the categorial feature matrices of other elements in the non-finite clause. Hence, the derivation produces a non-finite clause containing several negatively specified features which are not in a local configuration with a corresponding positive feature. In particular the categorial feature matrices of non-finite T and of the subject DP are not licensed, and the structure in (9e) is therefore ruled out. As for the grammaticality of (9e) without an overt subject, I will return to this point briefly in fn. 24 below.

5.2.5 Exceptional Case Marking (9f-9g)
As is well known, infinitival clauses do contain overt subjects in certain contexts, as for example when they are complements of verbs like *expect*. The traditional analysis for cases like (9f) is that certain verbs allow Case assignment to the subject of their infinitival complement clause (exceptional Case marking, ECM). It has been proposed that the particular property of these verbs is that they take infinitival complements which lack a CP level in their structure and that the lack of CP makes infinitival clauses transparent for Case assignment. I will adopt the first assumption here, namely the assumption that ECM verbs select infinitival clauses without a CP.

Let us consider the relevant steps for the derivation of (9f). Up to the point where non-finite T selects features for categorial bonding, the derivation proceeds as described above in section 5.2.4. *Search* cannot unambiguously identify an element for categorial bonding with T and the features on T and the subject cannot be licensed for the moment. However, co-indexation indicates which features would match (i.e. [-D]/[-N] on T with [+D]/[+N] on DP, [-V] on T with [+V] on VP). The next step in the derivation is that, since ECM complements lack CP, the matrix V and non-finite TP merge. V has a negative N-feature. Hence, a [+N] feature has to be selected for categorial bonding, and the closest candidate bearing [+N] is the subject DP of the subordinate clause. Thus, the [-N] feature on the matrix V selects and then attracts the subject DP for categorial bonding. This result has several desirable side effects. Non-finite T has marked D and N on the subject DP as candidates for categorial bonding. However, attraction of the DP by T was impossible due to conflicting requirements of T. But now that the DP gets attracted independently, the DP can move to non-finite TP on its way to matrix V. The categorial bonding relation established in [Spec, TP] means that [-N] and [-D] on T and [-T] on D (as a "free rider") are licensed. Non-finite T is therefore left with only one negatively specified feature, namely [-V], and the VP with its [+V] feature can now be identified unambiguously as the element for attraction to T. Thus, the *Search-Attract-Bond* process can be continued. Given that TP is the complement of V, *Attract* can be met vacuously and T establishes a categorial bonding relation with the VP. [-V] on T is therefore licensed as well. Finally, the subject DP moves to [Spec, VP] of the matrix clause and the result of this movement is that V and the subject DP enter a categorial bonding relation, thereby licensing [-N] on V and [-V] on the subject.23

23 This analysis has three immediate consequences. First, constructions like (9f) involve movement of the embedded subject into the matrix domain as originally proposed by Postal (1974). Second, since the matrix verb precedes the subject of the embedded clause, we have to assume that main verbs in English undergo short movement (see also e.g. Johnson 1991, Kozzumi 1993). And finally, it should be pointed out that [+N] on the subject DP enters a cate-
The question that arises now is why ECM is restricted to certain verbs and why other verbs such as *try* in (9g) do not allow overt subjects in their infinitival complements. As mentioned earlier, I adopt the traditional GB proposal according to which the contrast between ECM and non-ECM verbs is related to the absence or presence of CP in the non-finite clause. ECM verbs take infinitival complements without a CP, whereas non-ECM verbs take non-finite complements which have a CP-level. This assumption provides the basis for accounting for the ungrammaticality of (9g). The derivation for the infinitival CP proceeds as described in section 5.2.4. Then matrix V and the non-finite CP are merged. V has a feature with a negative value [-N] and it therefore initiates a *Search* process to determine the closest potential candidate for categorial bonding. The result is that since C is defined as [+D, +T; +N, +V] and since the CP is the closest element containing [+N], the non-finite CP is selected for categorial bonding by matrix V. The subject of the embedded clause therefore cannot enter a categorial bonding relation with matrix V. The consequence is that several negatively specified features within the embedded clause are not licensed, as it was also the case for (9e). Overt subjects are therefore ruled out within infinitival complements of non-ECM verbs due to the presence of a CP.

5.2.6 NP-Movement (9h-9i)

The next issue that has to be addressed is the phenomenon referred to as "NP-movement" within the GB framework as found in passives (9h) and in raising constructions (9i). Again Case Theory has played a prominent role for the analysis of these phenomena. The claim that has generally been made is that passive verb forms, raising verbs like *seem* and other ergative verbs are not able to assign Case to their complements and that therefore the categorial bonding relation both with non-finite T and with matrix V. Thus, the features on the head of a movement chain and of a lower copy can be used for the purposes of categorial bonding.

The question that arises at this point is how non-overt subjects in infinitival clauses are licensed (e.g. *Michael tried to leave*). One possibility would be to assume, following Hornstein (1999), that the subject of the matrix clause is generated in the infinitival clause and then gets attracted into the matrix domain. The categorial features of the infinitival clause then are licensed as a side effect of attraction into the matrix domain in the same way as in ECM cases. Alternatively and in more traditional terms, the subject of the non-finite clause is a non-overt DP, i.e. PRO. Assuming that non-overt elements have to satisfy certain structural licensing requirements (cf. e.g. Rizzi 1986), it could be argued that the licenser of PRO is non-finite C and that in order to establish a licensing relation C attracts PRO. The result is then again that the categorial features within the infinitival clause can be licensed as a side effect of attraction to a position above non-finite TP (cf. Haeberli 1999:68f. for a more detailed discussion of an option along these lines; cf. also Rizzi 1997:305 for relating PRO licensing to C).
stituents occurring in subject position in (9h/i) have moved there in order to satisfy the Case Filter. However, the lack of Accusative Case assignment by certain verbs is not a necessary consequence of Case Theory but simply an additional assumption as to how abstract Case functions. In particular, it is not clear for example why the Case assigning capacity of a verb should be lost through passivization (cf. also e.g. Brandner 1995, Sobin 1985 for objections against the idea that NP-movement should be related to Case).

Here, I will adopt an alternative analysis of "NP-movement" which has been proposed in the literature (see e.g. Goodall 1996, Kural 1998, Marantz 1992), namely an analysis according to which "NP-movement" is triggered simply by the requirements of the EPP (i.e. [-D][-N] on T). Thus, the assumption is that although a nominal argument may be licensed in lower positions, it still gets attracted to finite TP for categorial bonding and this attraction is the source of "NP-movement".

Another issue that has sometimes been related to abstract Case is the absence of raising out of finite clauses as illustrated in (15).

(15) * John, seems [CP that t is intelligent ].

Within the GB approach, NP-movement is Case-driven. Since John already can get Case in the subordinate clause in (15), there is no trigger for NP-movement any more. Within the Minimalist Program (see Chomsky 1995:284), the analysis of (15) takes a different form but it is still related to abstract Case. The DP John starts out with a Nominative feature. This feature is checked by the Nominative feature of the embedded finite T. Hence, the subject DP does not bear a Nominative feature any more. Subsequently, the subject DP gets attracted by the EPP-feature of the matrix clause. However, since the DP does not bear an abstract Case feature any more, the Nominative feature of the matrix T cannot be checked and an uninterpretable feature remains unchecked at the end of the derivation.

In terms of the system proposed here, the reason why (15) is ungrammatical is simple. The subject of an embedded finite clause cannot be attracted by the negatively specified nominal features of finite T because C is defined as [+D, +T; +N, +V] and because C is closer to matrix T than the embedded subject. Thus, when T searches for features to establish categorial bonding, the closest element which meets the requirements of T is the embedded CP. Hence, the embedded subject is simply not a possible candidate for attraction by matrix T and (15) cannot be derived.  

25 The only way to get a grammatical derivation with the lexical items in (15) would be to leave the subject DP within the embedded clause and to front the CP beyond matrix TP, e.g. by topicalizing it. However, in contrast to contexts involving passivized verbs or contexts like
Superraising is another instantiation of the same phenomenon.

(16)  * John, seems [CP that it is likely [IP t, to leave]].

Again matrix T selects the highest CP for categorial bonding and the DP John therefore cannot get attracted from the lower embedded clause.

5.2.7 Licensing of Objects (1 and 9j)

To conclude this section, let us reconsider the phenomenon with which I started this paper, namely object movement out of the VP.

5.2.7.1 Overt Object Movement

The main claim made in section 2 was that object DPs move out of the VP to license their negatively specified T- and V-features. However, although the underlying motivation for object movement may be related to requirements of the object, the movement has to be triggered differently within a system based on attraction. Categorial bonding through attraction, as outlined in section 4.2.ii, means that a higher feature has to attract a feature that is lower in the structure. Hence, an object DP can only move to TP if it is attracted by a higher (dominating) feature.

To obtain this result, I will adopt an analysis proposed by Richards (1997). Richards (1997:90ff.) argues that object movement out of the VP is the result of multiple attraction by a single attractor which creates a multiple specifier configuration. Thus, I will assume that [-D] on T can attract more than once and that this property of T allows object DPs to move out of the VP. Expressed in terms of categorial bonding, we could say that [-D] on T can have the property of requiring more than one [+D] feature in a local configuration to be licensed. However, I will assume that multiple attraction by [-D] is not freely available but that [-D] has to be associated with a secondary nominal attractor (i.e. [-N]) in order to attract additional nominal elements to TP.26

(6) in section 5.1, CP-topicalization does not seem to be possible with seem (*That John is intelligent seems). I have to leave it open why there is such a restriction on CP-fronting, pointing out simply that this issue is not specifically related to the analysis developed here but that it arises for any other analysis which postulates a topic position in clause-initial position.

26 As discussed below, such a restriction may account for what has been referred to as Holmberg's generalization. In addition, this restriction ensures that D cannot attract an unlimited number of arguments. If D had this capacity, it would also be able to attract for example DP complements of adjectives, thereby licensing such complements productively. However, a language like Dutch licenses object DP movement out of the VP but DP complements of ad-
Given these main assumptions, reconsider example (1a), repeated here in (17):

(17)  Hann las bækurnar ekki                   
      (Icelandic)
      He read books-the not

The derivation of (17) looks as follows:
(i) Merge $V_1$ and object DP; [-N] on $V_1$ selects [+N] on DP for categorial bonding. Suppose now that in functional feature matrices, the functional features take priority over lexical features in the sense that lexical features only become accessible for categorial bonding once the functional features have entered a categorial bonding relation as well. Hence, the categorial bonding process is interrupted after Search and bonding between [-N] on $V_1$ and [+N] on DP is not possible at this point. But selection leads to co-indexation of the two features. At the same time, [-V] on the DP can select [+V] on $V_1$ for categorial bonding.
(ii) Merge $V_2$ (i.e. little v; see fn. 12 above) with VP$_1$, move $V_1$ to $V_2$.27
(iii) N on $V_2$ selects N on object DP for categorial bonding. However, for the same reason as in (i) (interfering functional features), no categorial bonding relation is established. Co-indexation of [-N]/[-V] on $V_2$ and [+N]/[+V] on DP.
(iv) Merge subject DP in [Spec, VP$_2$].
(v) Merge proxy head for negation and VP$_2$ (see section 4.1). Move $V_2/V_1$ to Neg.
(vi) Merge negation in [Spec, NegP].
(vii) Merge finite T and NegP, move Neg/$V_2/V_1$ to T.
(viii) [-D]/[-N] on T select [+D]/[+N] on subject DP for categorial bonding; [-D]/[-N] on T attract the subject DP; the categorial bonding relation in TP licenses [-D]/[-N] on T and [-T]/[-V] on D (“free riders”).
(ix) Due to the presence of additional secondary attractors on T (i.e. [-N] on $V_1$ and $V_2$ within the complex T head), [-D] on T can act as a multiple attractor. It therefore requires a second DP within its projection and the object DP also moves to TP. As for the landing site of the object DP, there are two possibilities: The outer [Spec, TP] or the inner [Spec, TP]. As Richards

27 Note that V-movement cannot be triggered by categorial features within the framework proposed here since verbal heads do not have negatively specified categorial features which would have to be licensed by a verb. I therefore assume that V-movement is triggered by an independent feature, but I will have to leave the nature of this feature open here.
(1997) argues, the latter option can be argued to be the optimal one in terms of Shortest Move (see also Mulders 1997 for an analysis of multiple specifiers along these lines). To move to the outer [Spec, TP] would mean that the specifier position occupied by the subject has to be crossed. Movement to the inner [Spec, TP] is shorter because it does not involve movement past the Spec occupied by the subject. Hence, movement of the object DP to the inner [Spec, TP] has to be chosen for economy reasons. In terms of this analysis, Richards derives the observation that has often been made that A-movement of several arguments out of the VP does not seem to change the order of arguments, i.e. that A-movement out of the VP leads to crossing paths rather than to nesting (see Chomsky 1993, Collins and Thráinsson 1996, Haegeman 1993). As the result of object attraction to T, T and D can enter a categorial bonding relation which licenses [-D] on T, [-N] on V₁/V₂ and [-V]/[-T] on the object DP due to the local presence of [+D] and [+N] on the DP, [+V] on V₁/V₂ and [+T] on T.

The analysis proposed here accounts for what has been referred to as "Holmberg’s Generalization", i.e. the observation that overt object movement generally goes together with verb movement. If, as suggested earlier, multiple attraction by [-D] on T is restricted and depends on the presence of a secondary nominal attractor, then the [-N] features on V₁/V₂ have to move to T to allow [-D] on T to attract a second DP. Object movement therefore goes together with verb movement.

(x) Finally, in an example like (17) where the subject is in initial position: move T/Neg/V₂/V₁ to a higher head (C or proxy AgrS) and subject DP to [Spec, CP] or [Spec, AgrSP], depending on the analysis adopted for subject initial V₂ clauses.

In terms of this derivation, all categories containing features with a negative value occur in categorial bonding relations with categories containing the corresponding positively specified features and the derivation converges. The main assumption that has to be made then to obtain overt object movement is that [-D] on T can act as a multiple attractor.

5.2.7.2 Non-Overt Object Movement

As shown in the Icelandic example in (1a), object DPs do not always precede negation but they can also follow it. This suggests that object DP movement is not obligatory in Icelandic and that [-D] on T is only optionally a multiple attractor. Furthermore, there are languages like English for which

28 Note that for [-N] on V₁/V₂ to be able to occur in a sufficiently local configuration with the DP (i.e. immediate dominance, cf. section 4.1), I assume here that these features can percolate within the projection of their host head T.
it has been assumed that object DPs never undergo overt movement (cf. example 9j). In such a language, [-D] on T simply cannot function as a multiple attractor. Thus, the multiple attraction property of [-D] on T is subject to parametric variation.

The question that arises at this point is how the categorial feature matrix of an object DP can be licensed which does not undergo object movement to TP as the result of multiple attraction by T. My analysis of this issue is based on a contrast with respect to object movement that can be found in a language like Danish. The relevant contrast is shown in (18) (example from Vikner 1994:502).

(18)  Hvorfor læste studenterne (den / *artiklen) ikke (*den / artiklen)
     Why read the-students (it / the-article) not (it / the-article)
     'Why didn't the students read it/the article?'     (Danish)

What (18) shows is that, in a language like Danish, object movement of full nominals is never possible whereas pronominal objects obligatorily move in V-movement contexts. (18) suggests that multiple attraction by D on T is not possible in Danish because otherwise full DPs should be able to undergo overt movement as in Icelandic. Hence, the movement of the pronoun illustrated in (18) must be triggered differently, i.e. by a non-categorial feature which only attracts pronouns but not full nominal constituents (see also e.g. Bobaljik and Jonas 1996, Holmberg 1986, Josefsson 1992, Thráinsson 1996 for analyzing pronominal object movement differently from full DP movement). What may be crucial here is the syntactic status of the two types of elements. Suppose that, as suggested for example by Abney (1987:281ff.), weak pronouns are analyzed as simple D elements which do not contain an NP-complement. Thus, we could assume that pronominal object movement is the result of a process which only allows a D-head to move to TP for the licensing of [-T] but not entire DPs. More precisely, the pronominal D-head gets attracted by a non-categorial feature on T, allowing the D-head to move to T and, as a side effect, to establish a categorial bonding relation with T.

I have to leave it open here what the precise nature of this feature on T is. However, the same analysis can now be used for the licensing of categorial feature matrices of in situ full DP objects. I propose that the D-head of a DP object can also get attracted to T. Yet, due possibly to morphological constraints which prevent the D-head from moving away overtly from its DP, this D-movement is non-overt. But the effect is that [-T] on D is licensed through categorial bonding in T even though the DP does not get attracted to TP as a whole through multiple attraction by [-D] on T.

Let me conclude this section by pointing out that the analysis of non-overt object movement proposed here gets some support from a phenome-
non that can be found outside the Germanic language group. Given the proposals made above, non-overt movement is represented in terms of a chain which has a head as its head and which has a full DP as its lower member. Such a structural configuration is reminiscent of clitic doubling configurations as found in Romance or Greek:

(19)  
Ton idha ton Petro  
Him-ACC I-saw the Peter-ACC  
'I saw Peter'

In (19), we have a clitic which is attached to the verb and a coreferential DP which follows the verb. Thus, we get the same kind of structure as in covert object DP movement: a head in a higher position and a full DP in a lower position. One way to interpret this parallelism would be to assume that the clitic is simply a phonological (postsyntactic) realization of the D-head which has been moved to T non-overtly (see also Alexiadou and Anagnostopoulou 1997, Fanselow 1995 for conclusions along these lines). Thus, the D-head of the object would be moved as outlined above but in clitic doubling languages this D-head can be spelled out as a clitic at PF.

This proposal has an interesting consequence. Within the framework proposed here, non-overt object movement only occurs with object DPs. As for object NPs, their categorial bonding is established VP-internally. Hence, object NPs do not move out of the VP, neither overtly (see 1b/c) nor covertly. Suppose that, as suggested above, clitic doubling configurations basically correspond to non-overt object movement configurations. What we would expect then is that clitic doubling should be restricted along similar lines as shown in (1) for Icelandic object movement.

This expectation indeed seems to be borne out (cf. Alexiadou and Anagnostopoulou 1997, Dobrovie-Sorin 1990, Suñer 1988). Like object movement out of the VP (see example 1), clitic doubling also seems to be sensitive to referential properties of the nominal element. For example, in the Greek example in (20), only the definite nominal argument can occur in a clitic doubling construction, but not the indefinite argument (example from Alexiadou and Anagnostopoulou 1997:149).

(20)  
To diavasa (to vivlio/*kapjo vivlio) me prosohi  
it-ACC I-read (the book-ACC/some book-ACC) carefully  
'I read the book carefully.'

The contrast in (20) can be explained if we assume that, as in languages like Icelandic, the referential properties of nominal arguments in Greek are re-
lated to the presence or absence of a D-head. The indefinite object in (20) has the status of an NP, whereas the definite object is a DP. Thus, only the definite object has to enter a categorial bonding relation in the TP domain and it is therefore only with the definite object that the occurrence of the object clitic is licensed. With an indefinite object, all licensing requirements related to the object are satisfied VP-internally and an object clitic in T is not licensed.

6 Conclusion

Starting from the observation that the standard concepts used for analyses of the syntax of A-positions (EPP, abstract Case) are simply stipulations, I proposed in this paper that the effects of the EPP and of Case Theory can be derived from an independently motivated component of the grammar, namely the theory of syntactic categories. More precisely, I argued that EPP and Case phenomena can be analyzed in terms of a model of categorial bonding in which categories have requirements that can be compared to those of atoms in the sense that categories interact in order to compensate for deficiencies in their categorial feature matrices.

Given the prominent role of categorial features, the consequences of the system proposed here are far-reaching and there are many issues that, for reasons of space, could not be explored here. However, I have shown that some of the main phenomena that have been analyzed in terms of the EPP and abstract Case within GB and Minimalism can be reanalyzed in a simple way in terms of categorial features. For certain phenomena, the categorial approach has clear advantages as compared to standard Case-theoretical analyses. For example the observation that A and N are not structural Case assigners within standard Case Theory follows directly from the categorial feature specifications of these two categories. Or while the movement behavior of objects has remained problematic in terms of Case-theoretical approaches, it follows from the independently proposed DP/NP distinction.

Finally, from a purely conceptual point of view, the categorial feature approach has two important advantages. First of all, it unifies two aspects of the grammar (EPP, Case) which have standardly been treated as independent components. In terms of the proposals made in this paper, the EPP and abstract Case are simply two different manifestations of the same underlying property of the grammar, i.e. categorial bonding. However, there is a second, even more important, consequence of the framework outlined here.

Since the property that has been argued to unify the EPP and abstract Case is related to another component of the grammar (the theory of syntactic categories), we obtain the result that the EPP and abstract Case can simply be eliminated entirely as independent components of the grammar.

References

Chomsky, N. 1974. The Amherst Lectures. Documents Linguistiques, Université de Paris VII.
36/ Eric Haeberli


TEGRAL FEATURES, EPP AND ABSTRACT CASE/37


Eric Haeberli


