FEATURE COPYING AND BINDING: EVIDENCE FROM COMPLEMENTIZER AGREEMENT AND SWITCH REFERENCE

Akira Watanabe

Abstract. This paper claims (a) that the formal features of the goal are copied onto the probe in the Agree operation, contra Chomsky 2000 and more in line with Chomsky 1998, and (b) that formal features copied onto the probe during the Agree operation participate in binding in a selective way, contra Lasnik 1999. These claims are supported by a modification of Zwart’s (1993, 1997) analysis of complementizer agreement and a minimalist reworking of Finer’s (1984, 1985) theory of switch reference. It will also be shown that formal features deleted during Agree will not be carried anymore when the category containing them undergoes further movement. This assumption plays a key role in providing an account of complementizer agreement more constrained than Zwart’s.

1. Introduction

In recent theorizing on the syntactic computation of human language, the behavior of morphosyntactic features is highlighted as holding the key to understanding the nature of displacement phenomena. These features, called formal features in Chomsky 1995, participate in checking, which is assumed to lie behind movement. Furthermore, Chomsky 1995 proposes that a distinction has to be made among formal features in terms of LF interpretability. This distinction is introduced to explain why some features continue to be available to the syntactic computation even after checking whereas others do not. Those that continue to be available do so because they are needed for LF interpretation anyway. Those that do not continue to be available must be deleted and become invisible before the end of the derivation in order to ensure LF convergence. Thus, interface requirements on the LF side have repercussions for the precise mechanism of feature checking. The exact behavior of formal features during the checking process is currently under debate. This paper explores the nature of feature checking and its relation to binding, by focusing on the switch-reference system and the complementizer-agreement phenomenon.

One immediate goal, therefore, is to improve on Finer’s (1984, 1985) analysis of switch reference within the GB framework by subjecting it to
minimalist scrutiny. Switch reference is revealing because referential dependency between subjects is mediated by the binding relation between functional heads, the locus of feature checking. The result of this investigation is not just a recasting of the GB analysis in terms of the minimalist machinery; it also sheds new light on the mechanism of feature checking and its interaction with binding, an interpretive process. Specifically, I demonstrate that the formal features of the goal are copied onto the probe in the checking operation, contra Chomsky 2000 and more in line with Chomsky 1998, and that formal features copied onto the probe in the checking operation participate in binding in a selective way, contra Lasnik 1999.

An important new insight into the nature of switch reference can be obtained by bringing complementizer agreement, studied in detail by Zwart (1997), into the picture. Consideration of complementizer agreement brings to light the missing links needed for a principled account of switch reference. The leading idea is that essentially the same mechanism lies behind complementizer agreement and switch reference. To the best of my knowledge, these two types of phenomena have never been compared before. The level of understanding achieved by the Minimalist Program, however, makes it possible to approach the apparently unrelated phenomena in typologically unrelated languages from a unified perspective. I also show that an account of complementizer agreement more constrained than Zwart’s becomes possible once we adopt the assumption that formal features deleted during Agree will not be carried anymore when the category containing them undergoes further movement.

This paper is organized in the following way. Section 2 reviews Finer’s (1984, 1985) theory of switch reference, pointing out the respects in which it is wanting from the current theoretical point of view. Section 3 turns to the phenomenon of complementizer agreement and proposes a modification of Zwart’s (1993, 1997) analysis of it. Here, the mechanism of feature checking (or Agree) is explicated. Section 4 considers the theory of switch reference again, now in light of the results from section 3. Section 5 takes up the issues concerning binding and movement. It is shown here that this analysis of switch reference comes into almost direct conflict with Lasnik’s (1999) conclusion that feature movement does not license binding, thereby creating a paradoxical situation. In section 6, however, I suggest that the binding paradox only becomes apparent on a closer look at the relevant cases.

2. Finer’s (1984, 1985) Theory of Switch Reference

Switch reference is the system that regulates referential dependency between the superordinate subject and the subordinate subject by means of a marker in the subordinate clause. There are two such markers, the same-subject marker (SS) and the different-subject marker (DS). Schematically, the possible range of referential dependency between the superordinate subject and the subordinate subject is shown in (1).
When the same-subject marker appears, the two subjects must be coreferential, and when the different-subject marker is present, the two subjects cannot be coreferential.

Finer (1984, 1985) claims that the same-subject marker (SS) and the different-subject marker (DS) in the switch-reference system are located in the C₀ head and that SS is an anaphor whereas DS is a pronominal. The schema in (1) is somewhat inaccurate in that it does not indicate the SS/DS markers on C₀. More precisely, this paper deals with the structures in (2).¹

¹ The head-final nature of the structures in (2) is not crucial for the purposes of this paper, but it is true that the languages discussed by Finer place SS/DS at the end of a subordinate clause. Haiman and Munro (1983:xv) also remark that most of the languages with switch-reference marking happen to be verb-final, though Lenakel, discussed in Lynch 1983, has an SVO order and switch-reference marking.
The Yavapai examples in (3) that Finer (1985:37) took from Kendall 1975 illustrate these structures.

(3)  
   a. [tokatoka-č savakyuva u-t-k] čikwar-kiň
      Tokatoka-SUBJ Savakyuva see-TEMPORAL-SS laugh-COMPL
      ‘When Tokatoka looked at Savakyuva, he laughed.’
   b. [tokatoka-č savakyuva u-t-m] čikwar-kiň
      Tokatoka-SUBJ Savakyuva see-TEMPORAL-DS laugh-COMPL
      ‘When Tokatoka looked at Savakyuva, he laughed.’

According to Finer’s analysis, SS must be bound by the superordinate C\(^0\)/T\(^0\); DS cannot be. When the same-subject marker is used, the subordinate C\(^0\) is bound by the superordinate C\(^0\)/T\(^0\). Given the well-known dependency between C\(^0\) and T\(^0\), this binding relation is translated into dependency between the higher T\(^0\) and the lower T\(^0\). The Spec-head relation in TP then ensures that the subordinate subject must be coreferential with the superordinate subject. The reverse situation holds when the different-subject marker appears. The subordinate C\(^0\)/T\(^0\) cannot be bound by the superordinate C\(^0\)/T\(^0\) in this case. Coreference between the subordinate subject and the superordinate subject is prohibited, because coindexing of the two subjects would result in binding of the subordinate C\(^0\)/T\(^0\) by the superordinate C\(^0\)/T\(^0\) via the Spec-head relation in TP. This system explains the paradigm in (1).

Finer’s theory is highly ingenious, but it is not principled enough from the minimalist point of view. First, the role of the Spec-head relation in translating the dependency between the higher T\(^0\) and the lower T\(^0\) into that of the subjects is simply given as a matter of coindexing and thus is not fully articulated. Coindexing is only a shorthand notation of indicating some contentful relation. In this case, that relation has to do with referential dependency between DP expressions, but Finer’s theory does not specify how the Spec-head relation contributes to it. The same question can be raised about the relation between C\(^0\) and T\(^0\). Second, it does not make much sense to characterize C\(^0\) as anaphoric or pronominal, given that the binding-theoretic notions of being anaphoric and pronominal have to do with referential dependency between DPs, not CPs.

Interestingly, hints as to the solution of these problems come from consideration of complementizer agreement found in some Germanic languages.

3. Complementizer Agreement

Several dialects of Dutch, German, and Frisian display a phenomenon called complementizer agreement, where the complementizer, in addition to the verb itself, shows subject agreement morphology, as in the Groningen dialect examples in (4).\(^2\)

---

\(^2\) The complementizer-agreement data in this paper are drawn from Zwart 1997, unless noted otherwise.
Zwart (1993, 1997) claims that the agreement morphology on complementizers is due to adjunction of Agr$_S$ to C$^0$, following Hoekstra and Marácz 1989. I will base my discussion on his analysis in the 1997 book, which deals with the syntax-morphology relation more carefully.

3.1 Zwart 1997

Zwart 1997 assumes that the verb remains within VP in embedded clauses in the relevant dialects. He claims that in overt syntax, the formal features of the finite verb, FF(V), are adjoined to Agr$_o$, and the resulting complex to T, which in turn moves to Agr$_S$, which then is raised to C. The result is a complex head in (5).

When this complex head is spelled out in the morphology component in the manner proposed by Halle and Marantz 1993, the $\varphi$-features of Agr$_S$ are realized on the complementizer. The finite verb itself contains the trace/copy of its own formal features, too, which guarantees the appropriate verbal inflection.

Zwart’s analysis is a straightforward way of handling the curious agreement phenomenon, but it has at least two problems in light of the articulated theory of checking in Chomsky 1995. A first, perhaps trivial point is the use of Agr heads, which are abandoned in Chomsky 1995 because of their dubious status. The structure in (5) will be reduced then to (6), where T$^0$ carries the $\varphi$-features for subject agreement.
Second, more seriously, Zwart abstracts away from the feature checking involving the subject itself, the determinant of the agreement morphology. Let us consider why this is a significant problem.

Chomsky 1995 suggests that formal features are distinguished in terms of LF interpretability: φ-features of nominals are interpretable, whereas those of V and T are not. Uninterpretable features must be deleted through checking before the end of a derivation or else the derivation crashes (at LF). Now the problem is this. Suppose that the subject is raised to [Spec,TP] in overt syntax, undergoing φ-feature checking with T₀. The φ-features of T₀ then are deleted, becoming inaccessible to the subsequent computation.³ This deletion crucially takes place before raising of T₀ to C₀ because of cyclicity. In other words, the φ-features of T₀ should be inaccessible to the raising operation that applies to T₀. One might argue that the entire category of T₀ can carry whatever features are contained in it, whether checked or not. There are, however, reasons to believe that deleted features are not raised even by category movement, as I will show later. It follows that the source of complementizer agreement must be sought somewhere other than the φ-features of the inflectional head. Furthermore, the same reasoning leads us to rule out the possibility that the formal features of the finite verb are responsible for complementizer agreement.⁴

Zwart (1997:187–188) does suggest that AgrₛP is interpretable at LF if it has a proper pairing of a subject and a verb, but does not elaborate on how this comes about. Put in our terms, the question is how the features on T₀ achieve LF interpretability.

To summarize, Zwart’s analysis of complementizer agreement becomes problematic once checking with the subject is taken into account, together with the articulated mechanism of feature checking based on interpretability. Let us focus further on the nature of feature checking in the next subsection.

³ An exception arises in multiple-specifier cases, which I put aside in this paper.
⁴ If the verb as well as T₀ has φ-features for subject agreement, the system looks redundant. We need to posit φ-features on T₀ for subject agreement, because there would be no probe that checks the φ-features of the subject. If the verb lacks φ-features for subject agreement, on the other hand, questions arise as to how to ensure the correct agreeing form for the verb, which has no access to the informational content of the subject, on Zwart’s assumption that the finite verb remains within VP in overt syntax. One possibility is that the verb can be merged with T₀ under adjacency in the morphology component of the kind posited in Distributed Morphology (Halle & Marantz 1993). One cannot appeal to this option, however, if Dutch is head-initial throughout as argued by Zwart. The reason is that various elements, including direct objects, can intervene between the finite verb and the position assumed for T₀ in Dutch embedded clauses, as shown schematically in (i).

(i) \( C^0 \ subj \ T^0 \ \text{XP}^*V_f \)

In the text discussion, I follow Zwart in assuming that Dutch is head-initial. We are thus forced to assume that the verb as well as T₀ has φ-features for subject agreement.
3.2 The Behavior of Features in Checking

Chomsky (1995) not only introduces the distinction between interpretable and uninterpretable features but also sharpens the idea that checking of formal features plays a central role in displacement phenomena. He argues that the ideal computational system for human language, as far as LF convergence is concerned, would be the one in which only features participate in displacement. In reality, however, we find entire categories displaced. Chomsky (1995:265) tries to reconcile the conceptual expectation and the reality by claiming that category movement is forced by PF considerations and that category movement creates a feature chain as well as a category chain. LF movement, which is free from PF considerations, always involves feature movement, taking the form of adjoining the raised formal features to the head that attracts them and thus creating a feature chain only. Chomsky 1998 pushes the idea further and proposes that feature checking always involves adjunction of the attracted features to the attracting head, whether or not an entire category is displaced at the same time. “Pied-piping” of a category is motivated by the need to shorten the feature chain created, he says.

Chomsky 2000, on the other hand, abandons feature chains altogether and claims instead that feature checking, now called Agree, takes place without displacement of features. Displacement of a category is forced by the generalized notion of EPP features, now available not only for T⁰ but also for C⁰ and for the light verb v. Chomsky 2000 further eliminates the distinction between overt and covert movement altogether, dispensing with the notion of feature strength.

I would like to suggest that the approach of Chomsky 1995, 1998, but not of Chomsky 2000, is on the right track concerning the question of feature adjunction, in view of the problems raised by complementizer agreement. I will nevertheless frame the discussion by using the terminology of Agree, probe, and goal as well as the EPP mechanism of Chomsky 2000 throughout this paper even when earlier proposals in Chomsky 1995, 1998 are discussed, because this more recent work takes up the internal mechanism of feature checking and displacement in detail. I will concentrate on the issue of whether feature copying takes place during the checking operation, mostly putting aside other differences between Chomsky 1995, 1998 and Chomsky 2000.

The difference between the two hypotheses can be illustrated by looking at raising of the subject to [Spec,TP]. The structure before raising takes place is shown in (7), abstracting away from the shell structure within VP.

(7) T⁰ [VP Subj . . .]

The probe in this case is the set of φ-features that T⁰ has. It locates its goal in the VP-internal subject. Under Chomsky’s (1998) approach, the operation
Agree copies the \( \varphi \)-features of the goal onto the probe and deletes the probe, as in (8a). If the EPP feature of \( T^0 \) is to be satisfied by raising the subject, the result is (8b), where combination of Agree and Merge, namely Move, has applied. If an expletive independently satisfies the EPP, we get (8c).

\[
\begin{align*}
(8) & \quad a. \quad [TP \ [T_0 \ [T_0 \ \Phi(T^0) \ ] \ + \ \Phi(subj)] [VP \ Subj \ . . . ] ] \\
& \quad b. \quad [TP \ Subj \ [T_0 \ [T_0 \ \Phi(T^0) \ ] \ + \ \Phi(subj)] [VP \ t_{subj} \ . . . ] ] \\
& \quad c. \quad [TP \ Exp \ [T_0 \ [T_0 \ \Phi(T^0) \ ] \ + \ \Phi(subj)] [VP \ Subj \ . . . ] ]
\end{align*}
\]

Chomsky’s (2000) system does not posit feature copying during the Agree operation. Thus, we have (9), instead of (8).

\[
\begin{align*}
(9) & \quad a. \quad [TP \ [T_0 \ \Phi(T^0) \ ] [VP \ Subj \ . . . ] ] \\
& \quad b. \quad [TP \ Subj \ [T_0 \ \Phi(T^0) \ ] [VP \ t_{subj} \ . . . ] ] \\
& \quad c. \quad [TP \ Exp \ [T_0 \ \Phi(T^0) \ ] [VP \ Subj \ . . . ] ]
\end{align*}
\]

The crucial difference has to do with the feature content of \( T^0 \) that is visible to the computational system. In (8), but not in (9), \( T^0 \) contains \( \varphi \)-features visible to the computational system after Agree has applied.

Now I will show that (8) is the desired result, but there is one question that I have to put aside in this paper—namely, whether uninterpretable features of the goal are also copied under Agree. In (8), the goal is the set of interpretable \( \varphi \)-features of the subject, which continue to be visible to the computational system. It is not clear whether there are cases in which the goal is an uninterpretable feature, and if so, whether that uninterpretable feature is copied onto the probe. An obvious candidate for an uninterpretable goal would be Case, but Chomsky 2000 essentially abandons the idea that Case itself undergoes checking. In the concluding section, I will turn briefly to head movement, which is probably driven by morphological consideration only. Investigation of the entire issue, however, goes beyond the scope of this paper.

### 3.3 Feature Copying and Complementizer Agreement

Recall that the significant problem raised by complementizer agreement is the source of the agreement morphology on \( C^0 \). The critical point is that the \( \varphi \)-features of \( T^0 \), being uninterpretable, will not be raised to \( C^0 \), once deleted.

I propose that the agreement morphology on \( C^0 \) comes from the \( \varphi \)-features of the subject itself, the only possible alternative source. These features are interpretable at LF and thus are visible to the computational system throughout the derivation. They can therefore get raised to \( C^0 \). I continue to assume with Zwart 1997 at the same time that complementizer agreement is due to adjunction of \( T^0 \) to \( C^0 \). In other words, I assume that there is no reason to posit direct \( \varphi \)-feature checking between \( C^0 \) and the subject raised to \([Spec,TP]\).
Now it is clear that we want (8), not (9), at the point where C<sup>0</sup> is merged with TP. The result of merger is the structure in (10), to use the option of subject raising to [Spec,TP].

(10) \[ CP \ C^0 \ [ TP \ Subj \ [ T_0 \ \Phi(T^0) + \Phi(\text{subj})] \ [ VP \ t_{\text{subj}} \ldots]] ]\]

Once the complex head T<sup>0</sup> is raised to C<sup>0</sup>, we get (11).

(11)
\[
\begin{array}{c}
\text{CP} \\
\text{C}^0 \\
\text{TP} \\
\text{C}^0 \\
\text{T}^0 \\
\text{T}^0 \ \Phi(\text{subj})
\end{array}
\]

The complementizer is now associated with the φ-features of the subject, which are going to be realized as agreement morphology.<sup>5</sup>

One indication that the φ-features of complementizer agreement do not come from the inflectional head is the fact noted by Zwart that the agreement morphology of the complementizer is not necessarily identical to the ordinary agreement morphology found on the finite verb. In some of the dialects displaying overt complementizer agreement morphology, different inflectional endings are used on the verb and on the complementizer. That is the case in Lower Bavarian, for example, as shown in (12).

(12) ... das-ma mir noch Minga fahr-n

Note that the features of first person and plural are realized differently on the complementizer and on the verb.

This phenomenon, which Zwart 1997 calls double agreement, can be understood in the following way under the system I am presenting here. Verbal morphology is a PF reflex of the features either on the inflectional head or on the verb itself.<sup>6</sup> These features, once checked, become invisible for the syntactic computation but will be sent to PF for morphological

---

<sup>5</sup> An alternative, suggested independently by Edson Miyamoto and Jan-Wouter Zwart (personal communications), is to say that φ-feature checking in TP is value assignment to the φ-features of T<sup>0</sup>, which are originally undifferentiated and hence uninterpretable. Assignment of concrete values renders them interpretable. Chomsky 2000 suggests something similar for PF purposes, but the alternative in question extends it to LF interpretability as well. It seems to be difficult to tease apart feature copying and value assignment empirically, but value assignment should not be applicable to the φ features of the verb itself, to be discussed shortly. For this reason, I will stick to feature copying in the text discussion.

<sup>6</sup> Recall from footnote 4 that the verb as well as T<sup>0</sup> has φ-features for subject agreement.
realization. Agreement morphology on the complementizer, on the other hand, is due to the interpretable features of the subject copied onto $T^0$ and then raised to $C^0$. Thus, it is not surprising to find languages where these two sets of features receive different forms of morphological realization.

Zwart’s analysis can also explain double agreement. Recall that (5), repeated here, is the relevant structure of $C^0$ under his analysis.

\begin{equation}
(5)
\end{equation}

The verb itself, which remains within VP, has a copy of FF(V). Because the verb and the complementizer have different structures, agreement morphology can be realized differently.

If we take into account parametric differences in double agreement, however, it turns out that our system is more constrained than Zwart’s. Zwart 1997 argues, following Travis 1984, that subject-initial root clauses in Dutch do not involve topicalization of the subject. The finite verb in such clauses is raised to Agr$_S$ and the subject to its specifier, according to Zwart’s analysis. In Lower Bavarian, the finite verb in subject-initial root clauses uses the agreement morphology found on complementizers, as shown in (13b).

\begin{enumerate}
\item[(13)] a. ... das-ma mir noch Minga fahr-n/*-ma
\quad that-1PL we to Munich go-1PL
\item[b.] Mir fahr-ma/*-n noch Minga.
\quad we go-1PL to Munich
\item[c.] Fahr-ma/*-n mir noch Minga?
\quad go-1PL we to Munich
\end{enumerate}

When the verb is raised to $C^0$ as in (13c), it displays the same morphology as complementizers again. In dialects such as Brabantish, however, the finite verb in subject-initial root clauses inflects in the same way as in embedded clauses, as shown in (14).

\begin{enumerate}
\item[(14)] a. ... da-de gullie kom-t/*-de
\quad that-2PL you come-2PL
\item[b.] Gullie kom-t/*-de.
\quad you come-2PL
\item[c.] Wanneer kom-de/*-t gullie?
\quad when come-2PL you
\end{enumerate}
The finite verb inflects like complementizers only when it is raised further to C⁰ as in (14c).

Zwart accounts for this parametric variation by saying that three distinct complex head structures are involved. Omitting irrelevant details about Agr₀ and T⁰, they are:

(15) a. \[
\begin{array}{c}
\text{FF(V)} \\
\text{Agr}_s \\
\text{C}^0 \\
\text{Agr}_s \\
\text{C}^0
\end{array}
\]

b. \[
\begin{array}{c}
\text{FF(V)} \\
\text{Agr}_s \\
\text{V} \\
\text{Agr}_s \\
\text{V}
\end{array}
\]

c. \[
\begin{array}{c}
\text{FF(V)} \\
\text{V}
\end{array}
\]

The tree in (15a) is a simplified version of (5), and (15b) corresponds to subject-initial root clauses. Zwart assumes that formal features move independently from the verb itself, so that when the verb is raised to Agrₚ, both FF(V) and the verb itself are adjoined to Agrₚ. The tree in (15c) is the structure of the verb with φ-features, which remains in VP. Lower Bavarian gives the same morphological realization for φ-features to (15a) and (15b), and a different form to (15c), whereas Brabantish classes (15b) and (15c) together. As Zwart (1997:260) notes, however, there is no dialect known to have three different forms corresponding to (15a,b,c). No explanation for this fact is given under Zwart’s account.

The system presented in this paper, which takes into account interpretability of φ-features, can predict the absence of a language or dialect that uses three different inflectional forms. An important point to make is that there are only two kinds of sources for inflectional morphology if viewed from the perspective of interpretability. The φ-features on T⁰ or V⁰ are uninterpretable at LF, so that once checked, they become inaccessible to the syntactic computation and will be sent to PF. The φ-features of the subject, on the other hand, are interpretable at LF, so that they continue to be accessible throughout the derivation leading up to LF. For morphological realization, a copy of that set of interpretable features must be provided for the PF derivation at the point of Spell-Out, because the original set is needed for LF interpretation. Given that only these two kinds of sources for
morphological realization are available, there cannot be a dialect or language that distinguishes three inflectional forms.

A further conclusion can be drawn about the relation between movement and deletion from the possible patterns of double agreement. The crucial observation is that no dialect is reported where the verb in embedded clauses and the complementizer exhibit the same inflectional morphology to the exclusion of subject-initial root clauses. This fact receives an explanation if deleted features are not carried by movement. To see why, let us be more precise about the mechanism of double agreement under this account. Remember that, following Chomsky 1995, this account does not assume Agr nodes. Schematically, we are dealing with three types of clausal structure shown in (16).

\[(16)\] a. \[\ldots [\text{CP} \, \text{C+T} \, [\text{TP} \, \text{subj} \, t \, [\text{VP} \, \ldots \, \text{V} \, \ldots]]]\]

b. \[[\text{TP} \, \text{subj} \, \text{T+V} \, [\text{VP} \, \ldots \, t \, \ldots]]\]

c. \[[\text{CP} \, (\text{WH}) \, \text{C+T+V} \, [\text{TP} \, \text{subj} \, t \, [\text{VP} \, \ldots \, t \, \ldots]]]\]

Embedded clauses have the structure in (16a), whereas (16b) corresponds to subject-initial root sentences. (16c) represents root questions with verb raising to C^0. Thus, we have (17) instead of (15) as the three relevant complex head structures.

\[(17)\] a. 

\begin{center}
\[\begin{array}{c}
\text{C}^0 \\
\text{T}^0 \quad \Phi(\text{subj}) \\
\end{array}\]
\end{center}

b. 

\begin{center}
\[\begin{array}{c}
\text{T}^0 \\
\text{V}^0 \quad \Phi(\text{V}^0) \\
\text{T}^0 \quad \Phi(\text{subj}) \\
\end{array}\]
\end{center}

c. 

\begin{center}
\[\begin{array}{c}
\text{V}^0 \\
\Phi(\text{V}^0) \\
\end{array}\]
\end{center}

These structures are formed under this analysis as follows. Assume first that the verb as well as T^0 has \(\varphi\)-features for subject agreement. Here, I represent them as right-joined to T^0 or V^0 just for the sake of convenience. (17b), corresponding to subject-initial sentences, is created by copying the \(\varphi\)-features.
of the subject and then raising and checking the $\varphi$-features of the verb. This ordering is forced by consideration of interpretability. The crucial observation is that checking between the uninterpretable $\varphi$-features of $T^0$ and the uninterpretable $\varphi$-features of the verb leads to elimination of both of them. This checking, therefore, cannot take place prior to checking between the subject and $T^0$. Suppose then that checking between the subject and $T^0$ takes place first. Under this operation, the $\varphi$-features of $T^0$ get deleted and those of the subject copied onto $T^0$. Copying is essential here. Notice that if the $\varphi$-features of the subject were not copied onto $T^0$, the $\varphi$-features on the verb would have nothing to agree with at this point, on the reasonable assumption that only an $X^0$ element can act as a probe. The verb is then raised, undergoing checking with the copied $\varphi$-features of the subject, resulting in deletion of the $\varphi$-features of the verb itself. When the verb remains within VP as in embedded clauses, it still has to check its own uninterpretable $\varphi$-features. The ordering of checking is exactly the same as in the case of (17b), except that the verb itself is not raised. The in-situ verb has the structure in (17c).

When $T^0$ is raised to $C^0$, (17a) results. The important point is that the $T^0$ head in (17a) does not contain its own $\varphi$-features that are deleted in checking with the subject within TP. Even when the finite verb is raised to $C^0$ as in (13c) and (14c), it will have the structure in (17a'), assuming that deleted $\varphi$-features of $V^0$ and $T^0$ are not visible to the operation that raises the complex $T$ head in (17b).

(17) a'.

The interpretable $\varphi$-features of the subject are the only source of the agreement morphology found in the $C^0$ area.

Notice that it is only (17b) for subject-initial root clauses which has alternative ways of morphological realization. If the interpretable $\varphi$-features of the subject win, verbal morphology takes the same inflectional endings as (17a). If the uninterpretable $\varphi$-features of $T^0$ or $V^0$ win, the verb shows up in the same form as (17c). There is no way in which (17a) and (17c) receive the same morphological realization in dialects that distinguish interpretable and uninterpretable $\varphi$-features morphologically. Thus, the patterns of double agreement suggest that deleted features are not carried by movement.

Let me stress that this result cannot be obtained without the assumption that deleted $\varphi$-features are not raised by further movement. If deleted $\varphi$-features were carried along, raising of $T^0$ to $C^0$ would be accompanied by the
deleted φ-features of $T^0$ (and those of $V^0$ if $V^0$ is also raised to $T^0$) in (17a,a'), making it impossible to distinguish between (17a,a') and (17b) in terms of φ-feature content. But then, the double agreement pattern of the Brabantish type would not receive an explanatory account. In fact, failure of deleted features to raise is a very plausible assumption, if movement involves making a copy of the item to be raised. If deletion of features means invisibility to syntactic computation, the copying process will also ignore deleted features. The copy to be merged again, therefore, will not contain the deleted features.

To summarize, I have shown that consideration of interpretability leads us to treat complementizer agreement as due to the φ-features of the subject copied onto $T^0$. This modification also turns out to be more constrained than Zwart’s (1997) original analysis in handling the double agreement patterns. Theoretically, the results in this section support the hypothesis that the features of the goal are copied onto the probe under checking and that deleted features are not carried by movement.

3.4 Feature Copying Without Phrasal Movement

Before leaving the topic of complementizer agreement, let us consider the option of (8c), repeated here—in other words, expletive constructions.

(8) c. $[_{TP} \text{Exp } [_{TO} \Phi(T^0)+ \Phi(\text{subj})] [_{VP} \text{Subj }\ldots]]$

Zwart 1997 observes that the paradigms of complementizer agreement are mostly defective, so to study expletive constructions, we need to focus on those languages having distinct complementizer morphology for third person. Fortunately, West Flemish, which has been studied extensively in a series of works by Haegeman, has rich complementizer agreement. Here, I base my discussion on Haegeman 1992. In the third person, singular and plural agreement are distinguished as in (18).8

(18) a. … da Valère morgen goat
    that Valère tomorrow go

b. … dan Valère en Pol morgen goan
    that Valère and Pol tomorrow go

Interestingly, the complementizer inflects in the presence of an expletive, as in (19).

(19) … dan-der nie vele mensen woaren
    that-there not many people were

7 Here, I am abstracting away from the distinction between deletion and erasure (Chomsky 1995), which is made to provide room for checking with multiple specifiers.

8 The final -t in the singular is dropped in (18a) due to a phonological process.
Here, the complementizer displays the form for third-person plural, corresponding to the indefinite plural associate of the expletive. This fact indicates that feature copying takes place without displacement of a phrasal category into [Spec,TP], assuming that the expletive itself is devoid of φ-features.

This result leads us to conclude that the conception of category displacement in Chomsky 1998 must be discarded. There, it is suggested that pied-piping of a phrase including the feature undergoing Agree must accompany the checking operation on the ground that a feature chain cannot extend beyond a minimal domain. That is, phrasal movement is required as a kind of repair strategy that fixes the inadequacy induced by feature movement under this conception. The complementizer-agreement fact in expletive constructions, however, shows that feature movement can take place without an accompanying phrasal movement. Thus, the EPP effect must be guaranteed by a factor somewhat independent of feature copying itself that takes place during the Agree operation.⁹

At the same time, this result paves the way for the model of grammar in which there is no covert movement inaccessible to PF. Since features can be displaced without generalized pied-piping in a way visible to PF, the general mechanism for feature movement in principle will guarantee that every feature checking can take place in “overt syntax.” In this respect, this result is consonant with Chomsky’s (2000) elimination of covert movement.

Let us next return to the problems in Finer’s (1984, 1985) theory of switch reference. The proposed account of complementizer agreement sheds new light on them.

4. Switch Reference Revisited

Recall that the problems in Finer’s theory of switch reference are, first, that the role of the Spec-head relation in translating the dependency between the higher T⁰ and the lower T⁰ into that of the subjects is not given much theoretical content with respect to referential dependency, and second, that characterization of C⁰ as anaphoric or pronominal makes little sense if we take seriously the idea that the properties of anaphoric and pronominal classify nominal expressions.

These problems disappear, however, if the mechanism underlying complementizer agreement is at work in switch reference as well. The φ-features of the subject are first copied onto T⁰ as a result of checking and get raised to C⁰. The binding feature generated on C⁰ then gets associated with the φ-features of the subject and determines referential dependency of these φ-features, in the typical switch-reference configuration given in (20).

⁹ Uriagereka 1999, however, suggests the possibility that the need for repair strategy may be parametrized. In that case, the case from West Flemish will not generalize.
The tree in (20) is an elaboration of (2), with the internal structure of $C^0$ added.

Under this picture, it is not the Spec-head relation within TP that mediates the referential dependency between subjects. The Spec-head relation within TP involves feature checking between $T^0$ and the subject, which results in copying of the interpretable $\varphi$-features of the subject onto $T^0$. The $\varphi$-features of the subject themselves then undergo movement and get combined with a binding-theoretic feature at $C^0$ to establish referential dependency (or its absence). Thus, the properties of anaphoric and pronominal do not characterize $C^0$, strictly speaking. These properties characterize the interpretable $\varphi$-features of a nominal expression, namely, the subject, which are raised to $C^0$. The idea here is that referential dependency is a relation that holds between two sets of $\varphi$-features. In this sense, the dependency between $C^0$'s is the referential dependency between subjects.

5. Binding Paradox

This refinement of Finer’s theory of switch reference at first seems incompatible with Lasnik’s (1999) argument, based on the contrast in (21), that it takes more than just formal features to act as a legitimate binder.

(21) a. Some linguists seem to each other [$t$ to have been given good job offers]
   b. *There seem to each other [$t$ to have been some linguists given good job offers]

In (21b), the $\varphi$-features of some linguists must be raised to the matrix $T^0$ to ensure the correct agreement with the matrix finite verb. But these $\varphi$-features are not sufficient to license an anaphor in the matrix clause, as indicated by the unacceptability of (21b). The entire category must be raised as in (21a) to act as a binder for each other. Chomsky 2000 goes even further to eliminate feature movement in cases like (21b) and claims instead that the operation
Agree can take place without movement between the matrix T<sub>0</sub> as the probe and the φ-features of the expletive associate in situ as the goal. Under this view, it follows from the in-situ status of the φ-features of some linguists that binding is impossible in (21b).

In (20), on the other hand, the φ-features of the matrix subject on the higher C<sub>0</sub> bind the φ-features of the embedded subject on the lower C<sub>0</sub> without involving the entire category of the higher subject, if the same-subject marker appears. As noted by Finer (1985) and Hale (1992), the higher subject in [Spec,TP] does not participate in binding because it fails to c-command into the lower clause. In (3a) repeated here, for example, an R-expression occupies the lower subject position and is coindexed with the higher subject, which would be a Condition C violation if the higher subject were able to c-command into the lower clause. 11

(3) a. [tokatoka-č savakyuva u-t-k] čikwar-kiň
Tokatoka-SUBJ Savakyuva see-TEMPORAL-SS laugh-COMPL
‘When Tokatoka, looked at Savakyuva, he laughed.’

It follows that the higher subject does not bind the φ-features of the lower subject in C<sub>0</sub>, either. The only possible c-commanding antecedent for the φ-features associated with the same-subject marker on the lower C<sub>0</sub>, therefore, is the φ-features raised to the matrix C<sub>0</sub>.

Thus, binding by formal features without an entire category is sometimes possible and sometimes not. What is going on?

6. Binding Paradox Resolved

A closer look at the binding possibilities in switch reference gives us a clue as to when binding by formal features is possible. Consider again (3a) and its structure in (22).
I noted earlier that c-command failure prevents the higher subject from binding the lower subject. The matrix complex $C^0$ with the $\varphi$-features of the matrix subject adjoined, however, does c-command the lower complex $C^0$ and hence also the lower subject. Thus, the lower subject is bound by the $\varphi$-features of the matrix subject in (3a), although not by the matrix subject itself. But (3a) is allowed. The reason why binding by the $\varphi$-features of the matrix subject does not induce a Condition C violation under Finer’s (1984, 1985) theory was that switch reference involves binding between A’ positions, namely $C^0$, so that the lower subject was A-free in (3a) even if it is bound by the $\varphi$-features of the matrix subject. With the A/A’ distinction gone in recent theorizing, we can no longer adopt this explanation. Furthermore, the binder on $C^0$ comes from the matrix subject. In this sense, the binder on $C^0$ is essentially the same entity as the subject itself, forming a chain with (part of) it. It is therefore highly implausible to distinguish between the $\varphi$-features on $C^0$ and the subject itself for binding-theoretic purposes. Some novel explanation for the absence of a Condition C violation is called for.

Notice at this point that the failure of binding of the lower subject by the $\varphi$-features of the matrix subject on $C^0$ in (3a) involves essentially the same structural configuration as the binding failure in (21b), under the assumption that the $\varphi$-features of some linguists get copied onto the matrix $T^0$. In (21b), the potential binder of the anaphor is just a set of $\varphi$-features. The same is true with the lower subject in (3a) (= (22)). The correct generalization then is that a phrase needs a phrasal antecedent as its binder whereas formal features can take another set of formal features as their binder. We can distinguish legitimate from illegitimate cases of binding by formal features by the condition:

(23) $\varphi$-features not accompanied by semantic features cannot serve as the binder of $\varphi$-features accompanied by semantic features.

12 See Abe 1993 in connection with binding.
(23) is a very plausible condition because the antecedent should be able to provide more semantic information than the bindee. A phrase possesses semantic features as well as formal features and thus has more semantic content than a set of formal features. Both in (21b) and in (3a), binding of the maximal projection by φ-features fails because of (23). In (3a), this means that the R-expression in the subordinate subject position is free, satisfying Condition C. In (21b), the anaphor each other is also free, violating Condition A. The φ-features of the lower subject in (22), which are adjoined to C⁰, on the other hand, are legitimately bound by the φ-features of the higher subject, which are likewise raised to C⁰. Thus, (23) correctly predicts when binding by φ-features alone is possible. Simply put, binding by formal features without an entire category is possible when the bindee is also formal features without an entire category, but it is impossible when the bindee is a maximal projection.¹³

It should be noted that φ-features without a category become candidates for bindees only in the switch-reference environments. The reason has to do with the role that binding features play for semantic interpretation. Binding theory, conceived as consisting of interpretive procedures in the manner suggested in Chomsky and Lasnik 1993, goes into effect only in the presence of binding features such as [±anaphoric] and [±pronominal]. These binding features determine referential dependency or absence thereof for φ-features associated with them. The binding feature of the subject is not copied onto T⁰, because it does not participate in checking. The φ-features of the subject copied onto T⁰ are thus not candidates for bindees. It is only when they are raised to C⁰ and associated with binding features there that binding becomes relevant. Binding here means referential dependency of a set of φ-features on another set c-commanding it. (23) is a condition on when binding in this sense holds.¹⁴

It is interesting to observe that Lasnik 1991 proposes a condition similar in spirit to (23), given in (24).

(24) A less referential expression may not bind a more referential one.

Lasnik motivates (24) on the basis of binding possibilities in languages where Condition C does not hold, either entirely or partially. In Vietnamese, for example, an R-expression must be free only in its governing category. Thus, (25a) is acceptable whereas (25b) is not.

¹³ It is not clear whether there are cases where a nominal head must be considered to be a candidate for the bindee, but the prediction is that such a head should pattern with a maximal projection.

¹⁴ It should be noted that the binding feature (represented as SS in (22)) on C⁰ associated with the φ-features to be bound does not count in comparing the semantic content of the binder and the bindee. This may be either simply a matter of stipulation, or due to a default binding-theoretic marking that the C⁰ head not marked as anaphoric nor pronominal receives automatically.
(25) a. John tin John sẽ thắng
    John thinks John will win
b. *John thu’ông John
    John likes John

Nevertheless, an R-expression cannot be bound by a pronoun or an epithet in a higher clause, as shown in (26).

(26) a. cái thắng chó de’ tin cái thắng chó de’ sẽ thắng
    the son-of-a-bitch thinks the son-of-a-bitch will win
b. *cái thắng chó de’ tin John sẽ thắng
    the son-of-a-bitch thinks John will win
c. *Nó tin John sẽ thắng
    he thinks John will win

Lasnik gives (27) as a summary of the relevant data.

<table>
<thead>
<tr>
<th>Binder</th>
<th>Pronoun</th>
<th>Epithet</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>pronoun</td>
<td>ok</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>epithet</td>
<td>ok</td>
<td>ok</td>
<td>*</td>
</tr>
<tr>
<td>name</td>
<td>ok</td>
<td>ok</td>
<td>ok</td>
</tr>
</tbody>
</table>

The condition in (24), which is most likely to be universal, captures the generalization found in (27).

Despite the similarity, (24) cannot be a special case of (23), however. (23) determines the presence or absence of binding, whereas (24) rules out some instances of binding as ill-formed. In other words, I claim that (23) is part of the definition of binding. This point is significant in (3a). If (23) were a condition that simply marks certain instances of binding as ungrammatical, as does (24), (3a) would be ungrammatical on the ground that the embedded subject were bound by the φ-features of the matrix subject raised to C₀. The reason why (3a) is allowed is that binding in this case does not hold in the first place. In (26), on the other hand, the embedded subject is bound by the matrix subject, and the presence of this binding relation leads to a violation of (24) in (26b,c). Thus, consideration of degrees of referentiality does not enter into calculation of semantic content as far as (23) is concerned. Otherwise, binding would not hold in any of the examples in (26), making it impossible to account for the contrast. The condition in (23) simply cares about the presence or absence of semantic features.

Another way of looking at the difference between (23) and (24) is to focus on the notion of binding again. If binding is a dependency holding between two sets of φ-features, questions about referentiality do not arise at this point.
φ-features have to do with such things as person, number, and gender. A given set of φ-features is no more or less referential than another. Referential content is provided by semantic features. Given this conceptual clarification, (24) should perhaps more appropriately be formulated as follows:

(24’) If the φ-features of a less referential expression bind those of a more referential one, discard that interpretation.

(24’) presupposes that the definition of binding is already given. By contrast, (23) is part of the definition of binding.

To summarize, I have proposed (23) as part of the definition of binding. Binding by formal features without an entire category holds just in case the condition in (23) is met.

7. Conclusion

I have shown that assuming feature copying during the Agree operation is crucial for a principled account of complementizer agreement and for a minimalist theory of switch reference. In both cases, the φ-features of the subject that are copied onto T^0 and get raised to C^0 play an essential role in filling the theoretical gaps of the previous accounts. This result, in turn, leads us to examine binding relations in the switch-reference system more carefully. This investigation has shown that the status of a potential binder must be taken into account in the definition of binding. Specifically, for α to bind β, α must have at least as much semantic content as β. It then follows that binding by formal features without an entire category is possible when the bindee is also formal features without an entire category, but not when the bindee is a maximal projection.

It has also been shown that formal features deleted during Agree will not be carried anymore when the category containing them undergoes further movement. This is a natural consequence of the idea that uninterpretable features are marked during checking as invisible to the computation leading to LF. This hypothesis is crucial in explaining the possible patterns of complementizer agreement. Recall also that complementizer agreement in West Flemish provides evidence that features alone can be copied without accompanying phrasal movement.

The discussion of the mechanism of Agree was limited to the case where the goal of the Agree operation is interpretable φ-features of the subject. To conclude the discussion, let us briefly consider other types of movement.

Cheng 1991 observes that wh-in-situ languages tend to have overt clausetyping particles, which obviate the need of overt wh-movement to type the clause as interrogative, in the position of C^0. Hagstrom 1998 argues extensively that these particles are raised from wh-in-situ in the form of head movement. Watanabe 1992, on the other hand, claims that wh-in-situ in Japanese involves movement of a phrasal null operator in overt syntax. Under
the proposed view of feature checking, these two approaches to wh-in-situ become compatible with each other, because movement of a phrase into [Spec,CP] can induce a morphological change on the shape of the C\textsuperscript{0} head by means of feature copying.\textsuperscript{15} The clause-typing interrogative particle can be a morphological realization of the copied feature. In other words, what Hagstrom perceives as head movement may actually be feature copying. Given that feature copying can take place without phrasal displacement, as shown by the expletive construction in West Flemish, it is an open question whether phrasal movement must be posited for wh-in-situ, as in Watanabe’s approach.

There are also languages like Irish (McCloskey 1979) where overt wh-movement triggers a morphological change on the C\textsuperscript{0} head. These languages are candidates for cases where overt phrasal wh-movement is accompanied by feature copying that is morphologically visible.

Turning finally to head movement, it has been assumed (Baker 1988 and Lasnik 1981, among many others) that head movement is driven by the affixal nature of either the raised head or the hosting head. But then, head movement is not driven by LF-interpretable features. It is therefore not surprising if no evidence for feature copying is found. In this sense, head movement is somewhat different from the other types of movement. Nevertheless, it has the properties of movement taking place in the computational system leading up to LF, if Watanabe 1999 is on the right track. It is left to future investigation to explore the nature of head movement further.

References


Akira Watanabe  
Department of English  
University of Tokyo  
7-3-1 Hongo  
Bunkyo-ku  
Tokyo 113-0033  
Japan

h32140@m-unix.cc.u-tokyo.ac.jp