The case against Case

RICHARD HUDSON

1 Introduction

In 1968 Charles Fillmore published his article 'The case for case', in which he argued that although languages like English have little morphological case, there are distinctions of a deeper kind, which he called 'cases', which deserve recognition in any grammar. His theory was called Case Grammar, but his cases are now generally called 'semantic roles' or 'theta' roles (though a few linguists still call them cases - e.g. Anderson, Starosta).

More recently the notion 'abstract case' has become important once again, as a major pillar of Government-Binding theory. As in Fillmore's work, abstract case (generally written with a capital C, 'Case') is carefully distinguished from surface, or morphological, case (with a small c); and it is claimed that every language has Case, whether or not it has case. But in contrast with Fillmore's cases, the GB Cases are purely syntactic. For example, the NP *the boy* has distinct Cases in (1a, b), although it has the same semantic role (or Fillmorean case), and no apparent morphological case.

(1)   a. The girl kissed the boy.
        b. The boy was kissed.

The GB theory of Case is interesting and challenging. Its interest comes both from its complex internal structure and also from its subtle empirical predictions. It is challenging because many of its predictions are correct, and GB work in this area, as in so many other areas, provides a benchmark against which other theories should measure themselves. The challenge is especially valuable because the fundamental assumption of Case Theory, that the same system of abstract Cases is equally valid for all languages, is not obviously true; indeed, it has often been challenged as an example of forcing English into the mould of Latin.

The main question is, of course, whether Case Theory is right. The only way to answer this question is to compare this theory with alternative theories about the same range of phenomena, in order to find which of them makes the best predictions. This is what I want to do in the present paper; and specifically, I want to compare GB Case Theory with Dependency Theory, as manifested in Word Grammar (Hudson 1990). Other comparisons could of
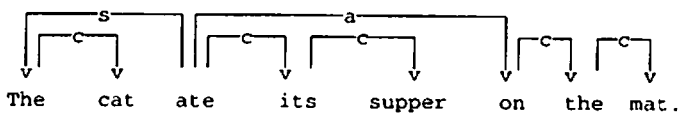
course be made, but I believe this one is particularly important because at least one of the notions which underpins Case Theory is derived from traditional dependency grammar, namely the notion of 'government'. Indeed, I shall show that the GB 'government' is almost the same as the notion 'dependency'; but whereas government is a derived notion for GB, derivable from constituency relations, dependency is (of course) basic to dependency grammar. Direct comparison of the two approaches may therefore throw some light on the very general question of whether constituent structures or dependencies are basic.

2 An outline of Word Grammar Dependency Theory

I shall start with an outline of the parts of WG which relate to dependencies. One reason for giving this before my outline of GB Case Theory (which constitutes the next section) is that WG is much less well known, and therefore possibly of more interest to most readers. The main reason, though, is that the dependency approach to the phenomena with which we are concerned is much more direct, and therefore simpler. Having presented the WG theory, I shall then be able to relate the GB account to it, thereby simplifying the presentation.

In WG, dependencies are basic elements of the structure of a sentence; indeed, the syntactic analysis of a sentence is nothing but the analyses of all the individual words, including the dependency relations among them. (Syntactic structure also includes coordination relations, which we shall ignore for the time being, though they will be briefly in focus in a later section). Larger-than-word constituents are never mentioned either in the grammar, or in general principles that control the grammar, or in sentence structures. For example, the structure of sentence (2) is built round the dependencies shown by the arrows, which point at the dependent of each pair of related words (e.g. *on* depends on *ate*, and *the* depends on *on*).

(2)


```
      ┌────s────┐  ┐┌─────────a────────┐      ┌──┐  ┌──┐
   ┌─c─┐        ┌─c─┐   ┌─c─┐        ┌─c─┐  ┌─c─┐  ┌─c─┐
   v   v        v   v   v   v        v   v  v   v  v   v
  The     cat  ate    its    supper    on   the   mat .
```

Each dependency can be classified more precisely in terms of the dependent's grammatical function; the labels in this diagram distinguish just subject, complement and adjunct, but more precise categories such as object can also be distinguished. You will notice that determiners (including *its*) are treated

here as heads of the following common nouns, as in the DP analysis (see Hudson 1984, 1987, 1990); but the main point to notice is the absence of NPs, VPs and PPs - and even of a 'sentence' node. The most important claim of WG is that such things are unnecessary.

Dependency diagrams are generated by a combination of conditions, some universal and general, some parochial and specific. A structure is well-formed if it satisfies all the relevant conditions. The general conditions concern the abstract geometry of the diagram, and include the following:

(3)   a.   Dependency arrows must not tangle (with certain stated exceptions for coordination and various kinds of 'raising' structures).
       b.   Every word must have just one head (with certain stated exceptions: the sentence 'root', generally a finite verb, has no head, and coordination and 'raising' permit more than one head).

If we ignored the 'certain stated exceptions', condition a would exclude discontinuity, and condition b would exclude double-motherhood; but the exceptions, which are allowed to override the general principles, give greater power to the system - a restricted context-sensitivity. This extra power is not directly relevant to the present paper, so I shall say no more about it.

The more specific and parochial conditions concern the details of the dependency structure, and the general principle is that every dependency must satisfy any relevant conditions. For any given dependency there are in principle five syntactic characteristics which may be controlled by these conditions:

(4)   a.   The type of word that can be the head.
       b.   The type of word that can be the dependent.
       c.   The type of dependency between them (defined in terms of grammatical functions such as subject and complement).
       d.   The status of the dependent (optional or obligatory).
       e.   The order of the dependent relative to the head (and in some cases, relative to other dependents of the same head).

A single condition must mention at least two of these characteristics. For example:

(5)   a.   If the dependent is an adjective, and its function is adjunct, then
its head must be a noun.
b.   If the head is a verb, and the dependent's function is object, then
it must be a noun and it is generally optional.
c.   If the head is a word, then the dependent follows it.
d.   Exceptionally, if the head is a verb, and the dependent is its
subject, the dependent precedes the head.

The formal details of how such rules should be expressed are irrelevant here -
the more so because we are going to compare the WG system with another
one in which conditions are typically expressed in prose.

Given our present interest in the treatment of case, I should explain that
restrictions on (morphological) case are a special case of a type (4b) restriction
- a condition which links a particular kind of head, plus a particular
grammatical function, to a particular kind of dependent word. We have already
seen one example of this general type - (5b), which requires a verb's object
to be a noun. In a language like German, there would be a general condition
on the case of an object, plus exceptional lexical entries which would override
this, as in (6).

(6)   a.   The case of a verb's object is accusative.
b.   The case of the object of FOLGEN, 'follow', is dative.

It is relevant to point out that the same kind of condition can also require the
dependent to be a particular lexical item, e.g. a particular preposition. For
example, the complement of a noun is, by default, the preposition OF, but this
can be overridden:

(7)   a.   The complement of a noun is OF.
b.   The complement of EXPERT is IN.

Many of these conditions on dependencies can be viewed as aids to the
hearer, because each condition restricts the range of possible analyses for a
string of words. In the examples that we have just considered, the conditions
apply to the dependent, but as is well known (Nichols 1986) dependencies can
also be 'marked' on the head word - e.g. by requiring the head to be in some
particular morphological form (such as the 'construct' form of Hebrew or
Arabic). The existence of such restrictions is predicted by dependency theory,
which easily accommodates them.

'Government' is the traditional name for restrictions on the morphosyntactic features of the dependent (other than those covered by 'agreement'). The classic example of this is, of course, government of case, but it is important to see that this is a particular instance of a more general phenomenon which also includes the government of verb-form:

(8) a. The form of a modal verb's complement is infinitive.
     b. The complement of OUGHT is TO.

Notice how the restrictions on the verb's complement may refer either to its inflectional 'form' (infinitive) or to a particular lexical item, TO, in a way which is exactly parallel to the choice between case and lexical prepositions as the marker of a nominal complement. (These similarities between nominal and verbal inflections were sufficiently obvious to the Arabic grammarians of mediaeval Iraq to treat them both as examples of 'government', and even to call the verbal inflections 'case' - Owens 1988.) In a dependency grammar, exactly the same mechanism can be used for imposing both kinds of restriction - case restrictions on nouns, and 'form' restrictions on verbs.

The concept 'government', then, is part of a typology of conditions in a dependency grammar. Another item in this typology is 'agreement', which describes a condition in which the morphosyntactic features of the head are linked to those of the dependent - as it were, a rather special kind of government in which the feature value is a variable. This typology is a helpful aid for the general grammarian, but it plays no part at all in a grammar: there are no conditions which mention either 'government' or 'agreement'. This is perhaps just as well, in view of the difficulties of drawing the line between the two (Matthews 1981: 246 ff).

To summarise, then, I should like to draw attention to the following characteristics of a WG dependency analysis:

(9) a. The dependency analysis is basic, and directly generated by the conditions of the grammar and the theory.
     b. The conditions imposed by the grammar may restrict various characteristics of a dependent, including its morphosyntactic features (e.g. case) and its lexical type.
     c. The theory predicts the existence of conditions which mention at least two of the following: the head, the dependent, the dependent's status, the dependent's function, and the word-order relation between the dependent and its head (or between two co-dependents).

3 An outline of GB Case Theory

It is impossible to do justice to the rich complexity of GB Case Theory in a brief summary, but I assume that most readers know the territory well, so all I need to do here is to clarify what I understand to be the structure of the theory. In the following I shall try to present a coherent account of the theory as explained by Chomsky (1986) and Haegeman (1991).

The theory is concerned primarily with arguments - subjects and complements - and has virtually nothing to say about adjuncts, though these can be NPs with morphological case, as in Latin *Romā venit.* 'He came from Rome', where *Romā* is an ablative NP. It says, first, that every argument must have Case. This is a consequence of a more general requirement that every argument should have a semantic role, combined with a principle that arguments are 'visible' to semantic-role assignment only if they have Case:

(10)   a.   *Theta Criterion*
            Each argument is in a chain to which just one theta role is assigned (by theta-marking).
            Each theta role is assigned to a chain containing just one argument.

       b.   *Visibility*
            α can theta-mark β only if β has Case (or is not an NP)[1].

In other words, any argument NP must have Case if it also has a theta-role. The requirement does not apply to predicative NPs, as in (11), because these have no theta-role (Chomsky 1986:95):

(11)   John is [a fine mathematician].

In general, however, all and only NP arguments must have a Case[2].

---

[1] The definitions of visibility in Chomsky 1986: 94, which is followed by Haegeman 1991: 181, implies that only Case-marked arguments are visible. I assume that this is a mistake, because there are arguments such as PPs and infinitival phrases that cannot receive Case, but which nevertheless have a theta role.

[2] According to Chomsky 1986:140, the clausal complement of a verb can also have Case provided it also has a theta-role; for example, BELIEVE assigns CASE to its clausal complement in (i).

(i)   I believe [that John came].

The next question, then, is how an NP receives Case. This is covered by the principles of case-assignment, which I present below in a form which I believe preserves the intention, though not the letter, of the originals.

(12) α assigns Case to ß iff
   a. ß is an NP,
   b. α governs B,
   c. α is adjacent to ß at D-structure,
   d. α assigns a value for Case to ß.

I have included the adjacency principle (clause c) as it is invoked in some of the most interesting explanations; however I recognise that not all GB linguists accept it. The other clauses introduce the notions of governing and assigning a Case value.

Before we consider what government means in this theory, let us briefly consider how the value of Case is determined.

(13) α assigns value γ for Case to ß iff
   a. α is tensed I, and γ is Nominative, or
   b. α is a V, and γ is Objective, or
   c. α is a P, and γ is Oblique, and α can assign Inherent Case to ß, or
   d. α is an N or A, and γ is Genitive, and α can assign Inherent Case to ß.

These value-assignments are apparently intended to be part of UG, i.e. they apply to all languages; so in those languages where Case is overtly manifested as case, the assignments allow interesting and testable predictions. Clauses c and d also mention the notion 'Inherent Case', for which the following definition is needed:

(14) α can assign Inherent Case to ß only if α theta-marks ß.

This analysis is incompatible with the adjacency principle (espoused by many GB linguists, though not by all), which Haegeman invokes (1991:167) to explain why an adverb can separate a verb from a clausal complement but not from an NP complement:

(ii)   a. *Poirot speaks fluently [English].
       b. Poirot believes sincerely [that English is important].

Chomsky's assumption does not seem to follow from any of the other parts of his theory; in particular, we can be sure that Case is not a prerequisite in general for having a theta-role, as I pointed out in fn. 1. I therefore ignore it in my presentation.

In contrast, clauses a and b deal with what is called Structural Case, but there seems to be no need to define this notion specifically.

We now turn to the most interesting part of the theory, the definition of government (whose importance is of course reflected in the name of the theory, Government and Binding Theory). In comparing the roles of government in GB and in dependency theory, we must first take account of a rather confusing semantic difference. We saw that in dependency theory government is one particular **consequence** of a dependency relation, in which the head selects ('governs') the morphosyntactic features of the dependent. In GB, however, we have just seen in (12) that government is a structural precondition for the selection of Case. Government in the GB sense can exist without government in the dependency grammar sense taking place - e.g. a PP complement is governed by the verb, but receives no Case (and is not governed according to dependency grammar).

This mismatch between the two theories is easily resolved if we recognise that in spite of its name, GB 'government' corresponds to **dependency**. If A governs B in the GB sense, then B depends on A; more specifically, B is a non-adjunct dependent of A[3]. To take the terminological comparison a step further, what dependency theory calls 'government' is what GB calls 'Case-marking'. In both theories, a clear distinction is drawn between the underlying structural relationship and its effect on the features of the governed element; the only difference is in the terminology. Here is a summary:

---

[3] The correspondence between GB government and dependency is not exact, and varies according to which definition of government we adopt. GB government excludes some dependencies, such as the dependency between a *that* complementiser and the following verb (Chomsky 1986:162), because a complementiser is not a lexical head, and therefore not a potential governor.

A more general mismatch between the two theories is that in dependency theory the dependent is a single word, the root of the dependent phrase, whereas in GB it is the entire phrase; e.g. in *without great difficulty*, the dependent of *without* is just *difficulty* for dependency theory, but *great difficulty* for GB. However it is interesting to see that for Chomsky (ibid) government 'reaches down' into the dependent phrase, so that the governor of XP also governs its head. This brings GB very much closer to dependency theory. Unfortunately Chomsky also allows the specifier of XP to be governed by the governor of XP; although this makes the comparison much harder, it is again interesting to notice that his example of a specifier is the determiner of NP, which in more recent analyses is taken as the head of its phrase.

(15)

	GB theory	dependency theory
structural relation	government	dependency
feature relation	case-marking	government

In short, the GB principles which define government can be seen as principles that map a constituent structure onto a dependency structure - or more accurately, as principles that do this for arguments, though not for adjuncts. It is true that the dependency structure is not usually displayed diagrammatically as a separate structure, but this is surely just a matter of notation and presentation. To show the similarity between GB government and dependency, I shall now refer to this shared notion as 'government/dependency', and avoid using 'government' at all in the traditional sense.

The GB principles for recognising government/dependency are as follows:

(16) α **potentially** governs ß iff
   a.   α is the head of its phrase,
   b.   α is A, N, V, P or tensed I, or
        α and ß are coindexed,
   c.   α m-commands[4] ß,
   d.   no barrier for ß intervenes between α and ß.

---

[4] Haegeman 1991:487, 511 uses c-command, rather than m-command, in the definition of government; however her earlier definitions (e.g. 404) use m-command, and none of the intervening discussion appears to require a change so I take this to be an error. Rather confusingly, the definition of c-command in Chomsky 1986:162 is exactly equivalent to that of m-command.

In dependency terms, A is the root of a phrase which 'potentially governs' all the other words that it 'm-commands' - i.e. all the other words in that phrase[5] - except for those which are separated from it by a 'barrier'.

(17)   α is a barrier for β iff
    a. α is a blocking-category for β and is not IP, or
    b. α is XP, α immediately dominates another node γ, and γ is a blocking-category for β.
(18)   α is a blocking-category for β iff
    a. α dominates β,
    a. α is not a theta-marked complement.

The simplest example of a barrier node is an adjunct node - a node which counts as a blocking-category because it is not an ordinary complement to which the head assigns a theta-role. By counting adjunct links as barriers we prevent them from being 'potentially governed', whereas dependency theory includes all the adjuncts of a word among its dependents. As noted earlier, this is a major difference between GB government and dependency.

    More generally, however, any node is a barrier between the head A and a subordinate word B if A does not assign it a theta-role. The assumption is that a head assigns theta-roles to some of its complements, but not to all of them; for example, it doesn't assign a theta-role to a predicative complement as we saw above in connection with (11). Another complement-type that is not theta-marked is an infinitival (i.e. IP) 'Exceptional Case Marking' complement, as in

(19)   I believe [him to be honest]$_{IP}$.

We should expect such complements to be barriers to government, and therefore to Case-assignment, but they are not, because the subject of the infinitive (*him*) receives Case from the main verb. The special status of these infinitival phrases is the reason for the exception made for IP in (17a). Conversely, there are other infinitival phrases that do receive a theta role and

---

[5] The definition of m-command refers to the lowest maximal projection above the word in question; if A is this word, and its category is X, then A m-commands all the other words within the lowest XP that contains it. By this definition, it m-commands any adjunct of X', but no adjuncts of XP. This distinction is not recognised in dependency theory, so it may be that some dependency adjuncts of A are not m-commanded by A. However, this makes no difference to government, because no adjuncts of A are governed by it in any case (thanks to the mention of blocking-categories).

should therefore not count as barriers, but that do count as one nevertheless; an example is (20). Such cases motivate (17b).

(20)   I decided [to see the film].

The main point of these details is to explain why an infinitive must have an overt subject if governed by BELIEVE, but not if governed by DECIDE.

Finally we come to the most obvious restriction of all, in dependency terms: that if B depends on C, and C depends on A, then B (typically) does not depend on A as well - in other words, the dependency principle that a word has only one head. This is expressed in GB in terms of 'minimality'.

(21)   α **actually** governs ß iff minimality is respected by α and ß.
(22)   α and ß respect minimality iff there is no other node γ such that:
       a. γ potentially governs ß,
       b. α c-commands γ.

For an example of minimality, consider the constituent structure for the VP *look for Mary*.

(23)

```
            VP
          /    \
        V        PP
        |       /   \
       |      P      NP
      look    for    Mary
```
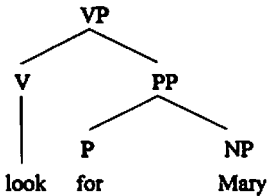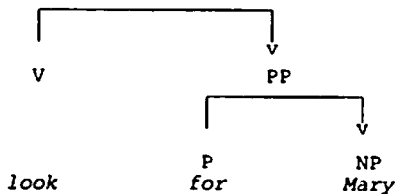
Minimality allows *look* to govern the PP, and *for* to govern *Mary*, but it prevents *look* from governing *Mary* because *for* is also a potential (and actual) governor of *Mary*.

These GB principles allow us to construct a 'government/dependency structure' for each phrase structure; for example, a structure for (23) showing that *look* governs *for Mary* and *for* governs *Mary*. I shall show each government/dependency relation by an arrow pointing from the governor to the dependent:

(24)

```
        ┌──────────────┐
        │              ▼
        V              PP
                ┌──────────┐
                │          ▼
                P          NP
look           for        Mary
```

Such structures already look very similar to pure dependency structures, but a full translation requires two further assumptions: that if a phrase XP is governed by A, then its head X is also governed by A (a principle which is accepted by Chomsky, as noted in fn 3); and that if the relation between A and X is shown, that between A and XP is redundant. If we apply the various definitions given above to the structure in (23), we get the government/dependency structure in (25).

(25)

```
    ┌────┐  ┌────────┐
    │    ▼  │        ▼
look   for        Mary
```
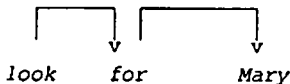
And if we apply them to a complete GB-style sentence structure for *The cat ate its supper on the mat*, we will get the government/dependency structure given above in (2), except for the adjunct link between *ate* and *on*.

This almost completes my presentation of the GB theory of Case and government. There is just one further element that deserves to be mentioned, which is the notion that Case-assignment is directional, the direction varying from language to language (possibly as a parameter distinct from the head parameter - Travis 1989). In English, Case is assigned uniformly to the right (Chomsky 1986:193). This claim corresponds to a dependency analysis in which all complements follow their heads[6].

4 Subjects of infinitives

We now move to a direct empirical comparison of the two theories, starting with the data that Chomsky describes as the primary source of his ideas about

---

[6]It is unclear how the principle of rightward Case assignment applies to subjects in English, if these receive Nominative Case from I when they are on its left.

Case, infinitival clauses with subjects (1986:186). I shall use Chomsky's examples, though I shall change the order.

(26)   *[John to be the winner] is unlikely.

According to GB, this is ungrammatical because *John* has no Case, and is therefore invisible. The reason why *John* has no Case is that it has no governor, infinitival TO not being on the list of recognised governors in (16). This restriction is simply stipulated (by listing). Similarly, WG stipulates that an infinitive (which for present purposes includes TO) normally has no subject.

(27)   An infinitive has no subject.

In the absence of any head for *John*, the sentence is ruled out.
       One empirical difference between the two theories lies in their generality. GB predicts that subjects are impossible with any non-finite verb; WG allows the restriction to apply only to infinitives. Examples like the following are relevant:

(28a) [John winning the race] surprised us.
(28b) [John being the winner], we gave him the prize.

These are both impeccably grammatical, but are predicted ungrammatical by GB. WG accommodates them by a general rule which allows any non-finite verb optionally to have a subject, with (27) as an exception.

(29)   A non-finite verb has an optional subject.

       Another difference between the predictions of GB and WG concerns subjects that are not NPs, and which therefore do not need Case. Take a sentence like (30), in which a PP is the subject.

(30)   By the river is a good picnic spot.

According to WG, this prepositional subject should be possible with some non-finite verbs, but not with an infinitive, following the pattern established in (26) and (28). For GB, on the other hand, the principles responsible for (26) and (28) are irrelevant, so we predict grammaticality throughout. Here are some relevant sentences, which support the WG predictions.

(31a) *By the river to be a good picnic spot would be unusual.
(31b) By the river being a good picnic spot supported my theory.
(31c) By the river being a good picnic spot, we decided to go there.

The rule in (27) also accounts for most of Chomsky's other examples:

(32a) *proud John to be the winner
(32b) *the belief John to be the winner
(32c) *I wonder to whom John to give the book.

These infinitives contrast with tensed verbs, which must have a subject.

(33a) proud that John is the winner
(33b) the belief that John is the winner
(33c) I wonder to whom John is to give the book.

Such examples are allowed in WG by a special rule for tensed verbs, which corresponds to the stipulation in (16) that tensed I is a potential governor.

(34)   A tensed verb has a subject.

One of Chomsky's remaining examples is (35).

(35)   I believe John to be the winner.

This is allowed in GB by another stipulation (17), that IP does not count as a barrier, combined with the assumption that *John to be the winner* in (35) is an IP, not a CP; in other words, BELIEVE is identifiable in the lexicon as an 'exceptional Case marker'. Most theories other than GB take a very different view of sentences like this, and analyse *John* as 'shared' by the two verbs - as object of *believe*, and subject of *to be*. The relevant facts are expressed in (36), where 'incomplement' is the WG equivalent of the LFG 'XCOMP', a complement which shares its subject with the preceding verb.

(36a) BELIEVE has an incomplement.
(36b) BELIEVE has an object.
(36c) If a word W has an incomplement I, then I's subject is also W's object
        (if it has one); otherwise it is W's subject.

This combination of rules automatically overrides the general ban on subjects for infinitives in (27), because being an incomplement is more specific than being an infinitive. Therefore the infinitive not only may have a subject, but in fact must have one because BELIEVE has to have an object.

One empirical difference between the two analyses involves the status of the shared NP, as seen in pairs like (37).

(37a) I believe John to be the winner.
(37b) *I believe # to be the winner.

Why is the shared overt NP obligatory? The GB explanation is that the element which I have indicated as '#' must be PRO, which must not be governed (as it would be, in this case, by *believe*). This explanation predicts that the same will be true of all ECM verbs. In contrast, the WG analysis of (37a) takes *John* as the object of BELIEVE, and since some verbs have optional objects (e.g. WASH), we may expect to find verbs which have optional objects but which are otherwise like BELIEVE. The facts are as predicted by WG: at least two verbs, CLAIM and PRETEND, are like BELIEVE except that the shared object NP is optional:

(38a) I claimed John to be the winner.
(38b) I claimed # to be the winner.

Presumably it is possible to describe these facts in GB, by recognising two distinct entries for CLAIM (and PRETEND), one taking an infinitival IP, the other an infinitival CP; but in contrast with WG, their existence is not predictable. Nor is it easy to see how the CP can be prevented from having an introductory FOR (as with PREFER):

(39)   *I claimed for John to be the winner.

Chomsky's last two examples both contain *for* before the infinitive:

(40a) For John to be the winner is unlikely.
(40b) I'd prefer for John to be the winner.

Why does the presence of *for* permit - in fact, require - the infinitive to have a subject? One would expect the GB answer to involve a government relation between FOR and the subject, but the details are unclear. This is because *for* is analysed as a complementiser, and Chomsky says explicitly (1986:162) that

the complementiser *that* is not a potential governor. Nonetheless, *for* is assumed to govern the infinitive's subject, and to assign it Case (ibid:172). Implicit in a later discussion (ibid:202) is the assumption that this *for* is a preposition (the assumption is explicit in Haegeman 1991:155), hence its Case-assigning and governing properties; but how can a word be both a complementiser and a preposition at the same time?[7] Moreover, how is it that a preposition can have an IP (the infinitival phrase) as its complement? If this is possible for *for*, why not for *of*, *by* and so on as well?

Suppose the above problems can be solved within the existing Case Theory, one interesting prediction emerges. This is that the distribution of CP containing FOR as C will be no more similar to that of CP without any overt C (but with an infinitival IP) than it is to the distribution of, say, a CP with THAT as C[8]. But this is clearly not true: without exception, wherever *for* + NP + infinitival phrase is possible, so is just an infinitival phrase.

In short, GB Case Theory throws remarkably little light on the effects of *for* before an infinitive, and where it does generate predictions these turn out to be false.

The WG analysis of the sequence *for* + NP + infinitive phrase is once again stipulative, but simple. Ignoring for the moment the classification of FOR, it is allowed to have two complements, an object and an incomplement -

---

[7] The status of the category 'complementiser' has been very uncertain ever since it was introduced into linguistics in the 1960s. On the one hand it is part of the set of basic categories recognised in a grammar, with exactly the same formal status as 'V', 'N' and so on; but on the other, it is generally treated as the name for a position or function, which puts it on the same footing as 'subject', 'complement', etc - categories which are claimed to be derivable from the structural configuration, and not basic. One analysis in which it is particularly clear that Comp is the name of a function is the Barriers analysis of English subject-auxiliary inversion, where the auxiliary verb is moved into Comp; e.g. in (i) the verb *has* has the function 'Comp'.

(i)   Has he finished?

If infinitival *for* is a preposition as well as Comp, this is another example of Comp being the name of a function, and not of a word-class.

[8] It is true that recent work in GB has assumed that a functional category such as C projects to its maximal projection not only its own category but also that of its complement; for instance, a DP also counts as an NP (Abney 1987). This would allow the 'infinitival' features of TO to project up to its IP, and then up to the CP. The trouble is that these features will have no privileged status among all the other features projected up to CP (e.g. from the verb in the bottom VP, or from other functional categories like NegP), so there is no reason to expect them to have any special relevance to the distribution of the whole CP.

i.e. just the same pattern as we recognized above after BELIEVE. Roughly speaking, the incomplement is the infinitive phrase, but it would be more accurate to say that the incomplement is just the word TO, just as the object of FOR is the noun which is the root of the NP. To complete the structure, we apply the same 'sharing' rule (36c) as for BELIEVE, identifying the object of FOR as the subject of TO. The result is a structure like (41).

(41)
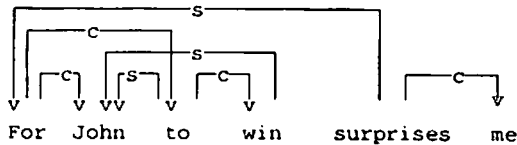
Therefore the WG explanation for the possibility of a subject after FOR is that FOR's valency requires it to have an object and an incomplement (TO), which means that its object also serves as the subject of TO.

What about the relations between infinitive phrases with an introductory FOR, and those without one, which we found to be problematic for GB? As in the GB analysis, FOR is the root of the subordinate clause, whereas a simple infinitive phrase like *to win* has TO as its root.

(42)
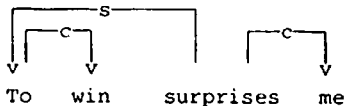
Since it is the root of a phrase which decides where the phrase can occur, this difference between phrases rooted in FOR and those rooted in TO would seem to raise precisely the same problem as I noted for GB: we can expect such phrases to be no more similar in their distribution than either is to a phrase rooted in, say, THAT. However WG offers a solution to the problem: it is possible to treat two distinct words as sub-cases of the same lexical item (e.g. *I* and *me* are both covered by ME), so TO and FOR can be subsumed under FOR/TO. Any construction which allows TO with or without FOR (+ NP) now allows FOR/TO, which may be realised either as FOR (with a dependent TO), or just as TO. (For further discussion see Hudson 1990:243-4.) This analysis makes further predictions about Chomsky's examples (40): that in each of these constructions the sequence *for* + NP is optional.

(43a) (For John) to be the winner is unlikely.
(39b) I'd prefer (for John) to be the winner.

This prediction does not seem to follow from the GB analysis.

What I have shown in this section is that WG can explain all the facts about infinitive phrases that Chomsky lists, and does so more successfully than GB. In the remaining sections of the paper I shall review some further empirical differences between the two theories.

5 Obligatory expletives

How do the two theories explain why expletives are obligatory in some structures? Let us consider the requirement that a tensed verb should have a subject, which explains the pattern in (44).

(44a) It rained.
(44b) *Rained.

We have already seen the WG rule which explains it - rule (34), repeated here:

(34)   A tensed verb has a subject.

Notice that this makes no reference either to semantic roles or to case. Although it is stipulative, it is no more so than the stipulation that tensed IP is a governor.

The GB explanation involves the Extended Projection Principle (Haegeman 1991:59), which requires any sentence to have a subject. Since an expletive has no theta-role, visibility is irrelevant and so is the whole of Case Theory. This allows an interesting prediction: in those constructions where an overt subject is excluded by Case Theory, an expletive should be possible. For example, IT should be possible without a supporting FOR in an infinitival clause:

(45)   *It to rain is unusual.

The prediction is wrong. The reason is easy to see: the rules which determine whether a subject is possible or not are purely syntactic, and blind to semantic

matters such as the presence or absence of a theta role. Basing Case Theory on visibility seems to be a step in the wrong direction.

Suppose we return to the earlier version of Case Theory, based on the Case Filter (Chomsky 1986:74):

(46)   Every phonetically realized NP must be assigned abstract Case.

As Chomsky notes (ibid:132), this solves the problem noted above, because IT, as a phonetically realized NP, can't occur in Case-free positions like (45). It also, however, extends the Case requirement to every single NP, predicting that overt NPs will always, without exception, be governed by Case-assigners. It is very easy to find exceptions. First, there are vocatives, which in languages like Latin are very explicitly case-marked (e.g. *Brute*, contrasting with nominative *Brutus*) but are ungoverned.

(47)   Pluit, Brute. 'It is raining, oh Brutus.'

And second there are NPs used as adjuncts (what Larson 1985 calls 'Bare-NP adverbs'), such as *this morning* in (48).

(48)   I saw him this morning.

In many languages NPs are more widely used as adjuncts than they are in English, often with the help of morphological case marking. Like vocatives, such examples are excluded by the Case Filter[9]. It seems, then, that replacing visibility by the Case Filter creates as many problems as it solves.

6 Case selection

According to the version of Case Theory presented above, a major difference between GB and WG lies in their treatment of Case selection - i.e. the choice of a value for Case (nominative, objective, etc.). Taken at its face value, Case

---

[9] Larson's account of adjunct NPs rests on the assumption that they assign Case to themselves; but this analysis founders on the rather obvious fact that they are no better than any other NPs in those positions where (external) Case is said to be missing, as in infinitival phrases without FOR:

(i)   For this morning to be so cold is a nuisance.
(ii)   *This morning to be so cold is a nuisance.

Theory predicts not only that all languages have Case, but that the range of available Cases is uniform across languages, and that the Case selected for an NP with a given function is also uniform - an NP governed by V is always Objective, one governed by P is always Oblique, and so on. This follows from our (13), repeated here:

(13) α assigns value γ for Case to β iff
   a. α is tensed I, and γ is Nominative, or
   b. α is a V, and γ is Objective, or
   c. α is a P, and γ is Oblique, and α can assign Inherent Case to β, or
   d. α is an N or A, and γ is Genitive, and α can assign Inherent Case to β.

In contrast, WG predicts considerable variation, with some languages having no case (or Case) at all, and others having far more than the four values listed in (13).

At least as far as morphological case is concerned, there is no dispute about which prediction is correct. Some languages (e.g. Chinese) do indeed have absolutely no morphological case, and others have vast numbers; e.g. according to Kiefer 1992, quoting Mel'čuk 1986, the language Tabassarian has no fewer than 46 distinct cases. Moreover, the case assigned to an NP with a particular function may differ from language to language; e.g. in many languages an objective (accusative) case is often assigned to the object of a preposition, and in Modern Greek indirect objects receive genitive case, whereas in many other languages they receive dative case, which is distinct from genitive. There are even languages in which a subject NP need not be Nominative (e.g. Icelandic, according to Zaenen et al, 1985.)

If we assume that when morphological case does exist, abstract Case always receives the same values, then we can be sure that the predictions in (13) are false. However, such problems are well known, and a good deal of latitude seems to be intended in the interpretation of Case Theory. For example, Chomsky and Lasnik (1991) recognise 'Case differences among verbs', such that some verbs assign Case, but others don't (e.g. ASK does, WONDER doesn't); and Woolford (1991) even suggests that languages may differ as to whether case assignment is obligatory or optional. It would probably be best to recognise that GB linguists have not yet reached a firm position on such questions, so my presentation of the theory is actually too rigid. On the other hand, the looser the theory is, the fewer predictions it allows, so we may have to wait some time before we can assess it empirically.

However, it is worth raising the question of the choice of pronoun forms in English (*I* versus *me*, *he* versus *him*, and so on), because these are often quoted as clear evidence that English has case, and that it therefore has Case too (e.g. Haegeman ibid:144). It seems very unlikely that this is in fact a distinction of case, because the choice interacts with coordination:

(49a) I/*me went there.
(49b) John and me/?I went there.

The choice of *me* rather than *I* in examples like (49b) is normal for very many speakers of English (perhaps the majority in the UK), and is very hard to explain if the choice is due to a case assigned to the subject by the verbal inflection. In languages that have real case, such as German, there is no interaction whatever between case and coordination, so if the same mechanism is supposed to be responsible for the choice of pronoun form in English, we need an explanation for this fundamental difference between English and German[10].

WG allows a different account of these facts (Hudson 1990:230-2). There is no general requirement that subjects of tensed verbs should have case; indeed, English has no case whatever. Rather, there is a constraint on a handful of pronouns (six of them, to be precise) that when they are used as subject of a tensed verb, they have a special form. (This rule involves the two ends of a specified dependency relation, so it fits comfortably into the typology of possible WG rules given in (4), and to that extent is predicted by the theory.) In non-standard dialects the rule is restricted further, by referring to the pronoun's function in coordination as well as to its dependency function. (As I mentioned briefly, WG recognises coordination as the second main type of relation in syntax, alongside dependency; so this interaction is also not unexpected.)

---

[10] In a recent interview, Chomsky (1991:30) briefly discusses examples like *Him and me were here*, and suggests that the reason for this choice of forms is because 'nominative is used only as the subject of a tense sentence' - and not as a part of a compound subject. This clearly can't be the right explanation, partly because some English speakers opt for *he and I* in such sentences, and partly because in languages like German nobody, without exception, opts for accusative case forms in coordinated subjects.

7 Case manifestation

Once abstract Case is distinguished from morphological case, it is natural to look for manifestations other than the inflectional morphology that has always been traditionally associated with case. This has allowed proponents of Case Theory to claim that English not only has abstract Case, but that at least some of the values for Case are formally distinct. The primary example of this is of course the distinction between *I* and *me*, and so on, which I have just suggested are not in fact distinctions of case. Three other special forms have also been suggested as manifestations of Case: possessive pronouns, the apostrophe *'s*, and various prepositions, as in the following.

(50a) Her pictures cost £5.
(50b) Mary's pictures cost £5.
(50c) The pictures of Mary cost £5.

We shall consider these purported case-markers in turn.

Suppose *her* was a case-marked pronoun. Could it also be a determiner? Clearly it is highly desirable that it should occupy the same position, relative to the common noun, as a determiner, because it satisfies the rule that a singular countable common noun must have a determiner. Determiners are generally treated either as specifiers of NP (the conservative analysis), or as heads of DP (the more recent analysis introduced by Abney 1987). But the governor of a phrase also governs that phrase's head and specifier (Chomsky 1986: 162), so whichever of these analyses we adopt, the determiner is governed by the governor of the whole phrase; e.g. in (50a), *her* is governed by the inflection of *cost*. But this conclusion leads to a contradiction: the possessive pronoun receives the same Case as the whole NP/DP - e.g. Nominative in (50a) - while also receiving Genitive case from the noun within the phrase. Until this conflict is resolved, we must hold onto the one fact we can be sure of, that possessive pronouns are determiners. There is no evidence that they are also case-marked.

Now suppose that the apostrophe *'s* were a morphological reflex of the Genitive case, as suggested by Chomsky (1986:194). One rather ancient objection to such an analysis is that *'s* is demonstrably affixed to a whole phrase, and not to its head word:

(51a) The girl next door's pictures cost £5.
(51b) Mary and Jane's pictures cost £5.

Whatever it may mean to say that *'s* is a morphological marker, it is clearly not the same kind of morphological marker as (say) the plural *-s* suffix. The status of the *'s* is quite unclear in Chomsky's analysis, because he suggests that it is inserted by rule - presumably an old-fashioned construction-specific transformation - but one assumes that the possessive NP plus *'s* are located in the same position as a determiner; the reason for this assumption is precisely the same as in the discussion of possessive pronouns above, namely that a singular countable common noun is allowed to occur after such a pattern, as though it were a determiner. But again, if NP*'s* is either the head or the specifier of the larger phrase (according to one's analysis of determiners), why doesn't it receive two conflicting cases at the same time? The easiest conclusion is that *'s* is not after all a marker of Case, any more than possessive pronouns are; indeed, a very simple analysis is possible if we take *'s* as a particular kind of possessive pronoun (Hudson 1991:276-82).

Finally, we consider prepositions as putative markers of Case. For Chomsky (1986:194), the word OF is inserted transformationally, so it presumably need not be analysed as a preposition at all, but as a mere Case-marker. The most obvious weakness of this analysis is that OF can be stranded, just like any other preposition.

(52a) Who did you buy pictures of #?
(52b) Which athletes do you must admire the achievements of #?

This is very odd behaviour for a Case-marker, especially if, as Chomsky argues (ibid:199) a moved NP takes its Case with it. This means that in a sentence like (52a), there would be a Case-marked NP (*who*), whose marker was stranded at the other end of the sentence. In the absence of evidence that OF really is a Case-marker, it would be very much simpler to take it as an ordinary preposition.

My conclusion, then, is that none of these three supposed Case-markers in fact has anything to do with either case or Case. In the context of WG there is nothing at all unusual in finding that particular dependencies can be marked on the dependent by selecting some particular word-class (possessive pronoun) or individual lexical item ('S, OF). But this can be said without in any way implying that these markers are themselves examples of case.

8 Conclusion

The general conclusion to which this discussion leads is that the predictions of Word Grammar fit the facts better than do those of Case Theory. The discussion is obviously very tentative because both of these theories are still changing and developing, but my prediction is that as they develop, they will become harder and harder to distinguish, until the point is reached where GB theory turns out to be a version of dependency grammar, just like WG.

One especially important question that remains to be answered is whether it is better to generate government/dependency structures directly, as in WG, or via constituent structures, as in GB. A great deal of theoretical apparatus is needed in GB in order to translate from constituency to government/dependency, via the notions visibility, Case, govern, m-command, barrier, blocking category and minimality. Why not generate these government/dependency structures directly? In order to justify the GB approach, we need evidence that the constituent structure allows analyses and explanations which are not allowed by government/dependency structure alone. This may well turn out to be true, but we need to see the evidence; so far its truth has simply been taken for granted.

References

Abney, S. 1987. *The English Noun Phrase in its Sentential Aspects*. MIT PhD dissertation.

Chomsky, N. 1986. *Knowledge of Language: Its nature, origin and use*. New York: Prager.

Chomsky, N. 1991. Language, politics and composition: a conversation with Noam Chomsky. *Journal of Advanced Composition*, 11. 4-35.

Chomsky, N. and Lasnik, H. 1991. Principles and Parameters Theory. To appear in J. Jacobs, A. von Stechow, W. Sternefeld and T. Vennemann (eds.) *Syntax: An International Handbook of Contemporary Research*. Berlin: Walter de Gruyter.

Fillmore, C. 1968. The case for case. In E. Bach and R. Harms (eds.) *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston, 1-88.

Haegeman, L. 1991. *Introduction to Government and Binding Theory*. Oxford: Blackwell.

Hudson, R. 1984. *Word Grammar*. Oxford: Blackwell.

Hudson, R. 1987. Zwicky on heads. *Journal of Linguistics*, 23. 109-32.

Hudson, R. 1990. *English Word Grammar*. Oxford: Blackwell.

Kiefer, F. 1992. Case. In W. Bright (ed.) *International Encyclopedia of Linguistics*. Oxford: Oxford University Press, 1.217-8.

Larson, R. 1985. Bare-NP adverbs. *Linguistic Inquiry*, 16. 595-621.

Matthews, P. 1981. *Syntax*. Cambridge: Cambridge University Press.

Mel'čuk, I. 1986. Toward a definition of case. In R. Brecht and J. Levine (eds.) *Case in Slavonic.*, Columbus, Ohio: Slavica.

Nichols, J. 1986. Head-marking and dependent-marking grammar. *Language*, 62. 56-119.

Owens, J. 1988. *The Foundations of Grammar: An introduction to mediaeval Arabic grammatical theory*. Amsterdam: Benjamins.

Travis, L. 1989. Parameters of phrase structure. In M. Baltin and A. Kroch (eds.) *Alternative Conceptions of Phrase Structure*. Chicago: University of Chicago Press, 263-279.

Woolford, E. 1991. VP-internal subjects in VSO and non-configurational languages. *Linguistic Inquiry*, 22. 503-40.

Zaenen, A.; Maling, J. and Thráinsson, H. 1985. Case and grammatical functions: the Icelandic passive. *Natural Language and Linguistic Theory*, 3. 441-83.