Anaphoric INFL*

Jun Abe

0.   Introduction

In the literature there has been much discussion about the status of PRO in the Government and Binding theory.   Roughly, there are three positions in this respect:

(i) Chomsky (1981) takes PRO as a pronominal anaphor. From the fact that an element could not satisfy binding principles A and B in (1) at the same time, it would follow that PRO have no governing category and that, to do so, it must be ungoverned.

(1)   *Binding Theory*

(A) An anaphor is bound in its governing category

(B) A pronominal is free in its governing category

(C) An R-expression is free

(Chomsky   (1981:   188))

(ii) Bouchard (1984) takes PRO as either a pronominal or an anaphor; it is regarded as an anaphor when bound locally and as a pronominal when bound nonlocally.

(iii) Lebeaux (1984) takes PRO as an anaphor. He assumes:

(2)   a.   PRO is a necessarily dependent element, and must be bound in a local domain D (to be specified).

b.   PRO may be either A- or A'-bound; it is unspecified for the position of its binder.

Chomsky's position may be preferred to the other two in that PRO's distribution (i.e., PRO is ungoverned) follows from independently motivated principles, i.e., the binding principles.   This position may, however, be undesirable in that PRO's locality in binding cannot be captured under the binding theory since this element must not have its governing category, whereas the other two positions can capture such locality under the binding theory: according to Bouchard's position, as either an

anaphor or a pronominal, PRO may have its governing category, in which it must be A-bound or A-free, and according to Lebeaux's position, as an anaphor, it may have its governing category, in which it must be A- or A'-bound. Lebeaux (1984-5) notes similarities between PRO and overt anaphors such as reciprocals or reflexives: they are either *locally* or *nonlocally* bound:

(3) a. John likes himself (*local*)

   b. John knew that some pictures of himself would be on sale (*nonlocal*)

(4) a. John tried PRO to leave (*local*)

   b. John had forgotten that PRO shaving himself took so much time (*nonlocal*)

(Lebeaux (1984-5: 343))

He further claims that, when nonlocally bound, they "allow split antecedents, do not require s-structure c-command by the antecedent, appear in free variation with pronouns, and so on;" for the relevant data, see Lebeaux (1984-5: 345-51). Lebeaux (1984) also argues against Bouchard's position, according to which PRO may act as a simple pronoun. Consider the following examples:

(5) a.* PRO_i having shaved already shows that Mary arrived more than 5 minutes after $John_i$ did.

   b. His_i having shaved already shows that Mary arrived more than 5 minutes after $John_i$ did.

(Lebeaux (1984: 254))

While PRO "cannot in general look arbitrarily far down the tree for its antecedent," (*ibid.*: 254) pronouns can be coreferential with the antecedent in such a position. This seems to me to be a serious problem for advocates of Bouchard's position.

   In this paper, we will pursue Lebeaux's position and try to refine it. Our main point is that PRO is unnecessary and Infl takes its place. We thus assume:

(6)   Infl may be an anaphor.

In Section 1, we will outline Lebeaux's (1984) analysis and point out a few problems. Section 2 deals with details of the assumption (6). In Section 3, we will postulate our version of the binding theory and see how our system, together with the assumption (6), successfully explains the behavior of anaphoric Infl.   In Section 4, we will take up apparent counterexamples to our binding theory and solve them from a cross-linguistic point of view.

1.   Lebeaux's (1984) Analysis

Lebeaux (1984) assumes that PRO is an anaphor and, as such, is subject to binding principle A in (1), though he further assumes that PRO's relevant local domain of the binding theory is not its *governing* category but rather its *binding* category; *binding category for PRO* is defined as follows:

(7)   Binding category: β is the binding category for PRO
        if β is the minimal NP, S dominating γ , where γ
        is the minimal S' dominating PRO.

                                              (Lebeaux (1984:   264))

He further assumes the following:

(8)   Closure: If PRO is unbound in its binding category,
        adjoin e_i coindexed with PRO_i to the binding category.

(9)   Operator interpretation: An *e* in an A'-position, not
        bound by an A'-antecedent, is construed as an operator
        O.   If O_i is coindexed with an antecedent NP_i,   inter-
        pret it with the reference of that NP. Otherwise,
        interpret it as free (i.e. arbitrary), or, in   marked
        cases, with a pragmatically picked out referent.

                                                            (*ibid.*)

This system accounts for the following array of the relevant data:

(10)   a.* John would like PRO$_{a,b}$ to leave.

b. * John killed someone PRO$_{a,b}$ to get insurance money.

c. * Bill $_i$ said that John killed someone PRO$_i$ to get insurance money.

(11)   a. John admitted that (O$_i$(it was uncalled for (PRO$_i$ to buy oneself that))).

b. John$_i$ admitted that (O$_i$(it was uncalled for (PRO$_i$ to buy himself that))).

In (10a-b), the binding category for PRO is the matrix S, which is the minimal S dominating the minimal S' dominating PRO, i.e., *PRO to leave* in (10a) and *PRO to get insurance money* in (10b), and hence it cannot be arbitrary since a possible antecedent, *John*, exists within its binding category. In (10c), PRO cannot take *Bill* as its antecedent since a possible antecedent, *John*, exists within its binding category, i.e., the embedded S. In (11), there is no possible antecedent of PRO within its binding category, i.e., the embedded S, so an operator O is adjoined to the binding category and A'-binds PRO. The operator itself may or may not be coindexed with *John*, as shown in (11a-b).

Notice that, in this analysis, PRO is regarded as an anaphor. One may well wonder why PRO appears only in the subject position of a nonfinite clause in English, French, etc. Recall that Chomsky (1981) takes PRO as a pronominal anaphor and that, by virtue of the binding theory, its distribution is restricted: it appears only in ungoverned positions, which amount to subject positions of nonfinite clauses in these languages. But, under Lebeaux's analysis, in which PRO is taken as only an anaphor, its distribution does not follow at all. Why is PRO's distribution so constrained? This calls for an explanation.

The second problem with Lebeaux's analysis concerns the definition of the binding category for PRO. He assumes (7) as the binding category for PRO but it seems to be ad hoc and lack generality in that it only applies to PRO and does not apply to

any other anaphors. In other words, though he claims that PRO is an anaphor and, as such, is subject to binding principle A, he treats it differently from the other anaphors as to what domain is relevant to the binding principles. This is obviously undesirable and uniform treatment is asked for.

2. Infl as an Anaphor

In the previous section, we pointed out two problems with Lebeaux's (1984) analysis. This section concerns the first problem: why does PRO appear in the subject position of a nonfinite clause in English, French, etc.?

First consider θ -marking of the subject of a clause. It has been generally assumed so far that, in (12), the subject *John* is θ -marked by the whole VP *like Mary* since the semantic role of the subject is determined compositionally, depending on the meaning of the unit V-NP; for details of this discussion, see Chomsky (1986a: 59-60).

(12)   ($_{IP}$ John ($_{I'}$ Infl ($_{VP}$ like Mary)))[1]

On the other hand, Chomsky (1986b) considers the possibility of assuming the following condition on θ -marking:

(13)   The θ -marker and the recipient of the θ -role must be sisters.

(13) applies properly to θ -marking of a zero-level category; for example, in (12), the V *like* and its complement *Mary* are sisters and hence the former assigns a θ -role to the latter. It also applies properly to θ -marking of a derived nominal; in (14) the N *refusal* and its complement *the offer* are sisters[2] and hence the former assigns a θ -role to the latter, and furthermore the N' *refusal of the offer* and its specifier *John's* are also sisters and hence the θ -marking from the former to the latter.

(14)   John's ($_{N'}$ refusal of the offer)

However, (13) does not hold for θ -marking of subject by VP. In (12), I' dominates VP but not the subject and hence *John* and the VP *like Mary* are not sisters.[3]

Recently some generative grammarians propose that the subject of VP is in the specifier position of the VP at D-structure as in (15a) and that the subject moves into the specifier position of the IP in order to get Case, as in (15b):[4]

(15)   a.   ($_{IP}$ ϕ ($_{I'}$ Infl ($_{VP}$ *John* ($_{VP}$ like   Mary))))

        b.   ($_{IP}$ *John$_i$* ($_{I'}$ Infl ($_{VP}$ t_i ($_{VP}$ like   Mary))))

If this is the case, then (13) wins full generality; since VP and its specifier position are sisters, the former can assign a θ -role to the latter under condition (13). In (15b), the moved NP *John* and its trace t_i form a CHAIN in the sense of Chomsky (1986a). A CHAIN is defined to include a chain, the S-structure reflection of a "history of movement," as shown in (16a), and an expletive-argument pair, as shown in (16b):

(16)   a.   *John$_i$* was killed e_i

        b.   *There$_i$* is a *man$_i$* in the room

The general condition on a CHAIN is as follows:

(17)   If C = (α_1,   ...,α_n) is a maximal CHAIN, then α_n occupies its unique θ -position and α_1 its unique Case-marked   position   (Chomsky (1986a: 137))

In (16), *John$_i$* and *there$_i$*, the initial elements of a CHAIN, are assinged a nominative Case by the matrix Infl, and are in a θ '-position; on the other hand, e_i and a *man$_i$*,   the   final elements of a CHAIN, are assigned a θ -role by the verbs *killed* and *is*, respectively, and are not assigned any Case. Therefore, both the CHAINs in (16) are well-formed in accordance with (17). Now consider the case of a CHAIN in (15b). *John$_i$*,   the   initial element of a CHAIN is assigned a Case--a nominative Case if Infl is (+TENSE), an objective Case from the outside if Infl is *to*-- and is in a θ '-position; note that the VP *like Mary* cannot

assign a θ-role directly to *John$_i$* because of condition (13). On the other hand, t_i, the final element of a CHAIN, is assigned a θ-role by the VP *like Mary*, as suggested above, and is not assigned any Case. Thus the CHAIN in (15) is well-formed.[5]

Next consider the case where the subject in (15) is empty, aside from the case where it is empty due to NP- or *wh*-movement. Most generative grammarians assume PRO to exist in this position, in accordance with the Extended Projection Principle, proposed by Chomsky (1982), which requires that sentences have subjects. We claim instead that PRO is unnecessary and subject positions may be "literally empty." Thus, when the subject in (15) is empty, the structure is as follows:[6]

(18)   [$_{IP}$ *Infl* [$_{VP}$ like   Mary]]]

If no elements could not receive a θ-role from the VP, (18) violates θ-criterion, as (19) does:

(19)   *There liked Mary.

Suppose that in (18), since the VP and *Infl* are sisters, the former assigns a θ-role to the latter. It follows then that *Infl* itself is an argument in this case. However, note that it lacks an inherent reference like reflexives or reciprocals. If we assume, following Aoun (1985), that if it lacks an inherent reference it is an anaphor,[7] then *Infl* in (18) is an anaphor. To recapitulate:

(20)   If Infl possesses a θ-role, it is an anaphor.

If this analysis is correct, it predicts with respect to a "small clause" that its subject must be overt since it lacks Infl which could receive a θ-role assigned by its predicate. This is borne out by the following data:

(21) a.   Mary saw (John run)

     b.   Mary believes (John   honest)

(22)   a.*Mary saw (run)

b. *Mary believes (honest)ª

In (21), the predicates *run* and *honest* can assign a θ-role to *John* in accordance with condition (13) since they are sisters; in (22), the predicates *run* and *honest* cannot assign a θ-role to, say, *Mary* since it violates condition (13).

Recall that Lebeaux (1984) takes PRO as an anaphor. We have claimed above that PRO is unnecessary and subject positions may be "literary empty," as shown in (18), and that, in (18), *Infl* is an anaphor. In other words, in our system, Infl takes the place of PRO as an anaphor. One of the greatest advantages of this change is that the distribution of anaphoric Infl is restricted to the position which can only receive a θ-role from VP, one which should be assigned to a subject. Therefore, the question why PRO appears only in a subject position does not arise since PRO is substituted for by anaphoric Infl. However, another question still remains: why can't PRO appear in the subject position of a tensed clause? or, to put it in our terms,why does anaphoric Infl behave differently when (+TENSE); we will answer this problem in Section 4.

3. How to Observe Binding Principle A

In the previous section, we assumed that Infl can receive a θ-role from the VP and that, if it does, it is an anaphor. In this section, we will see how anaphoric Infl observes binding principle A. To do so, we will adopt a definition of the binding category with full generality and try to account for the behavior of anaphoric Infl in the same way as we do for that of the other anaphors.

3.1. The Definition of the Binding Category

We pointed out in Section 1 that one of the problems with Lebeaux's (1984) analysis is that the definition of the binding category lacks generality and that PRO is different from the other anaphors as to what domain is relevant to the binding

principles. We will pursue the possibility of uniform treatment below.

We assume a version of Aoun's (1985) Generalized Binding Principles, which apply to NPs whose antecedents are not only in A-positions but also in A'-positions: [9]

(23)   *Generalized Binding Principles*

A.   An anaphor must be X-bound in its binding category.

B.   A pronominal must be X-free in its binding category.

C.   A name must be X-free.

(where X = A or A')

We assume *binding category* as follows:

(24) β is a *binding category* for α iff β is the minimal maximal category containing α, α's Case-assigner and a SUBJECT accessible to α.

A SUBJECT is, in a sense, a possible antecedent of α. We then define a SUBJECT as follows:

(25)   a.   A SUBJECT is a nominal head or specifier.

b. α's SUBJECT is not less prominent than either α or its Case-assigner.

(26)   a. α is more prominent than β if α m-commands β, and not vice versa.[10]

b.   If α and β m-command each other:

a head $>$ a specifier $>$ a complement

We further assume *accessibility* as follows:

(27) α is *accessible* to β iff β is in the c-command domain of α and coindexing of (α, β) would not violate the *i-within-i* condition (28).

(28) $^{*}(\gamma_i \ldots \delta_i \ldots)$
where δ is not γ's head

Chomsky (1981) classifies Infl in English as follows:

(29) a. (+TENSE, AGR)

    b. (−TENSE)

It is clear that (29a) is projected into the tensed clause in English. Chomsky assumes (29b) to be projected into the to-infinitive alone in this language, but it is more natural to consider that it is also projected into the gerundive and participial clauses. Moreover, he claims that the element AGR is nominal since it has the features person, gender and number, like PRO, and that the Infl of the tensed clause is thus qualified as a SUBJECT. We will accept this claim, but how about the Infls of (−TENSE)?; are they nominal or not? We will use (P__) as a diagnosis of "nominality." It follows that the Infl of the gerundive clause is nominal since it appears in the environment (P__), as illustrated in (30), whereas the Infl's of the infinitival and participial clauses are not.[11]

(30)   I am against his/him running again.

To sum up, qualified as a SUBJECT are the Infl's of the tensed and gerundive clauses.

Note that in (24) what is relevant to a binding category is not a governor, as is usually assumed in the literature, but a Case-assigner. This change is trivial in the case where the Case-assigner of α is also its governor, but it is of significance in the other cases; consider the following data:[12]

(31) a. *($_{IP}$ we$_i$ Infl ($_{VP}$ t$_i$ ($_{VP}$ preferred ($_{NP}$ each
             (AGR)

        other's (ING read the   book)))))

    b. ($_{IP}$ we$_i$ Infl ($_{VP}$ t$_i$ ($_{VP}$ preferred ($_{IP}$ each other
             (AGR)

    (ING read the book)))))

If we took a governor as relevant to a binding category, we could not give a natural explanation of the difference in grammaticality between these sentences above; for the binding

category for *each other* would naturally be considered to be the same in both the sentences since its accessible SUBJECT is, by definition, ING (which is more prominent nominal element than *each other* in the sense of (26b)) in both cases and its goveror is naturally considered to be the same because of the uniformity of each structure in (31); it seems ad hoc to say that N governs its specifier position but I dose not, or vice versa. On the other hand, in our system, we can naturally account for the difference if we adopt Reuland's (1983) analysis with respect to the ACC(usative)–ING construction, according to which in (31b), *each other* receives an objective Case from the matrix verb *preferred* through ING. In (31a), the binding category for *each other* is the embedded NP since its Case-assigner is, to say the least, in the embedded NP (it is obvious that the matrix verb *preferred* is irrelevant to the Case-marking of *each other*) and its accessible SUBJECT is ING. In this category, *each other* is A-free, violating principle A. In (31b), the binding category for *each other* is the matrix IP since its Case-assigner is the matrix V *preferred* and its accessible SUBJECT is the matrix Infl (note that ING is not the reciprocal's accessible SUBJECT since it is less prominent than the reciprocal's Case-assigner). In this category, *each other* is A-bound by the trace t_i of we_i; hence the grammaticality of the sentence.

Next consider the case where the *i*-within-*i* condition (28) works (irrelevant details are omitted):

(32) a. we_i Infl $(_{VP}$ t_i $(_{VP}$ saw $(_{NP}$ each other's pictures)))

b. we_i Infl $(_{VP}$ t_i $(_{VP}$ thought $(_{CP}$ (that) $(_{IP}(_{NP}$ pictures of each other) Infl would be on sale))))
                                                        (AGR)

If, following Huang (1983), we assume:

(33) The referential index of a head N comes (i.e. percolates) from the maximal NP node.

                                                        (Huang (1983: 558))

(32a) can be naturally accounted for as follows: the embedded
NP is not the binding category for *each other* since, with the
head N *pictures* and its maximal NP coindexed, coindexing of the
head N and the reciprocal would violate condition (28) and thus
the head N is not a SUBJECT accessible to the reciprocal.   Hence
the binding category for the reciprocal is the matrix VP, which
contains its accessible SUBJECT, i.e., the subject trace t_i
(which is more prominent than both the reciprocal and its Case-
assigner in the sense of (26a)). In this category, *each other*
is A-bound by the trace of *we*, satisfying binding   principle   A.
Next, if we further assume, following   Chomsky   (1981):

(34)   AGR is coindexed with the NP it governs

(Chomsky   (1981:   211))

we can give a natural account of (32b); we have seen above that
the   head   N   *pictures*   cannot   serve   as   an   accessible   SUBJECT   to
each other because of the violation   of   condition   (28); moreover,
the embedded IP contains a Case-assigner of *each other*   but   not
its   accessible   SUBJECT   since,   with   the   embedded   AGR   and   its
subject   coindexed,   further   coindexing   of   the   embedded   AGR   and
each other would   violate   the   *i*-within-*i*   condition   (28).   The
matrix   VP,   by   contrast,   contains   *each   other's*   accessible
SUBJECT (the trace of *we*) as well as its Case-assigner and hence
counts as the binding category   for   the   reciprocal.   In   this
category it is A-bound by the trace of *we*   and   the   sentence   is
correctly marked grammatical.

In   the   next   section,   we   will   see   how   anaphoric   Infl
occurring   in   various   circumstances   observes   binding   principle   A
under the system proposed above. To   fulfill   this,   we   will
extend   the   notions   of   binding   category   (24)   and   a   SUBJECT   (25)
as   follows:

(35) β is a binding category for α iff β is the
minimal maximal category containing α , α 's   Case-
assigner if it is   not   a   Case-assigner   itself   and   a

SUBJECT accessible to α.

(36) α can be a SUBJECT for α itself iff α is the most prominent element in α's m-command domain.

Then, in a maximal projection XP such as:

(37) $(_{XP}\ \alpha\ (_{X'}\ X\ \beta\ ))$
where X is a head, α is X's specifier and β is X's complement

X can be the SUBJECT for X itself as well as α or β since it is a head and, by definition (26b), the most prominent element in XP. α can only be the SUBJECT for β, but neither for X nor for α itself since it is X's specifier and more prominent than β, but it is not the most prominent element in XP and less prominent than X.

3.2. How to Observe Binding Principle A
In this section we will see how the behavior of anaphoric Infl is accounted for by the theory just outlined.
First consider the case where anaphoric Infl is *to*:

(38) a. I_i Infl $(_{VP}\ t_i\ (\text{want}\ (_{IP}\ to_i\ \text{leave})))$
b. I_i Infl $(_{VP}\ t_i\ (\text{told Mary}_j\ (_{IP}\ to_j\ \text{leave})))$
c. I_i Infl $(_{VP}\ t_i\ (\text{promised (Mary)}\ (_{IP}\ to_i\ \text{leave})))$

Assuming that *to* is a possible Case-assigner because it is prepositional in character,[13] *to*'s Case-assigner is irrelevant to the choice of its binding category according to (35). *To*'s accessible SUBJECT is not *to* itself since it is not a nominal element; since t_i is more prominent than *to* in the sense of (26a), it is *to*'s accessible SUBJECT. Thus the binding category for *to* is the matrix VP, in which *to* is A-bound by t_i, the trace of *I* in (38a) and (38c), and by *Mary* in (38b).[14]
In (38), the *to*-infinitives are subcategorized by the verbs and hence are considered to be in the VPs. In this case, the antecedent of *to* can be either the matrix subject or object

since in either case the antecedent c-commands *to*. On the other hand, if a relevant clause is outside the VP, then only the matrix subject can be an antecedent of the anaphoric Infl. The relevant data are the following:

(39) a. $(_{IP}(_{IP}$ To$_i$ spend some time by herself/*himself), Ann$_i$ Infl $(_{VP}$ sent Ned to NY)) (AGR)

      b. $(_{IP}(_{IP}$ ING$_i$ Cursing himself/*herself), John$_i$ Infl (AGR)

      $(_{VP}$ left Mary))

In (39a), the binding category for *to* is the matrix IP since it contains *to's* accessible SUBJECT (the matrix Infl). In this category, *to* is A-bound by *Ann* and hence *herself* is allowed in the *to*-infinitive; by contrast, *Ned* does not bind *to* and hence *himself* is not allowed in the *to*-infinitive. As for (39b), if we assume, following Reuland (1983), that the participial ING has a nominative Case to assign to its subject, its Case-assigner is irrelevant to the choice of its binding category. The participial ING cannot be an accessible SUBJECT to itself since it is not a nominal element. Then, the matrix Infl is ING's accessible SUBJECT since it is more prominent nominal element than the ING. Thus the binding category for the ING is the matrix IP, in which it is A-bound by *John*, but not by *Mary*; hence *himself* is allowed in the participial, but not *herself*.

Next consider the case where anaphoric Infl is the gerundive ING:

(40) John discussed $(_{NP}$ ING scratch himself/oneself)

The verb *discuss* takes only the POSS(essive)-ING construction, but not the ACC-ING construction, as illustrated below:

(41) The investigating committee discussed $\left\{ \begin{array}{l} \text{the army's} \\ \text{*the army} \end{array} \right\}$

      killing Asian civilians for no reason.

(Horn (1975: 380))

It follows that in (40), the gerundive ING is a potential possessive Case-assigner, and hence its Case-assigner is irrelevant here. Its accessible SUBJECT is itself, since it is the most prominent nominal element in the NP. Hence the binding category for the ING is the embedded NP. However, the ING is free in this category and the sentence should be marked ungrammatical, but the fact is different.

We assume that the greatest peculiarity to anaphoric Infl is that it may be either A- or A'-bound; by contrast, the other anaphors are only allowed to be bound in one way. We have so far seen the case where anaphoric Infl is A-bound. The case where it is A'-bound is that of O operator-binding in the sense of Lebeaux (1984); see Section 1. We will assume his rule of *closure*, adapting it in accordance with our system:

(42)   Closure: If anaphoric Infl has no possible binder in its binding category, adjoin e_i coindexed with this Infl $_i$ to the binding category.

(43)   Operator Interpretation: An e in an A'-position, not bound by an A'-antecedent, is construed as an operator O. If O_i is coindexed with an antecedent NP_i, interpret it with the reference of that NP. Otherwise, interpret it as free (i.e. arbitrary), or, in marked cases, with a pragmatically picked out referent.

Thus, since the gerundive ING has no possible binder in its binding category in (40), operator O_i coindexed with the ING_i is adjoined to the category:

(40') $John_{i/j}$ discussed $[_{NP} O_i [_{NP} ING_i$ scratch himself/ oneself]]

In (40'), ING is A'-bound by O, satisfying binding principle A. Since this operator may be coindexed with *John* or may not according to (43), both *himself* and *oneself* is allowed in the POSS-ING complement.

Note incidentally that in the case where anaphoric Infl is A-bound in its binding category, it cannot have an arbitrary interpretation since a possible antecedent exists within its binding category and hence it prevents operator O from being adjoined to the category:

(44) a. *I$_i$ Infl ($_{VP}$ t$_i$ (want ($_{IP}$ to$_{arb}$ leave)))

     b. *I$_i$ Infl ($_{VP}$ t$_i$ (told Mary ($_{IP}$ to$_{arb}$ leave)))

     c. *I$_i$ Infl ($_{VP}$ t$_i$ (promised (Mary) ($_{IP}$ to$_{arb}$ leave)))

     d. *($_{IP}$($_{IP}$ To$_{arb}$ spend some time by oneself), Ann Infl sent Ned to NY)

     e. *($_{IP}$($_{IP}$ ING$_{arb}$ Cursing oneself), John Infl left Mary)

Next consider the case where anaphoric Infl is ING of the ACC-ING construction. The verb *imagine* takes not only the POSS-ING construction but also the ACC-ING construction, as shown below:

(45) We imagined $\begin{Bmatrix} \text{Fred's} \\ \text{Fred} \end{Bmatrix}$ singing old songs.

Hence (45) will be ambiguous when the gerund is "subjectless":

(46) We imagined singing old songs.

Horn (1975) observes that extraction from POSS-ING complements is forbidden whereas extraction from ACC-ING complements is permitted. Therefore, (46) can be disambiguated by extracting an element from the gerundive complement, and as a result, the complement is regarded as an ACC-ING construction:

(47) What did we imagine singing?

The relevant strcuture of (47) will be the following:

(47') what did ($_{IP}$ we$_i$ Infl ($_{VP}$ t$_i$ ($_{VP}$ imagine ($_{IP}$ ING$_i$
                         (AGR)

     sing)))

Assuming, following Reuland (1983), that ING of the ACC-ING construction receives a Case from the outside and transmits it to its subject, in (47'), the ING's Case-assigner is the matrix verb *imagine*. Its accessible SUBJECT cannot be itself since it is less prominent than the ING's Case-assigner in the sense of (26a). Likewise, t_i is not the ING's accessible SUBJECT since it is less prominent than the ING's Case-assigner in the sense of (26b). On the other hand, the matrix Infl can serve as the ING's accessible SUBJECT since it is more prominent nominal element than both the ING and its Case-assigner. Thus the binding category for the ING is the matrix IP and in this category it is A-bound by t_i, the trace of *we*, satisfying binding principle A. It follows that ING of the ACC-ING construction does not have an arbitrary interpretation, like the anaphoric Infl's in (44). This is borne out by the following data:

(48)   What did we imagine singing
$$\left\{ \begin{array}{l} \text{i. *as being fun for some people.} \\ \text{ii. but were afraid to try.} \end{array} \right\}$$
(Horn   (1975: 382))

Consider next the case where the *to*-infinitival clause appears in the interrogative clause:

(49)   John asked how to behave himself/oneself.

Since the anaphoric Infl *to* may have an arbitrary interpretation in (49), which allows *oneself* to occur in the interrogative complement, it must have no possible binder in its binding category so as to resort to O operator-binding. Then, the relevant structure of (49) would be the following:

(49')   John$_{i/j}$ asked [cp O$_i$ [cp how [ϕ [ip to$_i$ behave himself/oneself]]]]

Recall that we have used (P__) as a diagnosis of "nominality" in Section 3.1. If we apply this diagnosis to the interrogative

clause, this clause passes it, as shown below, and hence is regarded as nominal.

(50)   She wasn't sure about how to move the table.

It follows that in (49'), the head ϕ of the interrogative complement is considered nominal. Thus the binding category for *to* is the embedded CP since it contains *to's* accessible SUBJECT (the head ϕ of the embedded CP, which is more prominent nominal element than *to*). In this category *to* has no possible binder and hence operator O is adjoined to the category. As a result, *to* is A'-bound by O in its binding category, satisfying binding principle A.

Finally consider the following data:

(51)   John said that Mary thought that to feed himself/
        herself/oneself would be impossible.

This example will favor our "anaphoric-Infl" analysis rather than the "anaphoric-PRO" analysis. If we took the latter's analysis, the relevant structure of (51) would be the following:

(51')   John said that Mary$_i$ Infl* $(_{VP*}$ t_i (thought that
                              (AGR)

        ((PRO $_i$ to feed   herself/himself/oneself)$_i$ Infl$_i$
                                                            (AGR)

        would be impossible))

In (51'), PRO's accessible SUBJECT cannot be the most deeply embedded Infl since with this Infl and its subject coindexed under rule (34), coindexing of the Infl and PRO would violate the *i*-within-*i* condition (28). Instead, t_i serves to be PRO's accessible SUBJECT since it is more prominent than PRO in the sense of (26a). Hence the binding category for PRO is the embedded VP*, in which it is A-bound by t_i, the trace of *Mary*. It follows, then, that PRO should take only *Mary* as its antecedent. But the example (51) shows that this is not the case.

On the other hand, if we take our own analysis, the problem

will disappear; the relevant structure of (51) will be the following:

(51″) John said that Mary thought that [$_{IP}$ O$_i$ [$_{IP}$(to$_i$ feed herself/himself/oneself) $_i$ Infl$_i$ would be (AGR) impossible])

In (51″), to's accessible SUBJECT can be the most deeply embedded Infl; even with this Infl and its subject coindexed under rule (34), coindexing of the Infl and to would not violate the i-within-i condition (29) since to is the head of the i-indexed subject. Thus the binding category for to is the most embedded IP, in which it is free, and hence operator O is adjoined to this category; in the category to is A'-bound by O, satisfying binding principle A. It follows that, since O may be coindexed with *Mary, John* or none, all of *herself, himself* and *oneself* are allowed to occur in the *to*-infinitive.

4. The degenerate Infl and AGR

We have so far shown how we can account for the behavior of anaphoric Infl appearing in various environments under our system. In this section, we will consider two apparent counter-examples in English and try to get them over by reexamining the relevant data from a cross-linguistic point of view and by assuming the notion of the degenerate Infl and AGR.

One counterexample to our analysis is what we call the "S'-Deletion case," as shown below:

(52) a. I believed [$_{IP}$ myself to know her)
      b. *I believed [$_{IP}$ PRO to know her)

Under a standard GB framework such as Chomsky's (1981), (52b) is explained as follows: the *believe*-type verb is marked in that it can take IP (whether underlyingly or by a deletion rule) and hence is allowed to exceptionally govern and Case-mark the subject of the IP, while the other verbs take CP, which serves

to be a barrier to its government of the subject of IP. There-fore, the ungrammaticality of (52b) follows from the PRO Theorem, according to which PRO is ungoverned (see the Introduction). On the other hand, under our system, (52b) should be ruled in; its relevant structure will be the following:

(52b') $*I_i$ Infl $(_{VP} t_i$ (believed $(_{IP} to_i$ know her)))

In (52b'), the binding category for *to* is the matrix VP since it contains *to*'s accessible SUBJECT (t_i). In this category *to* is A -bound by t_i, the trace of *I*, satisfying binding principle A, and hence (52b) should be grammatical with the reading in which *I* is the understood subject of *to know her*.

Another counterexample to our system is the case where the subject is underlyingly empty in the tensed clause:

(53) *The men say that PRO eat at ten o'clock.

If we assume the PRO Theorem, then (53) will be ruled out since the subject position of the tensed clause is that assinged a nominative Case by the Infl and must be governed for Case-assignment. Under our system, by contrast, (53) cannot be ruled out; its relevant structure will be the following:

(53') *the men$_{i/j}$ say that $(_{IP} O_i (_{IP}$ Infl$_i$ eat at ten
             (AGR)

o'clock))

In (53'), the binding category for Infl$_i$ is the embedded IP since it contains Infl$_i$'s accessible SUBJECT (Infl$_i$ itself because it is the most prominent nominal element in this IP); note that Infl$_i$ itself is a nominative Case-assigner and hence its Case-assigner is irrelevant here. Infl$_i$ is free in   this category and operator O is adjoined to the category; hence Infl$_i$ is A'-bound by O, satisfying binding principle A. Then (53) should be grammatical with the reading in which the understood subject of *eat at ten o'clock* is either *the men* or arbitrary (or pragmatically picked out referent).

Interestingly, the two couterexamples (52b) and (53) *are* actually acceptable with the very readings suggested above in other languages; compare (52b) with (54) and (53) with (55):[15]

(54)　Je croyais la connaître.　　　　　　　　　(French)
　　　'I believed to know her. (i.e. I believed that I knew her.)'

(55)　Los hombres$_i$ dicen que$\phi_{i/j}$ comen a las diez.
　　　　　　　　　　　　　　　　　　　　　　　(Spanish)

　　　'The men say that (they) eat at ten o'clock.'
　　　　　　　　　　　　　　　　　　(Suñer (1982: 60))

It is, then, natural to consider that the structures of (52b) and (53) should be grammatical in universal grammar, but that in English they are ungrammatical because of its marked properties to be clarified below.

We assume that English has the following two properties:

(56)　a.　degeneration of Infl
　　　b.　degeneration of AGR

As for (56a), it is useful to take up Borer's (1986) claim that "the rule of "S'-deletion" associated with particular lexical entries (e.g. *believe* and *seem*) may be dispensed with at no empirical cost, and that the range of phenomena explained by S'-deletion may be explained by appealing to the notions of I-subject and privileged relationship." She proposes the following rule of Universal Grammar:

(57)　Coindex NP with Infl in the accessible domain of Infl.[15]　　　　　　　　　　　(Borer (1986: 375))

The NP in (57) is named *I-subject*. She further assumes that:

(58)　The I-subject, once coindexed with Infl, cannot enter any assignment relations with any other Case-assigner.　　　　　　　　　　　　　　　　(ibid.: 403)

She calls this relationship between Infl and the I-subject a

privileged relation. She proposes that "the obligatory (NP,S) position required by the Extended Projection Principle should be replaced by an *obligatory coindexing between Infl and some NP.*" This means in our system that:

> (59)   Infl must be coindexed either by (a) rule (57) or (b) X-binding when it is an anaphor.

With these as premises, we assume, following Borer, that the *believe*-type verb is different from other complement-taking verbs in selecting a clause with a "degenarate Infl" node. This Infl is degenerate in the following respect:

> (60)   The degenerate Infl does not undergo coindexing.

From (60) it follows that the subject of the complement the *believe*-type verb takes is not an I-subject and that it does not bear the privileged relation (58) with its Infl node. Consequently, the subject can be assigned a Case by the matrix verb. Furthermore, in (52b'), repeated here as (61) with indices changed, to, receiving a θ-role from the VP, to is an argument and hence is an anaphor.

> (61) $^{*}I_i$ Infl $(_{VP}$ t_i (believed $(_{IP}$ to know her)))

However, since *to* is a degenarate Infl, it cannot undergo coindexing. Consequently, the anaphoric *to* cannot have its binder, violating principle A; hence the ungrammaticality of (61).

(56b) is the property which what we call non-null subject languages have; these languages are ones which have an impoverished agreement system. We assume two kinds of agreement features: ϕ-*features* (i.e., the features person, gender and number) and *Case-features*. We further assume that the following principle holds in the grammar:

> (62)   *Agreement Licensing*
>   AGR must be licensed in the way in which either its

ϕ - or Case-features are realized.

AGR's ϕ -features are realized on the verb which is attached to it,[17] while its Case-features are realized on the NP it is coindexed with (or, strictly speaking, the Infl containing it is coindexed with). Furthermore, we assume that non-null-subject languages have degenerate AGR; degeneration of AGR means that:

(63)   AGR's ϕ -features are impoverished and their realization does not suffice for AGR to be licensed.

In other words, we are assuming that it is ϕ -features, not Case-features that are degenerate or impoverished in these languages.

With this in mind, consider the following data:

(64) a. Él_i $[_{IV}$ $\text{llegó}_i]$ t_i
                    (AGR)

      b. Ellos_i $[_{IV}$ $\text{llegaron}_i]$ t_i
                    (AGR)   (Spanish)

(65) He_i $[_{IV}$ $\text{arrived}_i]$ t_i
                    (AGR)   (English)

In (64), AGR's ϕ -features are realized: in (64a), *llegó* has the features of third person and singular, and in (64b), *llegaron* has the features of third person and plural. There- fore, AGR's in (64) are licensed. Notice that AGR's Case- features are also realized in (64): *él* in (64a) and *ellos* in (64b) are both the NPs coindexed with AGR under rule (57) and are realized with a nominative Case. In (65), on the other hand, *he*, coindexed with AGR under rule (57), has a nominative Case, whereas *arrived* realizes no ϕ -features at all. But this is enough to satisfy agreement licensing (62). Next consider the pro-drop case:

(66) a. $[_{IP}$ O_i $[_{IP}[_{IV}$ $\text{Llegó}_i]$ $t_i]]$
                        (AGR)

      b. $[_{IP}$ O_i $[_{IP}[_{IV}$ $\text{Llegaron}_i]$ $t_i]]$
                            (AGR)

(67) *$[_{IP}$ O_i $[_{IP}[_{IV}$ $\text{arrived}_i]$ $t_i]]$
                        (AGR)

In all the sentences above, since the subjects of the VPs are absent, the θ-roles the VPs should assign to them are assigned to the Infl's in our system and, as a result, the Infl's serve to be anaphors. Since these anaphoric Infl's are free in their binding categories, operator O is adjoined to each binding category. Hence these Infl's are A'-bound by O's in their binding categories, satisfying binding principle A. According to operator interpretation (43), O is interpreted either as arbitrary or with a pragmatically picked out referent. In (66), AGRs' ϕ-features are realized (though its Case-features are not) and hence the AGR's are licensed; since *llegó* and *llegaron* have the features of third person, singular and third person, plural, respectively, O's must be interpreted with referents with such features. In (67), by contrast, neither AGR's ϕ- nor Case-features are realized and hence the AGR is not licensed; hence the ungrammaticality of (67). The same explanation holds true for the contrast between (53) and (55).[18]

5.   Concluding Remarks

In this paper we have claimed that PRO is unnecessary and that Infl should take its place. Adopting Lebeaux's (1984) position, we have then assumed that Infl may be an anaphor and that, if it is, it is either A- or A'-bound. One of the great advantages of this change is that the distribution of anaphoric Infl is restricted to the position which can only receive a θ-role from VP, one which should be assigned to a subject position and hence that the question why PRO appears only in a subject position does not arise while it is a crucial question to Lebeaux (1984). Another advantage is that, with this change, we can explain the behavior of anaphoric Infl under our fully general version of the binding theory, which also applies to the other anaphors.

Notes

* An earlier version of this paper was presented at the

59th annual meeting of the English Literary Society of Japan. I would like to thank Hiroaki Tada, Toshifusa Oka, Yuji Takano and Manabu Hashimoto for their valuable comments and criticism. Remaining errors are my own.

[1] Chomsky (1986b) assumes that the X-bar system shown in (i) extends to nonlexical categories such as the clausal categories conventionally labeled S and S' and that these categories have the internal structures (iia) and (iib), respectively:

(i) a. $X' = X\ X^{"*}$

   b. $X" = X^{"*}\ X'$

(ii) a. $S = I" = (NP\ (_I\ (_{VP}\quad V\quad ...)))$

   b. $S' = C" = (...(_C\cdot\quad C\quad I"))$ (Chomsky (1986b: 3))

We assume (i) and (ii) throughout this paper. Using conventional notation, we will represent the maximal projections C" and I" as CP and IP, respectively.

[2] We are assuming, following Chomsky (1986a), that *of* in *of the offer* in (14) is a mere realization of genitive Case and not an independent category, say P.

[3] In fact, Chomsky (1986b) defines "sisterhood" as in (i) so that θ-marking of subject by VP can observe condition (13).

(i) α and β are sisters if they are dominated by the same lexical projections.   (Chomsky (1986b: 13))

For criticism of this claim, see Fukui (1986).

[4] For such subject-raising analyses, see Zagona (1982), Kitagawa (1985), Johnson (1985), Koopman and Sportiche (1985), Kuroda (1985) and Fukui (1986). We assume, following Stowell (1983), that, for what we call a "predicate," its external argument appears in an adjoined position of its predicate, as shown below:

(i) $(_{XP}\ NP\ (_{XP}\quad ...))$

With this assumption, we could restrict a possible θ-marker in

V's projections to a head or a maximal projection.

⁵ Note that we have claimed that the specifier position of IP, being always a θ'-position, is only needed as the initial member of a CHAIN in order to supply Case. It follows then that this position is never occupied by an argument at D-structure and is occupied only by an application of Move α. Defining an A-position as the position in which an argument may appear at D-structure, Fukui (1986) concludes that the specifier position of IP is an A'-position and that the A/A'-distinction and the θ / θ'-distinction completely overlap. I cannot calculate here exactly what results from this assumption; I will leave this matter to future research.

⁶ We assume, following Chomsky (1986b), that in the case where the specifier position is missing in a maximal projection X″, the structure is (ib) rather than (ia):

(i)   a. $[_{XP}[_{X'}\ X\ ...]]$
       b. $[_{XP}\ X\ ...]$

⁷ In fact, Aoun (1985) defines an anaphor as follows:

(i)   An element is an anaphor (= has an incomplete matrix) if it lacks an inherent reference (reflexives, reciprocals) or if it does not satisfy the necessary condition for phonetic realization (NP-trace, wh-traces, PROs).   (Aoun (1985:   94))

Aoun proposes an original version of the binding theory, called Generalized Binding Theory, which generalizes A-binding theory so as to subsume A'-binding relation such as that of a wh-phrase and its trace. Therefore, in his system, wh-trace is treated as an anaphor.

⁸ In relation to (21b), (i) is still ungrammatical though the embedded predicate could apparently assign a θ-role to its Infl.

(i)   *Mary believes (to be honest)

For what we call an "S'-deletion" case such as this, see Section 4.

[9] In Aoun's (1985) system, variables such as *wh*-traces are also anaphors and as such must be A'-bound in their governing categories. In what follows, we will confine ourselves to the analysis of A-anaphors and anaphoric Infl.

[10] Following Chomsky (1986a), we define *m-command* as follows:

(i) α m-commands β iff α does not dominate β and every maximal projection γ that dominates α dominates β .

[11] One may claim that if subordinate conjunctions are reasonably assumed to be prepositions, Infl(+TENSE) passes the diagnostic test, as shown below:

(i) It will be some time [$_{PP}$ before [$_{IP}$ we know the full results]]

In (i), the IP *we know the full results* appears after the P *before* and hence the head of the IP, i.e., Infl(+TENSE), can be considered to be a nominal.

[12] For an expository purpose, we will follow Horn (1975) in assuming that the POSS-ING construction has the categorial status of NP whereas the ACC-ING has the categorial status of IP. For details of the relevant discussion, see also Abe (to appear).

[13] For the argument for the prepositionality of *to*-infinitives on syntactic and semantic grounds, see Abe (1986).

[14] The question why *to* is controlled by an object in such verbs as *tell* and by a subject in such verbs as *promise* is beyond our points. For the relevant discussions, see Saito (1985) and Williams (1980).

[15] In fact, in (55), the understood subject of *comen a las diez* is only either *los hombres* or pragmatically picked out referent, but not arbitrary. This is because *comen* has the

features of third person and plural and hence it requires its antecedent to have such features by the effect of agreement. For the relevant discussion, see below.

[16] She assumes *the accessible domain of Infl* as follows:

(i) α is in the accessible domain of $Infl_i$ iff $Infl_i$ c-commands α and there is no β_j, β_j I-subject   of $Infl_j$, such that $Infl_i$ c-commands $Infl_j$ and $Infl_j$ c-commands α .   (Borer   (1986:   375))

[17] It is immaterial to our analysis whether a verb is attached to an Infl or vice versa, but for an expository purpose, we follow Chomsky (1986) in assuming the former. We will also ignore the question of when this "verb-raising" should take   place.

[18] Note that we are assuming that in our system there is no empty subject base-generated at D-structure and that the θ -role which should be assigned to a subject is assigned to an Infl when the subject is not overt. With this assumption, we can dispense with the postulation of PRO and pro for the most part; that is, we need not postulate PRO or pro as an empty subject. However, we cannot say that there is no empty element, whether a subject or an object, base-generated   at   D-structure;   Japanese apparently has an empty pronominal appearing in an object position; moreover, Rizzi (1986) pointed out the existence of null objects in   Italian. Even in English, there   is   an   empty element occurring in an object position at D-structure, as shown below:

(i)   John tried $(_I$ to be loved ϕ by everyone)

In   (i), ϕ must be in this position in order to receive a θ -role assigned by the verb *loved*; otherwise, the violation of the Projection   Principle. Assuming   the   visibility   condition proposed by Joseph Aoun, according to which an argument must have a Case to receive a θ -role, ϕ cannot get a θ -role since it has no Case. Assuming that Case can be assigned to

only an overt element, ϕ can be salvaged in no way since it is empty.

One may claim that it can be visible by forming a CHAIN with *to*, which is a potential Case-assigner and can supply a Case to have ϕ receive a θ -role. Then, the appropriate representation of (i) will be the following:

(ii) John tried ($_{IP}$ to$_i$ be loved ϕ_i by everyone)

In (ii), the CHAIN (to$_i$, ϕ_i) lacks an inherent reference and hence is an anaphor. Consequently, it searches for a possible binder in its binding category in the way suggested above.

References

Abe, J. (1986) "Prepositional *To*-infinitives in English," *English Linguistics* 3, 87–106.

––––––––––. (to appear) "Generalized Binding Theory and the Behavior of Anaphors in Gerunds," *English Linguistics*.

Aoun, J. (1985) *A Grammar of Anaphora*. Cambridge, MA: MIT Press.

Borer, H. (1986) "I-Subjects," *Linguistic Inquiry* 17, 375–416.

Bouchard, D. (1984) *On the Content of Empty Categories*. Dordrecht: Foris.

Chomsky, N. (1981) *Lectures on Government and Binding*. Dordrecht: Foris.

––––––––––. (1982) *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.

––––––––––. (1986a) *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.

––––––––––. (1986b) *Barriers*. Cambridge, MA: MIT Press.

Fukui, N. (1986) *A Theory of Category Projection and its Applications*. Ph.D dissertation, MIT.

Horn, G.M. (1975) "On the Nonsentential Nature of the POSS-ING construction," *Linguistic Analysis* 1, 333–87.

Huang, C.T.J. (1983) "A Note on the Binding Theory," *Linguistic*

Inquiry 14, 554-561.

Johnson, K. (1985) *A Case for Movement.* Ph.D dissertation, MIT.

Kitagawa, Y. (1984) "Clauses," unpublished manuscript. University of Massachusetts.

Koopman, H. and D. Sportiche (1985) " θ -theory and Extraction," paper presented at GLOW meeting.

Kuroda, S.-Y. (1985) "Whether You Agree or Not: Rough Ideas about the Comparative Grammar of English and Japanese," Unpublished manuscript. University of California, San Diego.

Lebeaux, D. (1984) "Anaphoric Binding and the Definition of PRO," in P. Jones and C. Sells (eds.) *Proceedings of the Fourteenth Annual Meetings of NELS.* 253-74. Amherst: GLSA, University of Massachusetts.

----------. (1984-5) "Locality and Anaphoric Binding," *The Linguistic Review* 4, 343-63.

Reuland, E.J. (1983) "Government and the Search for AUXes: A Case Study in Cross-linguistic Category Identification," in F. Heny and B. Richards (eds.) *Linguistic Categories: Auxiliaries and Related Puzzles.* Vol.1 99-168. Dordrecht: D. Reidel.

Rizzi, L. (1986) "Null Objects in Italian and the Theory of pro," Linguistic Inquiry 17, 501-57.

Saito, S. (1985) *Control and Binding in English.* M.A. Thesis, University of Tsukuba.

Stowell, T. (1983) "Subject across Categories," *The Linguistic Review* 2, 285-312.

Suñer, M. (1982) "On Null Subjects," *Linguistic Inquiry* 9, 55-78.

Williams, E. (1980) "Predication," *Linguistic Inquiry* 11, 203-38.

Zagona, K.T. (1982) *Government and Proper Government of Verbal Projections.* Ph.D dissertation, University of Washington.

Institute of Literature and Linguistics
University of Tsukuba