Under the assumptions of Chomsky 1995, lexical items consist of bundles of features, and syntactic structure is the result of Merge, which links two syntactic items together to form a new one. Together, these two assumptions raise the question of what principles, apart from the linking operation itself, determine the distribution of features in a tree. The first element of an account of the distribution of features is projection: of the two items linked by Merge, one projects, its label becoming the label of the new constituent. Since Chomsky (1995:246) assumes that partial projection is impossible, projections of a head are copies of it, identical to the head in terms of feature composition.

The papers Chomsky 2000 and 2001 introduce a further operation Agree that plays a major role in determining the distribution of features in structures resulting from Merge. Agree operates when an agreement target ("probe"), typically the structural case assigner T or v, enters the derivation with unvalued, uninterpretable features (in the case of T or v, phi-features) and searches its c-command domain for an agreement trigger ("goal") that will supply values for them. Agreement under Agree is linked with casemarking and with movement: in the derivation of Sue seems [t to be happy], for example, valuation of the phi-features of matrix T by those of Sue, casemarking of Sue (i.e. valuation of Sue’s case feature), and movement of Sue to matrix Spec(T), taken to satisfy the EPP feature of T, are all aspects of a single operation.

Agree is designed to implement what is traditionally known as predicate-argument or (argument-)external agreement. With apparent reference to modifier-head or argument-internal agreement, Chomsky (2001:42 (fn. 6)) suggests that “There is presumably a similar but distinct agreement relation, Concord, involving Merge alone.” Carstens (2001) has treated DP-internal agreement in the framework of Chomsky 2000, 2001 in terms of Agree, without postulating a distinct operation Concord. Here, I argue that such an operation is in fact necessary, and that it differs crucially from Agree in feeding projection. I further claim that the traditional distinction between internal and external agreement is in fact a three-way distinction between agreement that precedes theta-marking ("internal"), agreement in raising constructions, which follows theta-marking but precedes case-marking, and agreement that coincides with case-marking ("external"). I explain the observation (Lehmann 1988:57) that internal agreement does not involve person and external agreement does not involve case, showing that agreement in raising constructions may involve either. Finally, I consider the phenomenon of case agreement, which superficially presents the paradoxical appearance of agreement in feature F needing to occur before a value for F is determined. I claim that both internal case agreement and long-distance case agreement in raising constructions receive a maximally simple treatment if nouns are lexically supplied with a variable-value case feature that is propagated both by Concord and Agree before having its value determined by an application of Agree involving a (structural or inherent) case assigner.

The paper begins, in section 1.1, by arguing that the nominal target (goal) of Agree must be D(P) rather than N(P). Section 1.2, building on that result, presents the argument that the operation responsible for internal agreement (Concord) must operate directly after the operation Select, thus feeding Merge. Section 1.3 motivates a restriction on Concord, so defined, to the effect that it may not involve a full argument—that is, an XP that has been theta-marked, and concludes the investigation of internal agreement and the Concord operation. Section 2 investigates the place of internal agreement in the typology of agreement generally, proposing a three-member classification based on derivational stage as defined by theta- and case-assignment. Section 3 considers the problem of case agreement, particularly in raising constructions, reaching the conclusion indicated above. Section 4, finally, is a brief review of the paper’s main claims.
The Operation Concord

1.1 Agree Targets D(P)

In the passage that introduces the probe/goal distinction, Chomsky (2000:122) proposes that in [T be elected an unpopular candidate], the (unvalued) phi-features of T, acting as probe, take as their goal the phi-set of candidate. In this section, I will suggest, as a preliminary to the discussion of Concord below, that Agree targets D rather than N. The evidence for this is that while D must (eventually) include all the features that trigger external agreement, namely features of the categories person, number, and gender, the same cannot be said for N.

Person features, in particular, are clearly associated with D rather than with N (we/you/those linguists). Further, the literature on DP structure (for an overview, see Bernstein 2001) displays a broad consensus that number features inhabit a separate head Num (here Nm). This suggests that the lexically determined (i.e. pre-Concord) distribution of agreement features within the DP is as in (1), where P(erson), N(umber), and G(ender) stand in for feature names.

\[
\text{DP} \quad \text{D[γP]} \quad \text{NmP} \\
\text{Nm[βN]} \quad \text{NP[αG]}
\]

Since the form of D, however, varies in many languages with number and gender (e.g. Italian il/la/i/le "the (ms/fs/mp/fp)"), D must acquire number and gender features before spellout. In contrast to N, then, D will have values for all three types of agreement feature.

The argument of this section is thus as follows: D must include features for all three of the categories person, number, and gender; N, in contrast, apparently lacks person features and, according to widely accepted proposals, number features as well. It follows that D(P) must be the goal of the Agree relation and in that sense the object of casemarking; the alternative would be to force the Agree relation to assemble the agreement features it needs from multiple sources. A second instance of this argument will supply not only evidence that number and gender features must reach D but evidence about the derivational stage at which they must do so. In particular, in a language like English in which anaphor licensing requires both agreement in person, number and gender (I let Jim shave himself/*yourself/*themselves/*herself) and (local) binding, the need for simultaneous satisfaction of the agreement and binding conditions allows us to construct an argument that number and gender features must reach D early enough to project to the DP node.¹²

Consider (2), where [F] abbreviates the gender feature [Feminine].

\[
\text{I let Sue's[+F] brother[-F] shave *herself[-F]/himself[-F]}.\]

The ungrammaticality of the [+F] anaphor in (2) is of course the result of failure of Sue to (c-)command the anaphor node. Yet given that the structure of [Sue’s brother shave himself] is, omitting vP structure, essentially as in (3), with Sue internal to DP₂, brother = NP, and himself = DP₃, it is not immediately clear why the [-F] anaphor should be licensed either.
In particular, the node corresponding to *brother*, the apparent locus of [-F], is at least as deeply embedded in DP₁ as the node corresponding to *Sue*.

In fact, apart from the V node, the only node within VP that commands DP₃, the node corresponding to the anaphor, is DP₁. This means that for the binding and agreement conditions on anaphor licensing to be satisfied simultaneously, DP₁ must in some sense be characterized by the features [αG], [βN], and [γP], to the exclusion of parallel sets of features elsewhere—in particular, internal to DP₂. While in principle we could imagine defining an accessibility relation according to which the features in question, but not their counterparts in DP₂, would be accessible to the DP₁ node without actually being included in it, several related considerations suggest that we should take DP₁ to be characterized by the features in question in the most concrete possible sense—in the sense that those features are part of the node label in question.

First, person features will be part of DP₁ as a result of projection in any event. Further, since, as we have already noted, D must acquire number and gender features before spellout, those features too will reach DP₁ as a result of projection if they arrive at D early enough. Finally, and more generally, the fact that an accessibility relation of the sort in question would replicate in reverse the independently needed notion of projection suggests that we should do without it if possible. Below, then, I will assume that D in (3) acquires number and gender features by Concord and that these project, along with D’ s inherent person features, to DP₁.

2 Concord Feeds Merge

In the last section, we argued that D(P) rather than N(P) must be the goal of the Agree relation, given that only D(P) has all the requisite agreement features. More specifically, we claimed that D must acquire number and gender features by Concord in time for those features to project to DP along with D’ s inherent person features. The indicated timing of Concord, however, is problematic: if that operation is to feed projection, it cannot operate subsequent to Merge, as does Agree. Two possible solutions to this problem are imaginable, the first assuming decomposition of Merge into Link and Project, the second appealing to the operation Select (Chomsky 1995:226), here construed as selecting two syntactic objects to be merged. These two possible implementations of the idea that Concord must feed projection are schematized in (4) and (5), respectively.

\[
\begin{align*}
\text{(4)} & & \text{(5)} \\
& a. & a. \\
& Link & \text{Select} \\
b. & Concord & b. & Concord \\
c. & Project & c. & \text{Merge} (\text{Link} + \text{Project})
\end{align*}
\]

Alternative (5) is more restrictive in predicting that Concord can relate only Select-mates (i.e. eventual sisters): after Select and before Merge, only the Select-mate relation will be available, whereas after Link and before Project, many other relations, including asymmetric c-command and head-head relations, will be available as well. This restriction, in turn, will allow us to predict a well-known fact about internal agreement, namely that it does not involve person features (in
we/you/those intrepid linguists, for example the adjective and noun do not agree in person with D). The reason for this is that if Concord is available only between Select-mates, there will be a way to get features from lower to higher elements (in particular, from lower heads to higher heads and specifiers), namely the iterative application of projection and Concord, but there will be no way to get features from higher to lower elements. If, on the other hand, Concord were available between heads, for example, there would be no principled way to exclude internal agreement in person.

Let us assume, then, a Concord operation that applies after Select and before Merge. I will take Concord to operate on the basis of the same probe-goal distinction as does Agree and will assume that that D has unvalued (uninterpretable) gender and number features as a lexical property and that Nm has an unvalued gender feature. Under these assumptions, the derivation of DP1 in (3) above will proceed as in (6):

\[
\begin{array}{c|c}
\text{Operation} & \text{Output} \\
\hline
\text{Step 1: } \text{Select/Concord} & \text{Select/Concord} \\
\text{Step 2: } \text{Merge} & \text{Merge} \\
\text{Step 3: } \text{Select/Concord} & \text{Select/Concord} \\
\text{Step 4: } \text{Merge} & \text{Merge} \\
\text{Step 5: } \text{Merge} & \text{Merge} \\
\end{array}
\]

To this point, we have considered only instances of internal agreement involving a head X and its complement YP. Internal agreement, however, may target a phrasal projection (i.e. a specifier) as well as a head. In Latin, for example, attributive adjectives (specifiers in the analysis of Cinque 1994) and adjectival possessors display DP-internal agreement just as do demonstratives and numerals. The fact that a merged specifier does not project means that when a specifier is the target of internal agreement, there will be no question of the agreement process feeding projection. This would suggest that the internal agreement of adjectives and possessors could be attributed to Agree, as in the analysis of Carstens 2001, section 5, rather than to the Concord operation proposed here. It turns out that while this is true, it is only true if prior application of Concord is presupposed.

As the probe-goal relations of her diagram (15) show, Carstens (2001:154) needs to assume with Chomsky (2000:122) that both features relevant to agreement of adjectives and possessors (gender and number) are present in N. If, as we have assumed here, gender and number are introduced as properties of two distinct heads, then, in the absence of Concord or some other mechanism of feature copy, agreement of adjectives and possessors in those two features will not be attributable to Agree, since valuation and deletion of the phi-features of a probe under Agree presupposes that the goal possesses all of the requisite features (Carstens 2001:153), following Chomsky).

Internal agreement of adjectives and possessors, then, presupposes Concord, although given that assumption, it could in principle be carried out by Agree. In practice, however, when a ZP rather than a Z0 is selected to merge with [eXYP], Concord, as we have characterized it, will take place between ZP and XP just as it does between Z0 and XP, assuming that the relevant unvalued features are properties of the ZP node. As a result, all internal agreement will in fact be a result of Concord.
3 A Restriction on Concord

Let us now ask whether there are any restrictions that must be placed on the application to
selectmates of the Concord operation we have postulated. I will argue, citing a class of cases
where overapplication of Concord must apparently be prevented, that the operation is not triggered
by a full DP argument—that is, a nominal constituent that has already been theta-marked. As we
might expect pretheoretically, then, Concord is strictly an argument-internal process.

Consider the German example (7), in which two inherently casemarked DPs follow one another.

(7) Ich danke dir für deinen Brief.
I thank you_{fact} for [your letter]_{acc}

In (7), the case of you and the letter are not just inherent in the sense of being associated with
theta-role assignment; they are lexically idiosyncratic properties of thank and for, respectively.
I infer from this first that the cases in question are assigned by the lexical heads (V and P) of
which they are idiosyncratic properties and second (given that both Vs and Ps are known to agree
with their objects) that the V and P in question, like structural case assigners, have empty
phi-features, to be valued as a consequence of the Agree operation.

The argument against unrestricted Concord is then that, at least under the assumption of a
Larsonian structure for an example like (7), the postulation of unrestricted Concord has
counterfactual results. Let us look at how the derivation of (7) will proceed under the assumption
of unrestricted Concord (the phi-features of the prepositional object are represented as [αphi];
case features are omitted).

(8) Operation Output

Step 1: Select/Concord (P_{αphi}, DP_{αphi}) (P_{αphi}, DP_{αphi})
Step 2: Merge (P_{αphi}, DP_{αphi}) PP_{αphi}
Step 3: Select/Concord (V_{αphi}, PP_{αphi}) (V_{αphi}, PP_{αphi})

... The problem is already apparent at Step 3: while we want P to receive the phi-features of DP, we
want this to happen subsequent to Merge, as part of an Agree operation that also results in the
casemarking of DP. We do not want P to receive the phi-features of DP by Concord before Merge,
since this will result in their projecting to the PP node and then being transferred to V by another
application of Concord, as shown in Step 3; V must agree with the dative object, to be introduced
as Spec(VP), and not with the prepositional object.

The partial derivation (8) suggests, then, that the application of Concord must be restricted
in some way. We might imagine that the problematic aspect of applying Concord as indicated in that
derivation is that both P and V are casemarkers and thus have unvalued phi-features. On the other
hand, the need for Concord between D and NmP, as in the derivation (6), would seem to cast doubt
on the idea of restricting Concord so as not to apply between X₀ and YP when X₀ is a casemaker,
given that D is the assigner of structural genitive case. Rather, the appropriate restriction would
seem to be that XP may not be a full argument—that is, may not already have been theta-marked.
I will conclude that this is in fact the correct form of the constraint on Concord that prevents
its application in (8).
4 A Typology of Agreement Processes

We have concluded that Concord is argument-internal, applying between $X^0$ and $YP$ only when $YP$ has no theta-role. Agree, in contrast, arguably applies between $X^0$ and $YP$ only when $YP$ has already been theta-marked; in particular, we argued above that the nominal constituent that acts as the goal of the Agree relation is $D(P)$ rather than $N(P)$. Agree between $X^0$ and $YP$ typically results in casemarking of $YP$, but when $X^0$ is not a casemaker—for instance, when it is participial—Agree applies with no case-related effects (Carstens 2001:152, Chomsky 2001:46(fn.37)). These observations suggest the typology of agreement processes represented in (9), where Theta and Case abbreviate "theta-role already assigned" and "casemarking results" respectively.

\[(9)\] Agreement Types

<table>
<thead>
<tr>
<th>Theta</th>
<th>Case</th>
<th>Operation</th>
<th>Informal Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-$</td>
<td>$-$</td>
<td>Concord</td>
<td>Internal Agreement</td>
</tr>
<tr>
<td>$+$</td>
<td>$-$</td>
<td>Agree</td>
<td>Participial Agreement</td>
</tr>
<tr>
<td>$+$</td>
<td>$+$</td>
<td>Agree</td>
<td>External Agreement</td>
</tr>
</tbody>
</table>

Let us consider the correlation between agreement type and agreement features—that is, features in which agreement is observed. As is well known, while external agreement (potentially) involves person, number, and gender but not case, internal agreement (potentially) involves number, gender, and case but not person (see e.g. Lehmann 1988:57). Above, we proposed an explanation for the fact that internal agreement does not involve person based on the restriction of Concord to Select-mates. The fact that external agreement does not involve case, on the other hand, is naturally seen as the result of casemarking and agreement being two aspects of a single operation: for external agreement in case to be possible, casemarking would have to be capable, counterfactually, of feeding agreement.

Participial agreement of the Indo-European type typified by Icelandic, finally, groups with internal agreement regarding the features it involves.\(^6\) The apparent failure of participles to show person agreement, however, is not accounted for by the restriction of Concord to Select-mates, suggesting that in principle participles should be able to agree in person. The Swahili examples cited by Carstens (2001:150, exx. (5)) appear to exemplify this possibility, since they clearly show that the auxiliary and main verb of a single clause may both show person and number agreement. Further investigation of the agreement properties of participles, however, is called for.

5 Case Concord and Case Agreement

We have seen that Concord, the process responsible for DP-internal agreement, applies incrementally as the DP is built: as new elements are added, they inherit the (number and) gender features of their selectmates. In addition to number and gender, however, DP-internal agreement is typically observed in the category of case, and case differs from number and gender in not acquiring a value until after construction of the DP is complete. The late determination of case would in principle pose no problems for DP-internal concord as long as we assume that (1) every DP-internal head has a case feature and (2) all instances of an unvalued case feature within DP are valued and deleted simultaneously. A second kind of case agreement, however, namely agreement in case between a DP and a participle that assigns it a theta-role but is unable to casemark it, is less straightforward because it occurs over apparently unbounded distances.

Participial agreement is treated in terms of Agree in the framework of Chomsky 2000 and 2001;
from our perspective, this is appropriate, given that the agreement trigger in such cases is a full DP argument. Frampton et al. (2000), however, observe that case agreement of participles raises a problem in the system of Chomsky 2001 (although they discuss possible solutions): in examples where the agreement trigger moves leftward over the participle before being casemarked, the participle will be left caseless at the end of the derivation. For concreteness, let us first look in some detail at a derivation with participial agreement in which this problem does not arise and then consider examples in which it does. In displaying partial derivations, I assume with Carstens (2001:151) that structural casemarkers possess a case feature C and that this feature deletes when the casemaker’s phi-features are valued (this assumption will be simplified below). I further adhere to the following conventions:

(10) a. Interpretable features are shown with their values in square brackets (e.g. Ps[3]).
    b. Uninterpretable features are shown after valuation with their values in parentheses (e.g.
        Ps(3), Cs(N)) and before valuation with empty parentheses (e.g. Ps( ), Cs( )).
    c. Probe-goal relations are shown by asterisks (e.g. Ps( )* ... Ps[3]*).
    d. Space permitting, the input and output of an operation are shown on the same line, with
        the input on the left; changes in the output due to the operation are underlined.

Let us start with example (11) below. In the first relevant step of that derivation, shown as
(11a), the participial head Pt agrees with some arrests, the complement of make, in number and gender
(which I arbitrarily take to be Feminine). In step (11b), finite T casemarks Pt, although it does
not receive the latter’s phi-features because Pt has no person feature and partial agreement is
impossible. Finally, in step (11c), finite T casemarks and agrees with some arrests, which moves
to Spec(T) to satisfy T’s EPP feature: the result is shown in the second line of (11c).

(11) [Some arrests were made t]
    a. Pt make [some arrests] Pt make [some arrests]
       Ps[3] Ps[3]
       Nm( )* Nm[P]*
       Gn( )* Gn[F]*
       Cs( ) Cs( )
    b. T be Pt make [some arrests] T be Pt make [some arrests]
       Ps( ) Ps[3] Ps( ) Ps[3]
       Nm( )* Nm[P]* Nm( ) Nm[P]
       Gn(F) Gn[F] Gn(F) Gn[F]
       C Cs( ) Cs( ) C Cs(N) Cs( )
       EPP EPP
    c. T be Pt make [some arrests] T be Pt make [some arrests]
       Ps( )* Ps[3]* Ps(3) Ps[3]
       Nm( )* Nm(P) Nm[P]* Nm(P) Nm[P]
       Gn(F) Gn[F] Gn(F) Gn[F]
       C Cs(N) Cs( ) Cs(N) Cs(N)
       EPP EPP
In derivation (11), the participle and its DP complement, speaking loosely, are casemarked separately by finite T, the participle first, and no case feature is left unvalued. The same will be true in examples where an expletive is inserted to satisfy the EPP feature of the lowest T and the DP in question remains in situ, as in (12).

\[(12) \quad (\text{On that occasion,}) \quad \text{[there were made some arrests] (that shocked the community).}\]

But in examples like those of (13), the DP complement of the participle will move before a structural casemarker enters the derivation, and when a structural casemarker does enter, it will become inactive in the process of casemarking the DP in question, since all its phi-features will be valued and deleted. As a result, there will be no way to value the participle’s case feature.

\[(13) \quad \begin{align*}
\text{a. } & \quad \text{[Some arrests are certain [t to be made t]}} \\
\text{b. } & \quad \text{We believe [some arrests to be certain [t to be made t]]}
\end{align*}\]

I would like to suggest a solution to this problem based on the intuition that there is in fact no reason ever to have a participle casemarked directly by a case-assigning head or to have a casemarker assign case to more than one element. First, note that there is arguably a conceptual problem with (11b), the step in the derivation above in which the participle is casemarked. The participle is uncontroversially disqualified from valuing any of the phi-features of T, since it lacks a person specification and partial agreement is impossible. But then, if casemarking and agreement are really two aspects of the same operation, in line with Chomsky’s characterization of Agree, it should be equally impossible, under such circumstances, for T to casemark the participle. Casemarking, that is, should be taken to be possible only be possible when agreement is, with the result that no casemarker will be able to assign case to more than one element, and we can take the feature C that characterizes structural casemarkers to delete when case is assigned.

How, under these assumptions, will case be assigned to participles? Given that the observational generalization about participial case is that the participle invariably shows the case that is eventually assigned to the DP that was merged as its complement, the case specification of the participle should in principle be determined by the complement DP in the same way as are its number and gender specifications. Further, agreement in all three features should ideally be the result of the same instance of Agree. This result can be obtained if nouns are listed in the lexicon with a variable case specification Cs(x) along with their specification for gender. The variable case feature will propagate through the DP in parallel fashion to the gender feature and will in that way come to characterize the DP as a whole. In examples like (11)–(13), the first step, that in which the participle agrees with its complement, will take the form (14) rather than (11a) above.

\[(14) \quad \begin{align*}
\text{Pt} & \quad \text{make [some arrests]} \\
\text{Pt} & \quad \text{make [some arrests]}
\end{align*}\]

The specification Cs(x) must of course be treated by the computational system in the same way as
the lack of a specification, so that the DP in question remains active until \( x \) is replaced by a constant case value. At that point, all occurrences of \( x \) will receive the same value, and the agreement of the participle with its complement will be complete.

6 Conclusion

We have reached two main conclusions above. The first is that there is indeed an operation Concord that is distinct from Agree, notably in that it feeds projection. Concord, then, is not simply Agree without casemarking, as in Carstens 2001. Our second main conclusion is that the phenomenon of (nonlocal) participial case agreement motivates the attribution of a variable case specification to the lexical entry of every noun. Such a variable specification, properly interpreted, allows us to capture the fact that with regard to both argument-internal and argument-external (in particular, participial) agreement, case behaves like phi-features even though its value has not yet been determined.

Notes

* I would like to express my gratitude to Andrew Nevins for insightful and stimulating comments on certain of the ideas developed here.

1 LF (or semantic (C-I) interface) licensing of anaphors will dictate a checking account, with overgeneration and filtering, of anaphor-antecedent agreement, given that the phi-features of the anaphor have phonetic consequences and thus must be determined by spellout. This in turn will mean that the phi-features of the anaphor and those of the antecedent must both be considered to be interpretable, since they must survive to the interface for checking purposes.

2 Under the assumptions of fn.1, the need for simultaneous satisfaction of agreement and binding also shows that the results of DP-internal concord must be available at the semantic interface, making it impossible to treat concord as a purely morphological phenomenon (cf. Schütze 1997:163).

3 Serious questions arise here about feature interpretability. We are assuming that [P] is interpretable on D, that [N] is interpretable on Nm, and that [G] is interpretable on N; other occurrences of these features are being assumed to be uninterpretable. In particular, [G] and [N] are being taken to be uninterpretable on D. As indicated in note 1, however, all of D(P)'s phi-features must be considered interpretable at LF for the purposes of verifying the conditions on anaphor binding. I will not attempt to reconcile here the apparent clash between what features are naturally taken to be interpretable internal to the DP and what features must be taken as interpretable for the purposes of binding theory.

4 At the same time, she postulates a head Num; this is problematic, given that Num is generally taken to be the locus of the number feature and not an abstract, featureless position.

5 This would seem to have the consequence that we must assume theta-marking to involve features or otherwise leave a residue of its operation, an assumption that is not clearly consistent with Chomsky's (1995:228) inclusiveness condition.

6 Verb forms identified as participles in other languages may, while retaining casemarking properties and tense contrasts, show no agreement at all (e.g. Dravidian relative participles).

References


