Covert Syntax for Anaphoric Interpretation*
Yoshihisa Kitagawa, Indiana University

Abstract

In this paper, I will propose and argue for the view that non-deictic proforms are introduced into syntactic representations as "referentially underspecified" entities, and that they must undergo a covert computational operation "Affect" at LF, as required by the Principle of Full Interpretation. While the proposed approach is developed under the leading ideas of the minimalist program, we will take significant departure from the current Chomskian view and suggest that good portion of anaphoric interpretation should still be dealt with in covert syntax with the assumption that the driving force of "Affect" is not limited to the need for the checking of "formal features". In the process of arguing for such an approach, we will also point out that it is not necessary to postulate a special class of entities called expletives in UG.

1. Introduction ................................................................................................................................. 1
2. Proposals ...................................................................................................................................... 6
3. Motivation ..................................................................................................................................... 15
   3.1 Attract ...................................................................................................................................... 15
      3.1.1 The Background .................................................................................................................. 15
      3.1.2 Attract Locative .................................................................................................................... 20
      3.1.3 Attract Clause ....................................................................................................................... 44
      3.1.4 Theoretical Consequences: ................................................................................................. 53
   3.2 Identify .................................................................................................................................... 56
   3.3 Copy .......................................................................................................................................... 59
4. Summary and Conclusions .......................................................................................................... 69
1. Introduction

We all know that a pronoun can be used "deictically" (or "referentially"), that is, to
directly select an entity in the physical or mental context, with either physical or mental
ostensity as in (1).¹

(1)  a. [ Pointing at John, someone says: ]
    The executives are satisfied with his performance.
   b. [ Reading a newspaper article on Ted Kaczynski, someone says to himself: ]
    He is insane.

The pronoun here is more or less on a par with a name in (2a-b) below in that its referential
content is identified without any recourse to its relation to some other linguistic expression
(Lasnik (1976)).

(2)  a. The executives are satisfied with John's performance.
   b. Ted Kaczynski is insane.

Deictic pronouns as well as names, in other words, show up in syntactic representations fully
specified in the sense that the identification of their referential content does not require any
syntactic computation. Let us describe this state of linguistic expressions as "referentially
saturated" — the notion we will define more precisely in Section 2 below.

We also know that a pronoun can be used non-deictically and be interpreted by way
of an "anaphoric" relation with a linguistic antecedent. In (3a-b) below, for example, we can
identify the referential content of his without need for imagining any particular speech context.

(3)  
   a. John is satisfied with his (own) performance.
   b. [ The executives who hired John ] are satisfied with his performance (and raised his salary).

If we continue not to associate these sentences with any specific speech context, and in addition, somehow fail to identify the "linguistic antecedent" of the pronouns, as indicated by xxx in (4a-b) below, the referential content of his remains unidentified, and the full interpretation of these pronouns must await the unveiling of xxx. Imagine, for instance, a situation in which you overheard someone utter one of these sentences when you passed in front of his or her office and some noise precluded you from catching the name for xxx:

(4)  
   a. xxx is satisfied with his (own) performance.
   b. [ The executives who hired xxx ] are satisfied with his performance (and raised his salary).

This indicates that the pronoun in each of (3a-b) as well as (4a-b) can be introduced into the syntactic representation in a 'referentially unsaturated' state, and its referential content comes to be identified only after it establishes an anaphoric relation with some other linguistic item in the course of the derivation. Without such a derivational process, the referential content of each pronoun in these sentences (with no specific speech context associated) would remain unspecified, and the interpretation of each sentence would remain incomplete.

One of the main goals of this paper is to reevaluate and reanalyze these simple observations on anaphoric interpretation, making an appeal to some of the leading ideas, though not necessarily their execution, of the minimalist program (Chomsky (1993)).
particular, I will propose and attempt to defend the following view of anaphoric interpretation:

(5)  
   a. All non-deictic proforms are introduced into syntactic representations underspecified with respect to their referential contents — the state we will refer to as "referentially unsaturated".
   b. Proforms are legitimate at LF only if they are "referentially saturated".
   c. The Principle of Full Interpretation as in (6) below requires all proforms to become legitimate objects at LF.
   d. Each underspecified proform may undergo covert computation to establish an anaphoric relation with some other linguistic item in order to become referentially saturated and satisfy the Principle of Full Interpretation (henceforth FI).²

(6) The Principle of Full Interpretation (FI):

   A representation at interface must consist entirely of legitimate objects.

If tenable, covert syntax applying in this fashion makes the syntax-semantics mapping for anaphoric interpretation highly transparent. Underlying this approach is the essence of the following view of language sketched out in Chomsky (1995). First, language acquisition basically amounts to the acquisition of the lexicon and parameter setting. Second, the function of each interface component is to sort out and organize the information introduced from the lexicon in such a way that each final interface representation contains the information both minimally and maximally necessary for the proper interpretation of sounds and meanings, respectively. Third, and quite importantly to us, whenever potential ungrammaticality arises in the interface components, grammar can make an appeal to the operation Affect. The application of this computational process, however, must be carried out in an optimal way. For instance, the Last Resort Principle as in (7) below prohibits its over-application.

(7) The Last Resort Principle: (Chomsky (1993))

   A step in derivation is legitimate only if it is necessary for convergence — had the step not been taken, the derivation would not have converged.
Affect, in other words, applies if and only if such a derivational step is necessary for the derivation to yield a well-formed interface representation — the state referred to as "convergence" of derivation in (7). Once it turns out that an appeal to Affect is inevitable, it must also apply in a most economical fashion, minimizing its computational search space, and perhaps avoiding redundancy like "undoing" the effect of the previously applied Affect. Thus, the general picture underlying the proposed approach is quite simple. Proforms in a referentially underspecified state will give rise to ungrammaticality at LF, but in so far as there is a derivation in which a covert operation applies and such proforms manage to have their referential contents specified, the sentence can be grammatical. If there is no such derivation, on the other hand, ungrammaticality results.

While we will pursue this "Affect" approach to anaphora, which is minimalist in spirit, we will take our departure from the current Chomskian "minimalist program" in several important respects. For instance, we will argue that driving force of covert syntax is not confined to the checking of formal features but may be an essentially semantic factor like referential saturation necessary for anaphoric interpretation, and that it may affect not only formal features but also semantic features. We will propose, in other words, to bring some aspects of anaphora back into syntax. We will also entertain an "interarboreal" covert operation across sentence and utterance boundaries for establishing some type of anaphoric relation (cf. Epstein, et al. (1998)). For this and other reasons to be mentioned below, our approach is "minimalist" only in the most general sense of this term, but is not faithful pursuit of Chomsky's proposals.

Section 2 below presents the proposed approach, and Section 3 provides its empirical and theoretical motivations. In particular, I will attempt to motivate, among others, the
following LF mappings achieved by the application of covert Affect. First, the "referentially unsaturated" locative proform there in (8a) below undergoes "Attract" as in (8b) for "referential saturation".

(8) Attract:
   a. LF: There arrived a spy [ in this city ].
   b. LF: [ in this city ] arrived a spy ↑

This analysis eventually leads us to the conclusion that it is not necessary to postulate a special class of entities called expletives.

The "referentially unsaturated" empty proform ([NP e]) in the discourse in Japanese in (9a) below, on the other hand, undergoes another option of Affect "Copy" as in (9b) to be "referentially saturated".

(9) Copy (and Identify):
      -TOP self’s PIN number-ACC forgot
       Demo Okusan-wa [NP e ] oboeita.
      but wife-TOP remembered
      'John forgot his own PIN number. But his wife remembered her own (or his) PIN number.'
   b. LF: John-wa [NP zibun-no ansyoobangoo-o ] wasuretesimatta
       [___________]
       Okusan-wa [NP zibun-no ansyoobangoo-o ] oboeteita
       [___________]

The resulting LF can yield sloppy identity of another proform zibun 'self' (contained in the copied NP) when it undergoes "Identify", still another option of Affect, as indicated by ↑ in each clause.
2. Proposals

Before we spell out our proposals, let us illustrate the notion "guises" with the example in (10) below. We will make a crucial appeal below to this notion in defining "referential saturation". The example is from Higginbotham (1985, 570).

(10) A: Who was that?
   B: Well, **he** put on **John**'s coat.

In (10B), when the speaker intends to imply that the person in question is John, the condition C (and/or D) of the binding theory apparently is not violated despite the fact that the pronoun **he** c-commands what appears to be its antecedent **John**. Heim (1982) and Heim (1992, pp. 14-17) ascribe the obviation of the condition C/D (henceforth BT (C/D)) observed here to the possibility of "multiple guises" associated with this discourse, as graphically represented in (11).

(11) Multiple Guises:

The core hypothesis here is that an utterance context supplies referring expressions not with referents but with "guises" which are in turn associated with actual referents. In (10B), for instance, the utterance context permits two distinct guises for two referring expressions—
guise₁ (= G₁ in (11)) as a visual impression of a male person who put on John's coat, and guise₂ (= G₂ in (11)) as an entry labelled as John in the memory of the speaker. The sentence "He put on John's coat" in (10B) therefore is interpreted roughly as "Whoever caused the visual impression in question put on the coat that belongs to the memory entry labelled as..."
John". He and John in (11) then can be interpreted as coreferential because the speaker implies that $G_1$ and $G_2$ happen to be the guises of a single referent $R_1$, which permits their association. Crucially, however, there is no direct anaphoric relation holding between he and John, and the absence of such a syntactic relation accounts for the absence of BT (C/D) violation. The apparent obviation of BT (C/D) here crucially relies on the availability of multiple guises and their association with a single referent. Since this state of affairs is made possible by the combination of utterance contexts and the speaker's implication, the sentence "He put on John's coat" uttered out of context fails to escape the problem. The notion guise thus proves to be a quite useful tool to capture otherwise puzzling obviation of BT (C/D).⁵

In this work, we will adopt the notion guise and hypothesize that guises and their association with linguistic expressions constitute part of formal syntactic representations. Let us first introduce the term $G$-link to refer to the association between a guise and a linguistic expression, and $R$-link to refer to that between a guise and a referent. We then hypothesize that $G$-link has the properties in (12a-b) and $R$-link has the property in (12c).

(12) a. Each R(eferential)-expression is introduced into a syntactic representation with a G-link.
   b. Each R-expression is linked to one and only one guise, and each guise is linked to one and only one R-expression in a given discourse.
   c. Each guise must be linked to one and only one referent but not vice versa in a discourse.

We assume here that G-link is already indicated on each referring expression when it is introduced to syntax.⁶ "R-expression" mentioned in (12a) refers to not only names but also deictic proforms. Those proforms introduced to syntax with a G-link, in other words, are deictic proforms while those introduced without a G-link are non-deictic proforms. In short, (12a-b) define an R-expression in terms of its biunique association with a guise.⁷ Unlike G-linking, however, we hypothesize that R-linking is established post-syntactically after LF.
The uniqueness condition on R-link asymmetrically imposed on guises as in (12c) permits more than one guise to be linked to a single referent as in (11) above while prohibiting more than one referent from being linked to a single guise.

With the notion guise, we now are ready to define "referential saturation" as in (13).

(13) A linguistic expression $\beta$ is referentially saturated at LF iff:

a. $\beta$ is G-linked,
b. $\beta$ is referentially dependent upon a G-linked item,

or
c. $\beta$ is a copy of a G-linked item.

Satisfaction of the condition stated in (13a) yields what we may characterize as "base-generated" referential saturation when referring expressions such as names and deictic proforms are introduced to syntax with a G-link, as illustrated in (14).

(14) "Base-generated" Referential Saturation:

Guises: $G_1$

Linguistic Expressions: The executives are satisfied with {his/John's} performance.

Satisfaction of the condition stated in (13b), on the other hand, permits what is base-generated as a referentially unsaturated item to become saturated by virtue of being dependent upon some other item that is G-linked, where the notion "referential dependency" is defined as in (15).

(15) $\beta$ is referentially dependent upon $\alpha$ iff:

(i) $\beta$ is referentially identified with $\alpha$, and
(ii) $\beta$ is c-commanded by $\alpha$.

"Referential identification" refers to the state that arises from the application of a covert computational process "Affect" indicated by the line connecting John and his in (16).
Referential Saturation by Referential Dependency:

(16) \[ \text{Guises: } G_1 \]
\[ \text{Linguistic Expressions: } \text{John is satisfied with his performance.} \]
\[ +cc \]

(16) illustrates referential saturation of a non-deictic pronoun via its referential dependency upon a name, which satisfies the condition in (13b).

As part of the definition of referential dependency, we adopt Reinhart's (1983, 18) definition of c-command as in (17a).

(17) **C-command:**

a. Node \( \alpha \) c-commands \( \beta \) iff the branching node immediately dominating \( \alpha \) also dominates \( \beta \), where \( \alpha \neq \beta \).

b. 

\[
\begin{array}{c}
\alpha \\
\ldots \beta_2 \ldots \\
\end{array} 
\begin{array}{c}
\ldots \beta_1 \ldots \\
\end{array}
\]

Note that the definition (17a) permits \( \alpha \) to c-command \( \beta \) even when the former dominates the latter. Therefore, in (17b) for example, \( \alpha \) c-commands not only \( \beta_1 \) but also \( \beta_2 \). The condition "where \( \alpha \neq \beta \)" in (17a), on the other hand, guarantees that \( \alpha \) does not c-command \( \alpha \) itself.

We will see the important consequences of the definition in (17a) in 3.3 below.

A case that remains to be accounted for is (3b), the relevant part of which is repeated here as (18).

(18) \[ \text{[ The executives who hired John ] are satisfied with his performance.} \]

The grammaticality of this sentence suggests that the pronoun \textit{his} is "saturated". (From now on, we will customarily use the terms "identified/identification", "dependent/dependency" and "saturated/saturation" to refer to "referentially identified/referential identification",}
"referentially dependent/referential dependency" and "referentially saturated/referential saturation", respectively, for ease of reference.) This saturation, however, cannot be attained in terms of either of (13a-b). As illustrated in (19) below, *his* is neither G-linked nor is dependent upon a G-linked item *John*, lacking a required c-command relation.

(19) Saturation without Dependency:

At this point, I would like to propose and spell out the entire picture of what I will refer to as the Affect Approach to anaphoric interpretation, which entertains a hypothesis as in (20) below and postulates (21a-c) as options of Affect.

(20) Unsaturated anaphoric items undergo a covert computational process Affect in order to become referentially saturated and satisfy FI at LF.

(21) Options of Affect for Anaphoric Interpretation:

a. Identify
b. Copy
c. Attract

In (18), the pronoun *his* introduced from the lexicon without any G-link is unsaturated. It therefore would fail to satisfy FI if it remains to be unsaturated. The first and necessary step toward convergence it takes is to undergo "Identify" (21a), which anaphorically associates two and only two linguistic expressions that agree in relevant grammatical features. "Identify" as a computational process is essentially identical to Higginbotham's (1983) "Linking" except that it applies in covert rather than overt syntax and that it crucially lacks directionality. It is also similar in its function to Heim's (1992) "(co)linking" defined in terms of "outer" and "inner" indices. Just as Higginbotham's original Linking, "Identify" as a
syntactic operation is intended to replace "coindexation" often used in the generative 
literature. We do, in other words, eliminate referential indices from syntactic representations. 
As will be discussed in Section 3.2 below, the referential identification established by the 
application of Identify does not necessarily coincide with that established by coindexation. 
Let me also emphasize here that "referential identification" refers strictly to a syntactic state 
rather than a semantic state. Both in (22) and (23) below, for example, Identify has applied 
and identification has been established between the two highlighted items. Such syntactic 
states, however, do not necessarily lead to a semantic state in which the two items 
"identified" at LF actually refer to a single individual.

(22) LF$_i$: **John** saw **him**.
   \[\text{[_______]}\]

(23) LF$_i$: **He** saw **John**.
   \[\text{[_______]}\]

Since Identify as an option of Affect is triggered solely for referential saturation, the Last 
Resort Principle prohibits G-linked items like names from invoking this computational 
process. Therefore, if two names as in the example (24) below ever come to be interpreted as 
coreferential, that must be made possible due to multiple guises rather than the application of 
Identify.

(24) **John** thinks that **John** is the best student in this school.

This is the interpretation of BT (C) under the Affect Approach.

    When an anaphoric item undergoes Identify and is identified with a c-commanding G-
linked item, dependency is established and saturation is achieved, as exemplified by (16) 
above. In the case of (19), on the contrary, saturation of **his** is yet to be attained due to the 
lack of dependency, although identity has been established between **his** and **John**. It is this
context in which another computational process "Copy" ((21b)) applies for saturation. Copy
operates on the two co-identified items and duplicates the semantic content of the more
contentful item and superimposes it onto the less contentful item in a non-overlapping
fashion.\textsuperscript{10} Thus, in (19), whatever semantic contents of John that his does not contain are
duplicated from the former and superimposed onto the latter. In effect, his "turns into" the
copy of John, as illustrated in (25).

(25) Copy for Saturation:
\begin{verbatim}
Guises: 

Linguistic Expressions: [ The executives who hired John ] are satisfied with John's performance.
\end{verbatim}

It is this LF-representation in which the non-deictic pronoun his comes to be saturated as the
copy of a G-linked item John, satisfying the definition (13c). One crucial assumption
underlying this mode of saturation is that Copy allows the "duplicate" to inherit the
legitimacy (or referentially saturated status) from its "original". The "duplicate" therefore
behaves on a par with the "original".\textsuperscript{11} In a sense, our analysis with covert Copy here
achieves at LF the reverse effect of the Pronominalization transformation in the Standard
Theory.

In Section 3.1 below, we will extend the Affect Approach to the analysis of expletives
and argue that they undergo covert "Attract" ((21c)). Attract, unlike Copy, transfers (rather
than duplicates) the entire content, including structure and G-link, of a more contentful item
onto the less contentful co-identified item, as illustrated by the derivation in (26).

(26) a. Identify:

\begin{verbatim}
LF: It is unlikely \[ \text{that he will win} \].
\end{verbatim}
b. Attract:

\[
\text{LF}_j: \quad [_{\text{CP1}} \text{that he will win}] \text{ is unlikely } \text{t.}
\]

\[
\begin{array}{c|c}
\text{-t} & \text{+t} \\
\hline
\text{θ} & \text{θ}
\end{array}
\]

In (26a), the pronoun \( \text{it} \) and a complement CP are identified but \( \text{it} \) is not dependent upon the CP. It is the unsaturated status of \( \text{it} \) in this context that triggers the attraction of the entire CP by \( \text{it} \), as illustrated in (26b). This computational process achieves the saturation of the pronoun \( \text{it} \) by turning it into a G-linked item, that is, with the satisfaction of the definition (13a).\(^{12}\) We will discuss the issue of agreement involved in this construction in 3.1.3 below. We will also discuss the case of BT (D) as in (27) below, and elucidate why (26) does not pattern with (27) in 3.3.

(27) \( \text{He} \) is satisfied with \( \text{John} \)'s performance.

\[
\begin{array}{c|c}
\text{θ} & \text{+θ}
\end{array}
\]

Although it may appear at first sight that the choice between the application of Attract and that of Copy is arbitrarily determined, it actually is not. First, if Copy instead of Attract is applied to (26a) as illustrated in (28) below, it would yield an LF-representation in which the duplicated argument CP is placed in a non-thematic position, resulting in the violation of FI.

(28) \( \text{LF}_j: \quad \text{*[}_{\text{CP1}} \text{that he will win}] \text{ is unlikely } [_{\text{CP1}} \text{that he will win}] \)

\[
\begin{array}{c|c}
\text{-t} & \text{+t} \\
\hline
\text{θ} & \text{θ}
\end{array}
\]

If, on the other hand, Attract instead of Copy applied in (19) as illustrated in (29) below, the resulting chain would not only involve two thematic positions but also leave a trace that is not properly bound.\(^{13}\)
The executives who hired $t$ are satisfied with John's performance.

Thus, the choice of the correct option of Affect at the right occasion follows from the general aspects of grammar. Note also that each of the proposed options of covert "Affect" for referential saturation establishes what is independently verified as a basic syntactic relation — Attract establishes a chain, Identify together with dependency results in binding, and Copy mimics the LF-reconstruction required for ellipsis. The proposed options of Affect, in other words, are nothing peculiar but are general in their nature.

To sum up the Affect Approach to anaphoric interpretation presented above, we hypothesized that linguistic expressions can be base-generated either G-linked or not G-linked. If the former, they are straightforwardly "referentially saturated" and satisfy FI at LF. In the latter case, the non-G-linked linguistic expressions may undergo covert Affect and become saturated by being directly or indirectly associated with G-links. First, they may undergo Identify and become dependent upon a c-commanding G-linked item. Second, if dependency cannot be achieved even after they undergo Identify, they may further undergo Copy and become saturated by turning into the copy of a G-linked item. Finally, if a non-G-linked item is located in a non-thematic position and c-commands a co-identified G-linked item, it may undergo Attract, and become saturated by virtue of turning into a directly G-linked item. Roughly speaking, we have defined an aspect of FI-satisfaction in terms of the notion "referential saturation," which in turn is syntactically defined in terms of G-link. We have also identified the application of one or more options of Affect for anaphoric interpretation as the quest for a direct or indirect G-link.
3. Motivation

I now start presenting motivation for the Affect Approach to anaphoric interpretation. In each following subsection, I will attempt to provide empirical and/or theoretical advantages this approach gains when each of Identify, Copy and Attract is adopted as a possible option of covert Affect for saturation. Among the three options, I will first attempt to motivate Attract, since, if successful, that will clarify and justify our theoretical stance, which departs from the current "Chomskian" view of computational processes as in (30) (Chomsky (1995, pp. 262-265)).

(30) a. Affect, overt or covert, applies only for the checking of formal features.
   b. Covert Affect affects only formal features.

More specifically, I would like to defend the view that covert Affect be characterized as in (31).

(31) a. Covert Affect can be triggered by semantic factors.16
   b. Not only formal features but also semantic features can be operated upon by covert Affect.

These properties of covert Affect certainly must be true if each of its options applies for the fulfillment of saturation in the way we argue for below. More generally, I would like to uphold the hypothesis that computational processes at each interface can be enacted by anything relevant to the interpretation at that interface.

3.1 Attract

3.1.1 The Background

Let us start our discussion of Attract for saturation with a brief (and certainly non-comprehensive) historical survey of the analysis of expletives as exemplified by (32a-b).
There are two observations that have determined the direction of the syntactic analysis of expletives in the generative grammar, especially that of *there*. First, at least in English, expletives in typical cases are subjects, as demonstrated by the applicability of Subject-Aux Inversion as in (33).\(^\text{17}\)

\[33\] a. **Did there** arise a storm in the South?  
   b. **Is it** likely that he will win?

Second, as a subject, at least *there* seems to agree in number with a verb (via a functional category), providing us with the impression that it inherits formal features from the "associate DP" located post-verbally:

\[34\] a. *There\(_1\)* seems \[ t\(_1\) to be a book \] on the table \].  
       \[..................................................\]

   b. *There\(_1\)* seem-Ø \[ t\(_1\) to be some books \] on the table \].  
       \[..................................................\]

Reflecting these observations, Chomsky (1986, pp. 131 ff., 143, 179) proposed an early minimalist analysis in which *there* is replaced by its associate DP at LF, as illustrated by the derivation in (35).

\[35\] a. S-structure: **There** arises a storm in the South.  
       b. LF: **A storm** arises in the South.

More recently, however, Chomsky (1995, pp. 265-276) decided to draw back from the idea that the associate DP actually replaces *there* at LF. Instead, he proposed a more complicated minimalist analysis summarized in (36).

\[36\] a. Base-generated *there* can check categorial feature D(eterminer) of T(ense) and satisfy the Extended Projection Principle (EPP).  
   b. It, however, cannot check Case- and Φ-features of T.  
   c. The formal features of the associate DP (henceforth FF (associate)) therefore are LF-adjoined to the (complex) T to fulfill the checking of Case- and Φ-features.
d. Agreement thus takes place directly between the adjoined FF (associate) and T without the
mediation by there.

Underlying this analysis is the hypothesis in (37).

(37) Overt movement moves the entire expression, while covert movement moves only formal
features.

One observation that motivates the hypothesis in (37) is that a quantified associate DP as in
(38) below cannot take scope over negation, as noted by Williams (1984).

(38) There aren't [ many pictures ] on the wall.

Thus, this sentence can be interpreted as in (39a) but not as in (39b).

(39) a. It is not the case that many pictures are on the wall. (NEG > MANY)
    b. Many pictures are such that they are not on the wall. (*MANY > NEG)

The lack of the scope interpretation as in (39b), however, can be accounted for if only the
formal features of the associate raise, leaving its semantic features behind within the VP as in
(40) below. Semantic features of the associate many pictures are represented here as SF
(many pictures).18

(40) LF: There [\( T \) FF (many pictures)-[\( T \) are,\( t \)] ] not [\( \text{VP} \) t, SF (many pictures) on the wall ]

Lasnik (1995) also argues that the contrast between (41a) and (41b) below supports the
hypothesis in (37).

(41) a. They seemed to each other [ t to have been angry ]
    \[
    \uparrow \text{________________________}_1
    \]

    b. There seem (to *each other) [ t to have been many linguists given good job offers ]
    \[
    \uparrow \text{________________________}_1 \uparrow \text{Covert} \] 1

When the antecedent of an anaphor is overtly raised and placed in a position from which it
can c-command the anaphor as in (41a), the condition A of the Binding Theory (henceforth
BT (A)) can be satisfied. When similar overt movement involves there as in (41b), on the
other hand, BT (A) cannot be satisfied, which suggests that the binding capacity of the
associate DP is not inherited by there (via a chain) when the trace of there is associated with
many linguists at LF within the subordinate clause. Again, if the movement of FF (associate)
(henceforth Move FF (associate)) leaves behind the semantic content of the associate DP
including its binding capacity, the contrast in (41) follows.

Chomsky (1995, pp. 272-276) considers that the contrast in (41) justifies his
assumption (36c) rather than (37), insisting that adjunction of FF (associate) to there would
deprive us of a means to distinguish the LF representations for (41a) and (41b). However,
since overtly raised there exhibits agreement based upon the number available in the
subordinate clause as exemplified by (34a-b) above, it must be assumed at any rate in this
approach that FF (associate) is inherited by there. Thus, the contrast in (41) does not
motivate the assumption in (36c), though it appears to motivate the assumption in (37).19

As Lasnik (1995) further points out, however, (37) leads us to some contradiction
when it is combined with Chomsky's (1993) covert raising of an accusative object to Spec-
Agr for Case checking (henceforth ACC-raising). It has been argued in the literature that
ACC-raising is motivated by the account it can provide for the facts as in (42).

(42) a. The DA proved [IP two men to have been at the scene ] during each other's trials.
   (Lasnik and Saito (1991))
   b. John [VP expected [DP no one that I did [VP e ] ] to be electable ].
   (Hornstein (1994))

In (42a), the exceptionally Case-marked (henceforth ECM) subordinate subject two men can
unexpectedly bind the anaphor each other located in the matrix clause (Postal (1974)). In (42b),
the antecedent contained deletion of VP located within the ECM subordinate subject can be
properly interpreted, again despite the expected difficulty, apparently without causing any
infinite regress.20 It is argued that these puzzling facts can be accounted for if the ECM
subordinate subject is assumed to undergo covert ACC-raising and maps (42a-b) onto (43a-b) at LF.

(43)  

a. LF: The DA \([_{\text{Agro}} \text{two men}], \text{proved}_{_{\text{Agr}}} [_{\text{VP}} t_2 [_{\text{IP}} t_1 \text{to have been at the scene } ] ] \) during each other's trials ]  

b. LF: John \([_{\text{Agro}} \text{no one I did } [_{\text{VP}} e ] ], \text{expected}_{_{\text{Agr}}} [_{\text{VP}} t_2 [_{\text{IP}} t_1 \text{to have been } ] ] \) at the scene]  

Under this analysis, the ECM subject is located in the matrix clause in (43a) and can properly bind each other (Lasnik and Saito (1991)). Likewise, the elided VP is no longer located within its antecedent VP in (43b), and infinite regress is resolved (Hornstein (1994)). Note, however, in order for these arguments to be valid, ACC-raising, despite its covert status, must be regarded as moving the entire content of accusative DPs including their semantic properties and even their structures rather than just their FFs. This, however, contradicts with the assumption (37).

Even more puzzling is that, in sharp contrast to (42a-b), similar ECM constructions involving there as the subordinate subject as in (44a-b) below exhibit a BT (A) violation and infinite regress, respectively.

(44)  

a. *The DA proved \([_{\text{IP}} \text{there to have been two men at the scene } ] \) during each other's trials.  

b. *John \([_{\text{VP}} \text{expected } [_{\text{IP}} \text{there to be } ] ] \) electable ].

Since ACC-raising should lift there in (44a-b) out of the subordinate clause just as it lifted the subordinate subjects in (43a-b), we must assume here that the covert operation which moves the associate DP to there fails to move the DP's semantic and structural content when it moves its FFs. Covert movement, in other words, must be assumed to affect the entire
content and structure of the moved item in one case, while it must be assumed to affect only FFs in another case.

In order to escape this serious theory-internal contradiction, Lasnik (1995) adopts Koizumi's (1993) proposal that ACC-raising to Spec-Agr_o applies overtly, being accompanied by overt verb raising (which allows him to maintain the correct surface word order), as illustrated in (45).

(45) \[
\text{The DA proved}_2 T-\text{ Agr}_x \left[ \text{ Agr}_o \text{ two men}_1 \left[ \text{ VP t}_2 \left[ \text{ t}_1 \text{ to be at the scene } \right] \text{ during } \right] \text{ each other's trials } \right].
\]

With this assumption, the dilemma is solved and (37) can be maintained. First, ACC-raising now regarded as overt movement moves the entire content of the ECM subordinate subject, which explains the well-formed interpretations in (42a-b). The associate-DP movement in the there construction, on the other hand, is a covert movement, and therefore moves only the associates' FFs, which explains the persistence of the interpretive problems in (44a-b).

To sum up Lasnik's arguments, Chomsky's hypothesis (37) is maintained with the assumption, first, that the there construction involves a covert computation Move FF (DP-associate), and that ACC-raising is an overt computational process which affects the whole item.

3.1.2 Attract Locative

I now would like to propose an analysis of expletive there as summarized in (46), which I will call the "Attract Locative" Analysis: 21
(46) "Attract Locative" Analysis:

a. The so-called expletive *there* is a non-deictic locative anaphoric item (Freeze (1992)), which attracts its locative associate at LF for its referential saturation.
b. The entire content of the associate locative is attracted to *there*.

The LF-derivation of the "Attract Locative" Analysis is exemplified in (47).

(47) a. **Identify:**

\[
\text{LF}_i: \text{There arrived a spy [ in this city ]}. \\
\text{[____________]} \\
\]

b. **Attract:**

\[
\text{LF}_j: \text{[ in this city ] arrived a spy } \text{there}. \\
\text{[____________]} \\
\]

One obvious and important departure of this approach from the standard generative view of the expletive *there* is that its associate is regarded no longer as a DP but as a locative expression.\(^{22}\) The important theoretical backbone of the proposed analysis is the Affect Approach to anaphoric interpretation. The idea is simple — *there* as an unsaturated proform must undergo covert Affect at LF so that it can be saturated and satisfy FI. As a locative proform, *there* is identified with a locative expression as in (47a). Since *there* c-commands its antecedent while being located in a non-thematic position, Attract (rather than Copy) applies as in (47b), to ensure that the resulting chain would involve only one locative role and one locative expression. In what follows, we will provide arguments to show, first, that a locative rather than a DP should be regarded as the associate of *there*, and second, that the entire content of a locative must be attracted to *there*. In some of those arguments, we will make a crucial appeal to the observations and logic Lasnik (1995) offered in his defense of Chomsky's hypothesis (37).
The first piece of motivation for the Attract Locative Analysis arises when we examine the locality of anaphor dependency. To begin with, anaphor binding is known to exhibit certain locality restriction as illustrated in (48).

(48) *They think [ that **the stolen picture** is hanging [ in each other's rooms ]]  

Here, the intervening subject **the stolen picture** prohibits each other to be bound by they. Strangely, on the other hand, an anaphor can be non-locally bound when the intervening subject is there. In (49) below, for example, each other can be non-locally bound by they.

(49) They think [ that **there** is a stolen picture (hanging) [ in each other's rooms ]]  

One crucial observation here is that each other is located in the locative expression appearing in the there construction. Anaphor binding suddenly ceases to exhibit the locality restriction, in other words, when the anaphor shows up in the locative associate of there. This otherwise puzzling fact follows quite naturally under the Attract Locative Analysis of there. After the locative associate of there is attracted at LF, the sentence in (49) is represented as in (50).

(50) After Attract:  

After Attract:  

\[
\text{LF: } \quad \text{They think [ that [ in each other's rooms ] is a stolen picture (hanging) ]} \quad \text{t}
\]

In this representation, the locality constraint on anaphor binding is in fact satisfied, since each other is locally bound by its antecedent they. This supports the analysis in which a locative rather than a DP is regarded as the associate of there, and the entire content of a locative (including the contained anaphor) is attracted to there. (From now on, we let the label "Attract Locative Analysis" entail these views.)
There is one thing that we must clarify in relation to the above discussion. In a sentence like (51) below, anaphor binding is successful even when an anaphor is contained in the DP associate of there, which strikes us as a piece of data neutralizing our argument for the proposed Attract Locatives Analysis.

(51) They think that there are [stolen pictures of each other] (hanging) in the next room.

It is well-known, however, that anaphors in the complement position of NPs can often be long-distance bound in English, in violation of BT (A). Such an exceptional behavior of anaphors, often referred to as "logophoric", is exemplified in (52a-b) below. The examples have been cited from Reinhart and Reuland (1993, 670, 682)).

(52) a. It angered him that she … tried to attract [a man like himself].
   (Quoted from Zribi-Hertz (1989))
   b. [The pictures of himself] that John saw in the post office was ugly.

How exactly this exceptional anaphor binding should be treated is controversial. Some researchers postulate PRO subjects within NPs along the line of Chomsky (1986), and others theorize lexical saturation of the thematic role of the subject of NPs along the line of Williams (1985). Whether one adopts one of these or some other approaches, in other words, some non-core treatment seems to be required in dealing with anaphors in the complement position within NPs. We should therefore refrain from employing the construction as in (51) and (52) in examining the locality effect of anaphor binding. On the contrary, the locality effect on anaphor binding with genitive-marked anaphors as in (53) below has been confirmed, and it can provide a reliable testing ground for the locality of anaphor binding.

(53) We believe that they thought that each other's pictures were on sale.

Thus, long-distance anaphor binding as in (51) does not necessarily impair our argument. Given the locality effect observed in (53), on the other hand, the contrast between (48) and
(49), especially the absence of the locality restriction in (49) demands explanation, and the Attract Locative Analysis fulfills this task.

The second argument for the Attract Locative Analysis arises when we examine the contrast between (54) and (55a-c) below concerning the possibility of interpreting a pronoun as a bound variable.²⁴

(54) **There** seem to [ **its** residents ] [ _vp t _ to be many serious problems **in every major city** ]

²⁴

(55) a. **There** seems to [ **his** friends ] [ _t _ to be **someone** quite suitable for that position ].

b. **There** seem to [ **their** advisors ] [ _t _ to be **many new graduates** quite suitable for that position ].

c. **There** seems to [ **its** residents ] [ _t _ to be [ a book about **every city** ] in the Library of Congress ].

In (54), the pronoun **its** can be interpreted as a variable bound by the quantified DP **every city**, yielding a distributive reading. So the sentence may indicate, for example, that, in each of New York, Tokyo, London, and so on, the residents there have the impression that many serious problems exist in the city they live in. In (55a-c), on the other hand, the pronouns **his**, **their** and **its** cannot be interpreted as a variable bound by the quantified DP **someone**, **many new graduates** and **every city**, respectively. The sentence in (55a) therefore cannot mean that there is at least one person such that this person's friends have the impression that he or she is quite suitable for that position, and (55b) cannot mean that many new graduates are such that their own advisors have the impression that they (= those graduates) are quite suitable for that position. (55c) likewise cannot mean that every city is such that its residents have the impression that a book about their own city exists in the Library of Congress. (55a-c) also contrast with (56) below, which may provide these interpretations without any problem.²⁵

(56) a. **Someone** seems to **his** friends [ _t _ to be quite suitable for that position ].

b. **Many new graduates** seem to [ **their** advisors ] [ _t _ to be quite suitable for that position].

²⁵
The contrast here indicates that the quantificational element c-commands the pronoun at LF in (54) as well as in (56), but not in (55a-b). The crucial difference, especially that between (54) and (55a-c), is that the quantificational element is (part of) a locative expression in (54) but it is a post-verbal DP in (55a-c). This contrast then can be captured if we can let there inherit the quantificational property of the locative (via the intermediate trace) but not that of the post-verbal DP, that is, if we adopt the Attract Locative Analysis characterized above.

There are, however, a few loose ends in the above argument that we must tighten in order to ensure its validity. First, as pointed out to me by Steven Franks, the quantificational locative expression might c-command the pronoun its in (54) from the outset, having been introduced as an IP-adjunct in the matrix clause. In fact, as has been pointed out by one of the anonymous readers, (54) can be also interpreted on a par with (57).

(57) **In every major city**, there seem to its residents to be many serious problems.

Since we can observe co-occurrence of two locative expressions as in (58) below, we should perhaps allow a locative as an IP-adjunct in addition to a locative associate of the expletive there.

(58) **In European countries**, there seem to its residents to be serious problems in **every major city**.

This means that the analyses as in (59a) below, and hence (59b) as well, should be also possible under the Attract Locative Approach.

(59) a. [ip **In every major city**, there seem to its residents [ip to be many serious problems [Loc e ]]]

   b. [ip There seem to its residents [ip to be many serious problems [Loc e ]] in **every major city** ]

We can, however, ensure the subordinate status of the locative associate of there in a sentence like (60).
There seem to its residents [IP to be [NP many serious problems t₁] in every city [CP₁ that have no obvious solution ]].

In this sentence, the application of Extraposition from NP placed the relative clause CP₁ to the right of the locative. This movement, however, should not go beyond the IP-boundary due to the Right Roof Constraint (Ross (1967)). It therefore is ensured that the locative in (60) is located within the subordinate clause in overt syntax. Now, the sentence is somewhat difficult to compute due to its extreme length. All of about a dozen speakers I checked with agreed, however, that the bound variable interpretation of the pronoun is still possible.

Compare, furthermore, (61) below with (60).

In European countries, [₁ many serious problems ] seem to its residents [ t₁ to exist in every major city (that have no obvious solution) ].

It seems to some speakers to be the case that the pronoun its in this sentence is interpretable as the variable bound by European countries despite the absence of agreement in number. The same pronoun, however, is not interpretable as the variable bound by the quantificational locative expression in every major city. It is, in other words, not the case that every quantificational locative expression in the subordinate clause comes to bind a pronoun in a higher clause. Rather, such a phenomenon takes place only when a quantificational locative expression shows up as the locative associate of there, which supports the proposed Attract Locative Analysis.

In order to make our argument involving (54) even tighter, however, we must also exercise caution and eliminate the focus interpretation of the quantified locative, which seems to permit the bound variable interpretation in question even in sentences like (55c) and (61) (repeated in (62a-b) below) at least for some speakers. Compare the interpretation by clearly
focusing a quantifier and then intentionally defocusing it. The speakers I have checked with
all agree that the bound variable interpretation in question is never available at least with the
defocused quantifiers.

(62) a. **There** seems to its$_1$ residents to be [ a book about { ÉVERY city$_1$ / every city$_{1}$ }] 
in the Library of Congress.

b. In European countries, many serious problems seem to its$_1$ residents to exist
   [ { in ÉVERY major city$_1$ / in every major city$_{1}$ }] (that have no obvious 
solution).

Finally, the analysis of (54) presupposes that the quantified DP can bind out of the
locative PP, and such transparent nature of PPs as well must be independently verified.

While how it becomes possible is not clear, this mode of binding, in fact, is known to have to
be postulated to account for various facts. For example, BT (C/D) is violated in (63a)
below, involving binding out of a PP:

(63) a. *It seems [pp$_1$ to him$_1$ ] that John$_1$ is a failure.                (Chomsky (1986, p. 183))
   b. It seems [pp$_1$ to [dp$_1$ his$_1$ friends ]] that John$_1$ is a failure.

Note that the problem in (63a) disappears in (63b) when the pronoun is further embedded into
a DP. This demonstrates that the problem in (63a) was that of binding, and binding out of a
PP is indeed permitted in this sentence. Likewise, anaphor binding out of a PP is possible in
(64). [27]

(64) Aeron talked [pp$_1$ with Rachel$_1$ ] about herself$_1$.

We thus regard the availability of a bound variable interpretation of a pronoun in (54) as a
legitimate motivation for the Attract Locative Analysis.

The third argument for the Attract Locative Analysis arises when we examine the
scope interaction of a quantified element and negation in the **there** construction. Recall first,
that when the post-verbal DP in this construction is quantificational as in (65) below, it cannot take scope over negation.

(65) There are₁-n't [vp ti many pictures on the wall ] (*MANY > NEG)

When a quantified element is located within the locative associate of there within the VP as in (66) below, on the contrary, it can take scope over negation.

(66) In Nagano, there are₁-n't [vp ti linguistics students [ in { many / all } private colleges]]
     (ok{MANY/∀} > NEG)

Thus, the sentence in (66) can mean that many (or all) private colleges in Nagano are such that there are no linguistics students there. Similarly, when there is overtly raised from the subordinate clause to the matrix as in (67) below, the quantified DP within the locative expression in the subordinate clause can take scope over the negation in the matrix.

(67) [ A statement made by a spokesman of TOYS "R" US: ]

    There do not seem [ t to remain Power Ranger dolls [ in { many / all } of our stores ]].
    ↑________________|  (ok{MANY/∀} > NEG)

The sentence, thus, can mean "Many (or all) of our stores are such that it does not seem to be the case that Power Ranger dolls remain there". Once again, we can ensure the subordinate location of the locative in overt syntax by applying Extraposition from NP as in (68) below, and still maintain the same scope relation as in (67).

(68) There do not seem [ip to remain [np Power Ranger dolls ti ] in { many / all } of our stores
     [cpi that we can sell to Christmas shoppers ]].
     (ok{MANY/∀} > NEG)

That a quantified element can establish its domain out of a PP can be also demonstrated in (69) independently of the there construction.

(69) [pp In every college ], the tenure system does not seem to torture the professors who do great research.  (ok∀ > NEG)
In this sentence, the locative PP is located in the top-most position of the matrix clause, and the quantified element in the locative PP can take scope over the negation in the matrix clause. This demonstrates that a quantified element can establish its domain out of a PP.

When the post-verbal DP is quantificational as in (70) below, on the other hand, the overt raising of there does not permit the quantified DP to take scope over the negation in the matrix.

(70) **There do not seem [ t to remain { many / all } Power Ranger dolls [ in our stores ] ].**

(*{MANY/∀} > NEG)

The sentence in (70) therefore cannot mean "Many (or all) Power Ranger dolls are such that it does not seem to be the case that they remain in our stores". Here, again, we should carefully avoid the focusing of the quantifiers as in (71), which may permit the wide scope of these quantifiers for some speakers.

(71) There do not seem to remain { MÁNY / ÁLL } Power Ranger dolls in our stores.

Observe also that when a quantificational element shows up as part of temporal or manner expressions as in (72) and (73) below, similar contrasts obtain between focused and unfocused cases.

(72) Temporal Expression:

a. There did not seem to arise a problem { ÉVERY time / every time } [ that required serious care ].
b. There did not seem to arise a problem { MÁNY times / many times } [ that required serious care ].

(73) Manner Expression:

a. There does not seem to arise divergence in { ÉVERY way / every way } [ that requires us to revise our theory ].
b. There does not seem to arise divergence in { MÁNY ways / many ways } [ that requires us to revise our theory ].
In (72a), for example, the interpretation 'A problem that required serious care never seemed to arise' is possible only when every is focused. Without such focusing, however, the only available interpretation seems to be something like 'It does not seem to be the case that a problem that required serious care arose every time.'

Finally, similar contrasts involving focusing can be observed when a quantificational element shows up in a locative expression that either is not a DP associate of there as in (74), or is a DP associate but there is located in a position lower than negation at surface, as in (75).

(74) a. On the East Coast, Power Ranger dolls do not seem to remain in \{ ALL / all \} of our stores (that we can sell to Christmas shoppers).
   b. On the East Coast, Power Ranger dolls do not seem to remain in \{ MANY / many \} of our stores (that we can sell to Christmas shoppers).
   c. The tenure system does not seem to torture the professors in \{ EVERY / every \} college who do great research.

(75) a. It does not seem that there remain Power Ranger dolls in \{ ALL / all \} of our stores that we can sell to Christmas shoppers.
   b. It does not seem that there remain Power Ranger dolls in \{ MANY / many \} of our stores that we can sell to Christmas shoppers.

For the sake of completeness and fairness of the argument, the reader is urged to reexamine (68) and confirm the possibility of the wide scope of the quantified elements even when they are intentionally defocused.

Hatakeyama (1999, 8) independently argues for the Attract Clause Analysis and offers another type of scope argument, citing Kuno's (1971, 365) observation concerning the example (76).

(76) There are [DP many girls ] [PP in every class ].

Here, the quantified locative expression can take scope higher than that of the post-verbal DP. Since both quantified elements appear within a single sentence in (76), however, this scope
phenomenon does not necessarily support the Attract Locative Analysis for those who find a scope ambiguity in a sentence like (77).

(77) $[\text{DP } \text{Many students }]$ attended $[\text{DP } \text{every lecture }].$

We can, however, provide a more reliable scope argument involving modality. First, we construct two distinct discourses in which a structurally identical $\text{there}$-construction is involved:

(78) a. In state universities, there $\textbf{must}$ be a female professor in $\text{every}$ department.
    Otherwise, the university could get sued.
    — MUST > $\forall$

    b. In this university, there $\textbf{must}$ be a female professor in $\text{every}$ department.
    Otherwise, the department will disqualify from receiving a fund from the state.
    — $\forall$ > MUST

The felicitousness of both discourses here suggests that the quantified locative expression can take scope either lower or higher than that of the modal $\text{must}$. Then, we construct similar discourses, this time involving the raising of $\text{there}$:

(79) a. In state universities, $\textbf{there}$ seems $[\text{IP } t_1 \text{ to have } [\text{IP } t_1 \text{ to be } [\text{a female professor } t_2 ] \text{ in } \text{every } \text{department } [\text{CP}_2 \text{ that teaches full-time } ] ]]$. Otherwise, the university could get sued.
    — MUST > $\forall$

    b. In this university, $\textbf{there}$ seems $t_1 \text{ to have } [\text{IP } t_1 \text{ to be } [\text{a female professor } t_2 ] \text{ in } \text{every } \text{department } [\text{CP}_2 \text{ that teaches full-time } ] ]$. Otherwise, the department could disqualify from receiving a fund from the state.
    — $\forall$ > MUST

Here as well, both discourses are felicitous, which suggests the possibility of a scope ambiguity similar to that observed in (78). Note especially that (79b) permits the higher scope of the locative expression, which can be ascribed to the presence of $\text{there}$ in the matrix clause under the Attract Locative Analysis but remains mysterious otherwise. The extraposition of $\text{CP}_2$ out of the post-verbal DP ensures the subordinate status of the locative
expression in (79b). The readers are again urged to confirm that the discourse in (79b) is felicitous even when we exercise by now familiar caution to avoid focusing of the quantified element in the locative expression.

In short, all of the scope facts presented above follow straightforwardly if a locative expression rather than a DP is the referential associate of there, and if there inherits the quantificational property of the locative expression with the application of covert Attract (via a chain in cases like (67) and (68)). That is, if we adopt the Attract Locative Analysis.30

The fourth argument arises when we examine antecedent contained deletion (henceforth ACD) involving a there construction. The examination, however, requires some careful untangling of the data since, as pointed out by Lasnik (1999, Chapters 3.6), more than one computational processes seems to be responsible for ACD resolution. Let us here briefly overview Lasnik's approach, which we will use as a stepping stone.

First, as in his earlier work, Lasnik extends Hornstein's (1994) proposal and regards ACC-raising as one major operation which resolves ACD. He presents the contrast in (80) below to verify the relevance of the Accusative Case-checking in ACD resolution.

(80) a. 7Dulles talked about [DP Philby, who Angleton did not [VP e]].  
   b. *Mary stood near [DP Susan, who Emily did not [VP e]].

It is pointed out that ACD can be resolved in (80a) but not in (80b) because reanalysis of the verb-preposition sequence is permitted only in the former, which is evidenced by pseudopassivization in (81).

(81) a. Philby1 was [t, talked about] t1 (by Dulles).  
   b. *Susan1 was stood near t1 (by Mary).

ACC-raising of the object DP, in other words, is possible in (80a) but not in (80b).
Lasnik also presents the contrast in (82) below and points out that ACD can be resolved even without ACC-raising.

(82) a. Mary stood near [dp everyone] Emily did [vp e].
   b. *Mary stood near [dp Susan, who Emily did not] [vp e]. (= (80b))

Extending Baltin's (1987) argument, he claims that the resolution in (82a) is induced when extraposition applies to a restrictive relative clause at LF, as illustrated in (83).

(83) LF: Mary [vp stood near everyone] [cp t] Emily did [vp e].

The contrast in (82) thus is reduced to that in (84) below. That is, restrictive relatives can but appositives can never extrapose.

(84) a. A man arrived [cp who was wearing a red hat].
   b. *John arrived [cp who was wearing a red hat].

Lasnik, on the other hand, rejects the Quantifier Raising (henceforth QR) account of ACD resolution (May (1985)). There are reasons, however, to believe that this conclusion is drawn prematurely, and it is not clear if we can completely eliminate QR from the list of the computational processes that can induce ACD resolution, although there is no need for us to defend the QR approach either. 31

Returning now to our main topic, suppose that we have a case of ACD involving a there construction as in (85).

(85) [vp V [there to be … [locj … [vp e]]]]

Here, an elided VP is located within the locative associate of there, which in turn is the subordinate subject of an ECM construction. Suppose now that we combine the Attract Locative Analysis with one crucial assumption underlying Hornstein's (1994) analysis of ACD — that the ECM subordinate subject in English is raised out of the higher VP at LF.
We then predict that the infinite regress will be resolved in this construction since the elided VP contained in the locative associate of there in (85) comes to be located outside its antecedent VP at LF, as illustrated in (86).\(^{32}\)

\[(86)\] b. LF: \[\text{Loc} \cdots \[\text{VP}\] \] \[\text{VP} \text{ V} \{t'_1 \text{ to be } \cdots \text{ t}_1\} \]

In constructing relevant examples to verify this prediction, however, we must exercise caution to avoid the possibility of involving other "ACD-resolving" processes discussed above. For this purpose, if we introduce an elided VP in a relative clause as we often do, we would like to make sure that this relative clause is appositive rather than restrictive, and that it does not modify an Accusative-marked or quantified nominal. First, let us examine the ACD in (87a)-(89a).

\[(87)\] a. *Watson \[\text{VP} \text{ believed there to have been a ring given to Susan, who Holmes didn't \[VP\] e \}].

b. Watson believed there to have been a ring given to Susan, who Holmes didn't \[VP believe there to have been a ring given to \].

\[(88)\] a. *Watson \[\text{VP} \text{ expected there to be a dance next Sunday, when Holmes didn't \[VP e \}].

b. Watson expected there to be a dance next Sunday, when Holmes didn't \[VP expect there to be a dance \].

\[(89)\] a. *NBC \[\text{VP} \text{ put a TV camera directly under the Golden Gate Bridge, where ABC didn't \[VP e \}].

b. NBC put a TV camera directly under the Golden Gate Bridge, where ABC didn't \[VP put a TV camera \].

In (87a) and (88a), all the conditions we imposed upon us above are satisfied but crucially the DP containing the elided VP is located in something other than a locative associate of there.

In (89a), the elided VP appear within a locative expression but there is no there for it to be associated with. Under these circumstances, the sentences cannot be interpreted on a par
with the non-elliptical sentences in (87b)-(89b), respectively. That is, ACD is not resolved in (87a)-(89a).

On the contrary, when we introduce the elided VP within the locative associate of there, satisfying all the other conditions, as in (90a) and (91a) below, it is noticeably easier to interpret these sentences on a par with those without ellipsis, as in (90b) and (91b).

(90) a. 7Watson [vp expected therej to be something hidden [Locj in Big Ben, where Holmes didn't [vp e ] ] ].
   b. Watson expected there to be something hidden in Big Ben, where Holmes didn't [vp expect there to be something hidden ].

(91) a. 7Watson [vp expected therej to be a riot [Locj in Paris, where Holmes did [vp e ] ] as well.
   b. Watson expected there to be a riot in Paris, where Holmes did [vp expect there to be a dance ] as well.

ACD in (90a) and (91a), in other words, can be resolved, as predicted in the Attract Locative Approach. The examples in (90a) and (91a) also contrast with those in (92a) and (93a) below, which cannot be interpreted on a par with the non-elliptical sentences in (92b) and (93b).

(92) a. *Watson [vp expected that therej would be something hidden [Locj in Big Ben, where Holmes didn't [vp e ] ] ].
   b. Watson expected that there would be something hidden in Big Ben, where Holmes didn't [vp expect that there would be something hidden ].

   b. Watson expected that there would be a riot in Paris, where Holmes [vp expected that there would be a riot ] as well.

In (92a) and (93a), there shows up as the subject of a finite clause, and hence is not expected to be covertly raised out of the matrix VP. The contrast here, in other words, is also predicated under the proposed analysis.
Steven Franks (p.c.) pointed out to me, however, that infinite regress can be resolved even when the elided VP is located in a temporal adjunct, as in (94a), or a causal adjunct, as in (94b), in the *there* construction.

(94) a. Watson always [{vp expects there to be something hidden }] [{cp whenever Holmes does [vp e ] }].
    b. Watson always [{vp expects there to be something hidden }] [{cp because Holmes does [vp e ] }].

These facts open up the possibility that the successful resolution of infinite regress in (90a) and (91a), after all, might not be caused by covert movement of the locative expression but by some unknown factor common to all of (90a), (91a) and (94a-b). There is, however, evidence which reveals that (94a-b) are structurally distinct from (90a) and (91a), and this structural distinction can be translated into two distinct ways the resolution of infinite regress can be achieved. VP-fronting provides us with a crucial test for the structural distinction in question. First, as illustrated in (95) and (96) below, VP-fronting is optional for both temporal and causal expressions — note that they may or may not be fronted as part of a VP.

(95) VP-Fronting with **Temporal** Expressions:

I thought he might expect there to be a problem whenever the offer was attractive, and

a. [{vp1 expect there to be a problem whenever the offer was attractive }], he really did e₁!
    b. [{vp1 expect there to be a problem }], he really did e₁ whenever the offer was attractive!

(96) VP-Fronting with **Causal** Expressions:

I thought he might expect there to be a problem because he doesn’t trust us, and

a. [{vp1 expect there to be a problem because he doesn’t trust us }], he really did e₁!
    b. [{vp1 expect there to be a problem }], he really did e₁ because, as I said, he doesn’t trust us!

Crucially, then, the possibility of VP-fronting as in (95b) and (96b) suggests that temporal and causal phrases can be analyzed as items located outside the relevant VP. From this, it
follows that the elided VP located in a temporal or a causal expression as in (94a-b) does not cause any infinite regress to begin with.

In sharp contrast to temporal and causal expressions, however, the locative associate of there undergoing VP-fronting has no choice but to be fronted as part of a VP, as illustrated by the ill-formedness of (97b).

(97) VP-Fronting with Locative Associate of there:

I thought he might expect there to be some problem in our plan, and

a. \([v_p] \text{expect there to be a problem in our plan}, \text{he really did } e_1!\]
b. \(*[v_p] \text{expect there to be a problem }, \text{he really did } e_1 \text{ in our plan}!\]

This demonstrates that the locative associate of there must be located within the VP. We have systematically shown, in other words, that (94a-b) are inherently immune to the problem of infinite regress while (90a) and (91a) can not escape the problem in the same way. This way, we can restore the validity of our argument for the Attract Locative Analysis above, which would account for the otherwise puzzling resolution of infinite regress in (90a) and (91a).33

Recall finally Lasnik's (1995) claim that the ill-formedness of (98) below arises because there attracts only the formal features of its DP associate.

(98) *John \([v_p \text{expected } [i_p \text{there to be } [d_p \text{no one that I did } [v_p e ] \text{electable }]]].\]

This approach fails to explain why (90a) and (91a) can ever escape the same problem since "Attract FF (associate DP)" completely disregards the content of the locative expression. The Attract Locative Analysis, on the other hand, can account for the well-formedness of (90a) and (91a) while staying perfectly consistent with the ill-formedness of (98) at the same time since this analysis does not entail attraction of any content of the post-verbal DP. (We will
briefly discuss how agreement should be handled in the Attract Locative Analysis shortly below.)

To sum up so far, we have observed that the phenomena of locality of anaphor binding, bound variable interpretation of pronouns, scope interaction and antecedent contained deletion all support the Attract Locative Analysis of there, which incorporates the claims in (99).

(99) a. A locative expression rather than a post-verbal DP is the referential associate of there.
   b. There inherits the entire content of the locative expression, not just its formal features.

There is, however, one empirical phenomenon that remains somewhat recalcitrant in the Attract Locative Analysis. Observe first the successful licensing of a negative polarity item any in (100a-b).

(100) a. Overt Raising (to Subject):

   No good linguistic theories seem to any philosophers [ to t have been formulated ]
   \[ \uparrow \]

   b. Covert Raising under ECM:

   The DA proved no one [IP t to be at the scene ] during any of the trials.
   \[ \uparrow \]

According to Lasnik (1995), the negative subordinate subject in both of these sentences undergo overt and hence total raising, and come to c-command the polarity item any located in the matrix. Similar negative polarity is not licensed, however, in (101a-b) below, in which there c-commands any after overt movement raises it to the matrix, but its associate DP as a negative item is located in the subordinate clause.
(101) a. Overt Raising (to Subject):  

*There* seem to *any* philosophers [ *t* to have been *no good linguistic theories*  

↑________________________|↑_________Covert________|  

formulated ]

b. Covert Raising under ECM:  

*The DA proved *there* [t_{IP} *t* to be *no one* at the scene ] during *any* of the trials.  

↑_____|

The ungrammaticality of (101a-b) thus seems to demonstrate that the negative content of the associate DPs is not attracted to *there*. This is one of the facts that urged Lasnik to support Move FF (DP-associate).

Note that all these facts are also compatible with the Attract Locative Analysis, which does not affect the post-verbal DPs. We predict in the Attract Locative Analysis, however, that, if a negative item shows up as the locative associate of *there*, its content should be attracted to *there* and license *any* in a similar construction. Contrary to our prediction, however, virtually all speakers still find negative polarity in (102) problematic.

(102) ?*~*There* seem to *anyone* [ *t_{1}* to be problems *nowhere in the city* ]  

↑________________________|↑_________Covert_________|  

Mamoru Saito (p.c.) pointed out that this might in fact be a predicted result under the Attract Locative Analysis since the combination of Attract Locative and a possible covert adjunction of *any* to the negative item involved in its licensing might create an ill-formed chain in (102). There, however, are speakers who report that the source of awkwardness involved in this sentence perhaps is not the failure of negative polarity, which would have caused even severer offense as in (101) above and (103) below.

(103) a. *Anyone did not* come to the party.  

b. *Anyone came to the party.
It can be shown, in fact, that there is a factor which yields awkwardness in (102) irrespectively of its covert computation. The contrast observed in (104) and (105) below suggests that a negative item must precede a polarity item in its surface order to become fully acceptable.

(104) a. **Never** in my life, have I felt anyone's affection.
   b. *I have [\textit{never} felt anyone's affection ] never in my life.

(105) a. **Nowhere in Germany**, does the possibility of meltdown [ seem to anyone to exist in a nuclear plant in the next province ].
   b. *Possibility of meltdown [ seems to anyone to exist in a nuclear plant in the next province ] nowhere in Germany.

The acceptability of (106a-b) below indicates that the awkwardness of (104b) and (105b) indeed stems from the licensing of negative polarity.

(106) a. I have felt her affection never in my life.
   b. The possibility of meltdown [ seems to a local official to exist in a nuclear plant in the next province ] nowhere in Germany.

The contrast in each of (107) and (108) below with respect to superiority effect also demonstrates that the negative expressions do c-command the polarity items in (104b) and (105b).^{44}

(107) a. On which occasion, did you [ feel whose affection ] t,?
   b. *Whose affection, did you [ feel t ] on which occasion,?

(108) a. In which German province, does the possibility of meltdown [ seem to which congressman to exist in a nuclear plant in the next province ] t,?
   b. *To which congressman, does the possibility of meltdown [ seem t, to exist in a nuclear plant in the next province ] in which German province,?

What went wrong in (104b) and (105b), in other words, indeed seems to be the relative linear order between a negative expression and a polarity item. Although I have no definite idea
why linear order must be relevant to the licensing of negative polarity, one possibility I can think of is that it involves a processing problem — when processing reaches the negative polarity item *any* before it hits a negative item, a language user might receive the impression that it is a "free choice" *any*, and when the negative expression is encountered later, a "garden path" type of effect might arise.

This discussion leads us to the Attract Locative Analysis of negative polarity in a sentence like (109).

(109) The FDA inspector could \[ _{VP} \text{prove} [ \text{there to have been cockroaches nowhere in the building} ] \text{ during any of the inspections} \].

In this sentence, a negative locative expression precedes a polarity item *any* at surface, and *there* shows up as an ECM subordinate subject. Negative polarity therefore is expected to be licensed here under the Attract Locative Analysis. Some of my informants indeed accept this sentence while others do not. It is not very clear what exactly is the source of this variation at this moment. It seems to be the case, however, that negation expressed in this sentence has a bit of handicap to begin with to be accepted as fully natural locution. In order to make this sentence natural, we must imagine a pragmatic context in which the temporal phrase containing the polarity item as well as the locative phrase is taken up as part of a focus and be strongly negated, but such a situation is not very easy to imagine. The best I can think of is the situation in which the FDA inspector in question desperately repeated his inspection, in vain, to prove that there had been cockroaches at least in some spot in the building. The sentence in (109) to be interpreted in this context will be most naturally pronounced with two (paired) intonational peaks placed on *nowhere* and *any*. Thus, the overall impression of
negative polarity facts is that they tend to fall short of offering clear support to the Attract Locative Analysis, but for good reasons.

One non-trivial issue that we have not discussed so far is agreement. The operation Attract Locative does not appear, at least at first sight, to directly control the number agreement observed between the post-verbal DP and the verb in the there construction. A couple of possibilities come to mind. First, the agreement in question, in fact, might be established directly between the verb and the post-verbal DP within the verbal projection. Since the notion "checking domain" (Chomsky (1993)) is rather artificially defined, and the notion like "Case of unaccusative" (Belletti (1988)) suggests independent need for establishing such relations, this option is not out of the question. Another possibility is that a locative expression inherits number from the post-verbal DP in one way or another, and agrees with the verb (via T) at LF when it is attracted to there. den Dikken and Næss (1993), for instance, argue that a locative expression as a predicate agrees with a post-verbal DP as its subject in the locative inversion construction. That locative PPs can participate in number agreement from a subject position can be also demonstrated by (110a-b).

\[(110)\]  
\[\text{a. [pp Under the chair] is a nice place for the cat to sleep. (Stowell (1981))} \]
\[\text{b. [pp Under the chair] and [pp on the TV set] are two nice places for the cat to sleep.}\]

Whichever of these analyses may turn out to be correct, we perhaps need to postulate a chain of agreement to explain why the agreement in the subordinate clause is inherited in the matrix in (111).

\[(111)\]  
\[\text{There seems [ip t_i to be a problem in the main cabin].}\]
If the DP-associate agrees with \textit{be} within VP as in (112a) below, \textit{there} must be assumed to agree with \textit{be} (in T) as in (112b) before it gets raised and agrees with the matrix verb \textit{seems} as in (112c).

(112) a. \textit{seems [ there [$_T$] [$_{VP}$ \textbf{be a problem in the main cabin} ] ]}

\hspace{1cm} [.............]

b. \textit{seems [ \textbf{there} [$_T$ \textit{be$_2$} ] [$_{VP}$ t$_2$ a problem in the main cabin ] ]}

\hspace{1cm} [.............]

c. \textbf{There}, \textit{seems [ t$_1$ [$_T$ \textit{be$_2$} ] [$_{VP}$ t$_2$ a problem in the main cabin ] ]}

\hspace{1cm} [.............]

If the locative and the DP-associate agree within VP as in (113a) below, on the other hand, the matrix verb comes to agree with the locative as in (113d) after \textit{there} is overtly raised as in (113b) and the locative is covertly attracted to \textit{there} as in (113c).

(113) a. \textit{seems [ there [$_{VP}$ be \textbf{a problem in the main cabin} ] ]}

\hspace{1cm} [.............]

b. \textbf{There$_1$}, \textit{seems [ t$_1$ [$_{VP}$ be a problem in the main cabin ] ]}

\hspace{1cm} [............................]

c. \textit{in the main cabin$_1$}, \textit{seems [ t$'_1$ [$_{VP}$ be a problem t$_1$ ] ]}

\hspace{1cm} [............................]

d. \textit{in the main cabin$_1$}, \textbf{seems [ t$'_1$ [$_{VP}$ be a problem t$_1$ ] ]}

\hspace{1cm} [............................]

If either one of these analyses is tenable, the Attract Locative Analysis can completely replace \textbf{Move FF (associate DP)}. Even if neither of them turns out to be correct, on the other hand, we can still adopt, \textbf{in addition to} Attract Locative Analysis, Chomsky's \textbf{Move FF (associate DP)} to T, in which agreement takes place directly between the adjoined FFs of associate DP and T without mediation by \textit{there}.  

43
3.1.3 Attract Clause

In this subsection, I will attempt to motivate what I call the "Attract Clause" Analysis of the expletive *it* as summarized in (114).

(114) a. The so-called expletive *it* is anaphoric to CP.  
    b. It covertly attracts the associate CP for referential saturation.\(^{36}\)  
    c. The entire content rather than just formal features of the associate CP is attracted.

The Attract Clause Analysis is illustrated by the derivation in (115).

(115) a. Identify:  
   \[
   \text{LF}: \quad \text{It is unlikely } [\text{CP that he will win}].
   \]
   \[
   \]
   b. Attract:  
   \[
   \text{LF}: \quad [\text{CP that he will win} \text{ is likely t.}]
   \]
   \[
   \]

Descriptively, we can state that Attract Clause takes place when the pronominal *it* is generated in a non-thematic position, and *c*-commands the CP it is identified with as exemplified by (115a). When the entire content of the associate CP is transferred onto *it* as in (115b), this proform becomes referentially saturated. (We will discuss this mode of referential saturation in further detail in 3.3 below, and show how it can avoid BT (D) violations.) In a sense, this LF-derivation is a covert counterpart of Emonds's (1970) "intraposition" analysis of clausal subjects.

It has long been recognized, both in traditional grammar and generative grammar, that the expletive *it* is anaphorically or otherwise associated with the post-verbal CP.\(^{37}\) It therefore is not entirely misdirected to state that the Attract Clause Analysis captures our intuition quite faithfully. The major observation which has made researchers shy away from this analysis in the minimalist tradition, I believe, is that the agreement comparable to that
observed in the there construction is missing in the CP-extraposition construction. In what follows, I will first present one empirical motivation for adopting the Attract Clause Analysis. I will then attempt to show that the premise of the rejection of the Attract Clause Analysis just described, namely, the assumption that the expletive it does not exhibit number agreement is erroneous.

The empirical argument for the Attract Clause Analysis is based upon the observation offered by Chomsky (1981) concerning the locality of anaphor dependency. We actually extended it to the analysis of there above, and the form of the argument is basically the same here. First, anaphor binding is disrupted by an intervening DP as in (116) (as we have already seen in 3.1.2).

(116) *They know [ that I believe [ that each other's books will be best-sellers ]].

Quite interestingly, an anaphor can be non-locally bound when the intervening SUBJECT is the expletive it as in (117) below:

(117) They think [ that it is unlikely [ that each other's books will be best-sellers ]].

One crucial observation here is that each other is located in the CP associate of the intervening "expletive" it. This situation certainly reminds us of a similar observation in regard to there, and naturally leads us to a solution similar to Attract Locative. If we adopt the Attract Clause Analysis here, the sentence in (117) comes to be represented as in (118) below after the CP associate is attracted to it at LF:

(118) LF after "Attract Clause":

↓
LF:  They think [ that [CP that each other's books will be best-sellers ] is unlikely  t ]

In this representation, the locality constraint on anaphor binding is in fact satisfied, since each other is locally bound by its antecedent they. This motivates the "Attract Clause" Analysis, under which the entire content of the CP associate including the contained anaphor is attracted to it. As Lasnik and Uriagereka (1988) pointed out, on the other hand, mere presence of the expletive it does not necessarily override the locality restriction on anaphor binding, as shown in (119a-b).

(119) a.  They said that [ we think that it is unlikely [ that each other's books will be best-sellers ]].

b.  They think that it surprised each other [CP that Bill won ].

These facts also follow naturally from the "Attract Clause" Analysis. In (119a), even after the associate CP is attracted to it, the intended anaphor binding between they and each other is not local because of the intervention of we. In (119b), Attract Clause does not improve the situation since the anaphor is not contained in the associate CP. When the order of the intervening plural NP and the expletive it as in (119a) is reversed, and when the anaphor in (119b) is contained in the associate CP, on the other hand, Attract Clause becomes relevant again and an anaphor can be non-locally bound across the expletive it. Compare (120a-b) below with (119a-b).

(120) a.  They think that it surprised us [ that each other's books were on sale ].
b. **They** think that **it** surprised Bill [that each other's books were on sale].

Let us now turn to the issue of agreement. Since his early minimalist analysis (Chomsky (1986)), Chomsky has maintained the claim that **there** triggers the covert movement of part or whole of its associate while **it** does not. Chomsky (1995, pp. 273-274) attempts to justify this asymmetry with the assumption that **it** has the capacity to check all of categorial D, Case- and Φ-features of T while **there** lacks the capacity to check Case- or Φ-features of T. **There**, in other words, drives the application of Move FF (associate) for the convergence of derivation, while the Last Resort Principle prohibits its application to **it**.  

McCloskey (1991) further argues for the absence of LF-replacement of the expletive **it** by its CP associate in the following way. First, the subject-verb agreement in (121) below shows that conjoined propositions denoting "distinct situation-types" are plural in number.

\[(121) \ [\text{CP} \ [\text{CP} \text{That the president will be reelected }] \text{ and } \text{CP} \text{that he will be impeached }] \text{ are equally likely at this point.}\]

The expletive **it**, however, does not exhibit plural agreement with a verb even when it has a plural conjoined CP of this sort:

\[(122) *\text{It seem equally likely at this point } \text{CP} \text{[CP that the president will be reelected ] and CP that he will be impeached }.\]

According to McCloskey, this demonstrates that the expletive **it** does not participate in number agreement with its CP associate, and hence that it does not undergo Attract Clause in our terms (whether it attracts only formal features or the entire content). Note, however, that **it** is inherently singular, and is disagreeing with the inflection of the verb **seem** in (122). The sentence, in other words, is bound to be ungrammatical irrespective of the agreement between
it and the CP-associate. In this regard, the example involving the plural form they as in (123) below is more relevant, and still supports McCloskey's point.

(123) *They are equally likely at this point \([_{CP} \text{that the president will be reelected}] \) and \([_{CP} \text{that he will be impeached}] \). 

At this point, there arise two questions that need to be answered under the Attract Clause Analysis. First, why and how the anaphoric relation in (124a) below is prohibited while such is not a problem in (124b).

(124) a. *They are equally likely \([_{CP} \text{that John will win and that he will lose}] \). 
   b. \([_{CP} \text{That John will win and that he will lose}] \) are equally likely, but they can't both happen.

Since the Affect Approach does not discriminate expletives from "regular" anaphoric items, this asymmetry is unexpected. Second, why can the plural proform they not serve as an expletive while the singular proform it can? How can we explain, in other words, the unsuccessful plural agreement as in (124a) above and the successful singular agreement in (125a-b) below at the same time?

(125) a. It seems \([_{CP} \text{that he will win}] \). 
   b. It seems equally likely at this point \([_{CP} \text{that the president will be reelected and that he will be impeached}] \). (McCloskey (1991, p. 565))

In what follows, I would like to show that there is a systematic reason why the plural agreement as in (124a) fails to surface. If successful, such an attempt will allow us to neutralize McCloskey's counterargument to the Attract Clause Analysis, which seems to be the strongest objection to this analysis raised in the literature.

I believe that the key to identifying the actual but hidden problem here is to understand the nature of plurality involved in the examples like (124a-b). Here, the crucial and perhaps obligatory use of the expressions equally and both in (124a-b) provides us with an important clue. First, a sentence of the form in (126a) is interpreted as in (126b).
(126) a. CP₁ and CP₂ are equally likely.
    b. CP₁ is likely, and CP₂ is likely, and the degree of likelihood in two cases is equal.

The likelihood (or at least possibility) of the realization of each individual situation in (126a), in other words, is taken up and their relation is asserted. Likewise, a sentence of the form in (127a) is interpreted as in (127b).

(127) a. CP₁ and CP₂ can both happen.
    b. CP₁ can happen, and CP₂ can happen, and the realization of each situation is not mutually exclusive.

Here again, the potential realizability of each individual situation in (127a) and their relation is asserted. The use of equally and both in the contexts like (126a) and (127a), in other words, seems to induce a distributive interpretation of a plural set of events rather than its collective interpretation.

Arguing that Binding Condition B (henceforth BT (B)) operates on semantic predicates, Reinhart and Reuland (1993, 677) also make a crucial appeal to this property of both to account for the contrast in (128).

(128) a. [ Max₁ and Lucie ] talked about him₁.
    b. *Both [ Max₁ and Lucie ] talked about him₁.

The preferred interpretation in (128a) is the collective one, suggesting an act of mutual talking. To most speakers, on the other hand, both in (128b) forces the distributive interpretation, entailing two separate acts of talking by Max and by Lucie. The two interpretations of (128a-b) therefore can be represented as in (129a-b), respectively.

(129) a. Max and Lucie (λx (x talked about him))
    b. Max (λx (x talked about x)) & Lucie (λx (x talked about him))

Under the collective interpretation in (129a), the two NPs making up the subject are taken as one set, and the predicate is not regarded as reflexive. Under the distributive interpretation in
(129b), on the other hand, the predicate is distributed over the two distinct NPs, and one of them \( (x \text{ talked about } x) \) is reflexive, which leads to the violation of BT (B).

Thus, we may characterize the plurality involved in (124a-b) (and possibly the plurality of CPs in general) as distributive (or aggregate) in nature rather than being collective. That is, the individuality of each conjunct CP is retained in the coordinated CP.

The plural pronoun they in each of these examples, in other words, is anaphoric to the two conjunct CPs rather than the entire coordinate structure, much in a way similar to the case of "split antecedents", as indicated by the coindexation in (130a-b).

(130) a. \([\text{CP1} \text{ That John will win }] \text{ and } [\text{CP2} \text{ that he will lose } ]\) are equally likely, but they\(_{1+2}\) can't both happen.

b. *They\(_{1+2}\) are equally likely at this point \([\text{CP1} \text{ that the president will be reelected }] \text{ and } [\text{CP2} \text{ that he will be impeached } ]\).

The sentence (131) below exemplifies similar split antecedents involving conjoined DPs denoting "aggregate" plurality.

(131) \([\text{DP1} \text{ Michael Jordan }] \text{ and } [\text{DP2} \text{ Larry Bird } ]\) are perhaps two of the most well-known players in the history of NBA, and coincidentally they\(_{1+2}\) were both named after an uncle.

Note the singular expression "named after an uncle", which clearly suggests that the individuality of each antecedent DP is retained. So, the question now can be restated — why are the split antecedents permitted in (130a) but not in (130b)? I would like to point out here that the contrast follows naturally when we adopt the Affect Approach to anaphoric interpretation. In particular, the answer lies in the two distinct ways the pronoun they is saturated at LF in these sentences. First, in (130a), there does not exist a c-command relation between the pronoun and the two conjunct CPs it has been identified with. What must take place at LF for saturation, therefore, is the application of covert Copy out of a coordinate structure, as illustrated in (132).\(^{39}\)
(132) LF: 
\[
[\text{CP}_1 \text{ That John will win }] \text{ and } [\text{CP}_2 \text{ that he will lose }] \]
are equally likely,
\[
\text{but } \{ \text{CP}_1 \text{ That John will win } \}
\]
can't both happen.
\[
\text{CP}_2 \text{ that he will lose }
\]

As we have already seen, similar split antecedents with DP conjuncts as in (131) above and (133a) below are also possible, and covert Copy out of a coordinate structure as illustrated in (133b) must be permitted.

(133) a. 
\[
[\text{DP John}_1 \text{ and Bill}_2 ] \text{ have returned, but they}_{1+2} \text{ are both injured.}
\]
b. 
\[
\text{LF: } [\text{DP John}_1 \text{ and Bill}_2 ] \text{ have returned, but } \{ \text{Bill}_2 \} \text{ are both injured.}
\]

In (130b), on the contrary, the pronoun they generated in a non-thematic position commands the two conjunct CPs it has been identified with. Saturation of this pronoun, therefore, must be achieved by the application of Attract rather than Copy, as illustrated in (134).

(134) 
\[
\text{LF: } *[\{ \uparrow \text{CP}_1 \text{ that the president will be reelected } \text{ at this point [ t and t }. \}
\]
\text{CP}_2 \text{ that he will be impeached ]}
\]

As illustrated in (135) below, however, extraction of conjuncts out of a coordinate structure is known to be prohibited (Ross (1967)).

(135) *What\text{$_{1+2}$} did you buy [\text{DP t}_1 \text{ and t}_2 ]?

We can, in other words, reduce the contrast between (130a) and (130b) to the different option of Affect to be adopted for the saturation of the pronoun in each construction.
One fact that we still must account for is why the singular "expletive" *it* is possible in (136) (= (125b)).

(136) **It** seems equally likely at this point [*CP that the president will be reelected and that he will be impeached*].

Here, it must be the case that the entire coordinate CP is covertly attracted to *it* and that this CP is singular. Somewhat surprisingly, in fact, many speakers accept (137) below not only with *are* but also with *is*.

(137) [*CP That the president will be reelected and that he will be impeached*] { *is* / *are* } equally likely at this point.

It must be the case, therefore, that *equally* does not necessarily force a distributive interpretation and the entire coordinate structure in (136) can be singular, though the semantics involved in this interpretation is not at all clear to me. This conclusion makes (136) at least compatible with our analyses above. Since this sentence involves covert attraction of the entire coordinate CP rather than that of the individual conjuncts out of a coordinate structure, no problem is expected to arise. Perhaps, the entire coordinate CP is regarded as a single set and hence as a singular element under the collective interpretation whereas the individual conjunct CPs as members of this set are regarded as plural under the distributive interpretation. While this asymmetry shows up only as distinction in number agreement in (137), the same asymmetry permits *it* in (136) but rejects *they* in (130b) as the attractor of CPs because only (130b) involves attraction of CPs out of a coordinate structure.

To sum up, the ungrammaticality of (130b) can be ascribed to the general prohibition against the extraction of individual conjuncts out of a coordinate structure under the Attract Clause Analysis of expletives. If credible, this conclusion permits us to consider that the CP-extraposition construction does in fact involve the subject-verb agreement that reflects the
anaphoric relation between the subject proform and its associate CP. When a singular proform \textit{it} is generated in a typical CP-extraposition construction as in (138) below, the singular number of its associate CP is straightforwardly reflected by the verbal inflection.

(138) \textbf{It seems} \textit{[\text{CP} that he will win]}. When a plural proform \textit{they} is generated in a CP-extraposition construction as in (130b) above, on the other hand, its association with a plural CP necessarily involves illicit application of Attract, and therefore the expected plural agreement fails to surface legitimately. There is, in other words, a systematic reason why the CP-extraposition strikes as not involving number agreement comparable to the agreement the \textit{there} construction exhibits. The strongest objection to the Attract Clause Approach therefore can be dismissed. By adopting the Attract Clause Analysis, on the contrary, we can capture the notion 'CP associate' of the expletive \textit{it} in terms of a pure anaphoric relation between these items.  

3.1.4 Theoretical Consequences:

In the literature of generative syntax in 1980s and 1990s, it has been simply assumed that UG makes available expletives as a small class of exceptional linguistic entities that lack a referential function. We can, however, ask a number of somewhat naive but perfectly legitimate questions as to the properties of expletives characterized in this way, as in (139).

(139) a. Why are expletives always proforms? 
b. Why can't they (co)refer unlike other proforms? 
c. Why are they the counterpart of \textit{it} (non-animate proform) and \textit{there} (locative proform) cross-linguistically, and not \textit{he} or \textit{she}, for example?

I believe that the approach incorporating Attract Locative and Attract Clause can provide quite straightforward answers to all of these questions. To begin with, the basic assumption underlying this approach is that exceptional linguistic entities called expletives need not be postulated in UG. They are proforms that are base-generated in a non-thematic position, and
come to head a chain after undergoing Attract at LF. If any referentially saturated item is
generated in this non-thematic position, the Last Resort Principle would prohibit it from
attracting its associate, and hence from establishing a chain involving a thematic position. If,
on the other hand, some meaningless non-proform is inserted in this position, it will fail to be interpreted in the end. In both cases, FI (and eventually the Bare Output Condition) will be violated. Therefore, nothing other than unsaturated proforms can show up as an "expletive", which answers the question in (139a). The question in (139b) is asked based upon the premise that if an item is generated in a non-thematic position and still satisfies FI, it cannot have a referential function. The Attract approach rejects this premise, however, by motivating a chain established at LF when the associate of the proform is attracted from an interpretable position within a lexical projection to a position under a functional projection. Proforms used as "expletives" therefore need not be regarded as different from proforms used in other positions.

Finally, why are expletives limited to there and it (= (139c))? The answer, I believe, lies in their Case properties. Suppose, first, that he is generated in a non-thematic position and identified with a referential DP as in (140a) below. This proform then should covertly attract its antecedent John, as illustrated in (140b).

(140) a. PF/LF; : He is likely [ John to win ]

    [___________]

   b. LF; : John is likely [ t to win ]

   ↑___________

We note, however, that the Case feature of one of the two DPs (He and John) in (140a) is bound to remain unchecked in this derivation since the intrinsic Case-assigning feature shows up here only on the matrix tense. Crash of LF-derivation in (140) thus can be ascribed to the
failure to check off the Case feature of one of the DPs. What then rules out the LF-derivation in (141) below, in which another inherent Case-assigning feature shows up on the subordinate tense?

(141) a. PF/LF₁: He is likely [ that John will win ]
   \[
   \text{id} \]

   b. LF₂: John is likely [ that t will win ]
   \[
   \uparrow \text{id} \]

Note that the Case feature of both DPs can now be checked off and the problem observed in (140) should not arise here. In the Principles and Parameters approach, it is generally assumed that the successful interpretation of a nominal argument requires biunique association of \(\theta\)-marking and a Case relation within its chain (including a trivial chain).

While the "Visibility" condition is often said to ensure the presence of a legitimate Case relation necessary for successful \(\theta\)-marking, we may also consider that the economy consideration prohibits \(\theta\)-marking from being associated with more than one superfluous Case relation in a chain. The LF-derivation in (141) then will be ruled out due to the involvement of two Case relations — one between He and the matrix tense and the other between John and the subordinate tense — for one instance of \(\theta\)-marking.

Among the proforms, only there can refer to a locative expression, and it can refer to a singular CP. One notable property of locatives and CPs is that both of them need not enter any Case-checking relation. It is true that a CP can enter a Case checking relation, for instance, as a subject of a finite clause, as in (142).

(142) \(\text{cp} \) That the President was not impeached ] surprised us all.
It, however, does not necessarily have to do so even as an argument. A CP may appear, for instance, as a complement of *seem* as in (143a) below, while the ill-formedness of (143b) indicates that this verb lacks a Case feature.

(143) a. It seems [\textit{CP} that he is innocent].
  b. *It seems [\textit{him} to like it].

Obviously, in other words, if the "visibility" condition is to be maintained, it constrains only nominal (DP) arguments. When *there* and *it* attract their associates and create chains at LF, therefore, neither of the problems observed in (140) and (141) above arise.41

To sum up, the Attract Locative/Clause Analysis can provide simple answers to all of the questions in (139), and eliminates "expletives" as special linguistic entities.

Finally, let us emphasize again that if Attract Locative/Clause Analysis is plausible, we will acquire quite reasonable grounds for adopting the view that covert Affect be characterized as in (144a-b) below, departing from the current "Chomskian" view of computational processes in general.

(144) a. Covert Affect can be triggered by semantic factors.42
  b. Covert Affect can manipulate not just formal features but the entire content of an item
  (including its semantic features).

With this conclusion in mind, let us now turn to the motivation for other options of Affect for anaphoric saturation.

3.2 Identify

The essence of "Identify" as a computational process has been adopted and argued for by a number of researchers ever since Higginbotham (1983) proposed "Linking" (e.g., Seely (1988), Fox (1998)). It therefore is the option of Affect which seems to require least justification. Here, I will only discuss the empirical phenomenon which allows us to confirm
why we should adopt Identify over coindexation, adopting and reinforcing Heim's (1992)
argument.43

First, the sentence in (145) below can be ambiguously interpreted as illustrated by the
two distinct near paraphrases in (146a-b).

(145) Every boy, thinks that only he, likes his, mother.
(146) a. Every boy thinks that other boys do not like their own mothers.
     b. Every boy thinks that other boys do not like his mother.

Roughly speaking, the interpretation illustrated in (146a) concerns "other boy's liking of their
own mothers", while that illustrated in (146b) concerns "other boy's liking of one boy's
mother". Crucially, however, the coindexation indicated in (145) fails to distinguish these
two interpretations. We can, on the other hand, establish two distinct binding relations as
illustrated in (147) in terms of two distinct modes of the application of Identify.

(147) a. Every boy thinks that only he like his mother. (Serial Binding)
               |___________________||______|
               ________________
     b. Every boy thinks that only he like his mother. (Parallel Binding)
               |___________________|

In (147a), two applications of Identify established two instances of binding in a "serial"
fashion, which yields the reflexive interpretation between he and his. In (147b), on the other
hand, two applications of Identify established two instances of binding in a "parallel" fashion,
which distinguishes he as the person who likes the mother from his as the son of the mother.

At first site, the reading in (146b) may strike as a case involving multiple guises — one guise
as the person who (thinks that he) likes the mother and another guise as the son of the mother,
and one may attempt to capture this reading with coindexation as illustrated in (148).
It seems inappropriate, however, to regard he in (148) as G-linked, since it is interpreted as a variable bound by its quantified antecedent every boy. Another possibility is to postulate G-linking and R-linking as in (149).

This analysis, however, would require us to analyze a quantified element as having a referent. Even if we ever allow ourselves to do that, we would have to regard a (semantically) singular noun his and every boy as sharing an identical referent. The reading in (146b), however, certainly does not require everyone to be interpreted as a unit set.

The same ambiguity persists in (150) below, and this fact also disfavors the multiple guise analysis of the reading in (146b).

Every child thinks that only he or she likes his or her mother.

Note that the disjunctive pronominal his or her in (150) cannot be associated with any particular referent (via a guise) when it is interpreted as a variable bound by the quantified element. Both instances of his or her in (150), in other words, are saturated by dependency even when the reading in (146b) is obtained. Thus, we have good reason to believe that the
reading as in (146b) is not made available by multiple guises, but with "parallel" binding as in (147b).

To sum up, the ambiguity in (145) can be captured with "Identify" but not with coindexation. This provides us with appreciable empirical motivation to adopt the former over the latter.

3.3 Copy

We can also motivate "Copy" as a covert computational process in several different ways. The first argument arises from examination of the so-called Null Object Construction in Japanese as in (151B) below, which may yield both sloppy and strict identity interpretation as indicated by the English translation.

(151) A: John-wa zibun-o nagusameta ga,  
    -TOP self-ACC consoled though
    'John consoled himself.'

B: Bill-wa [NP e] semeta.  
    -TOP blamed
    = 'Bill blamed himself.'
    = 'Bill blamed John.'

Here, sloppy identity is possible even when the two sentences involve VPs headed by distinct verbs, and hence apparently without involving VP-Ellipsis. While Otani and Whitman (1991) claim that the null objet construction in Japanese as in (151B) involves "VP Ellipsis in disguise", the validity of this hypothesis has been questioned by Hoji (1998) and Kitagawa (1999). Sloppy identity in this construction, on the other hand, can be derived without any recourse to VP-Ellipsis or \( \lambda \)-abstraction when we postulate covert "NP-Copy". All we must do is simply to combine our Affect Approach with Kuroda's (1965) claim that the null object in this construction is a phonetically empty pronoun, as illustrated in (152).
(152) a. **Identify:**

\[
\text{LF}_{j}: \text{John}-wa [_{NP} \text{zibun-o }] \text{nagusameta ga, Bill}-wa [_{NP} \text{e }] \text{semeta.}
\]

b. **Copy:**

\[
\text{LF}_{j}: \text{John}-wa [_{NP} \text{zibun-o }] \text{nagusameta ga, Bill}-wa [_{NP} \text{zibun-o }] \text{semeta.}
\]

c. **Identify:**

\[
\text{LF}_{j}: \text{John}-wa [_{NP} \text{zibun-o }] \text{nagusameta ga, Bill}-wa [_{NP} \text{zibun-o }] \text{semeta.}
\]

As in (152c), Identify applying after Copy correctly lets \text{zibun} become saturated via dependency in each clause and yields sloppy identity without recourse to any extra mechanism.

For the sake of completeness, the derivation for the strict identity in (151) is also given in (153) below, in which \text{zibun} undergoes Identify and becomes saturated before the NP containing it undergoes Copy.\(^{46}\)

(153) a. **Identify:**

\[
\text{LF}_{j}: \text{John}_{1}-wa [_{NP} \text{zibun-o }] \text{nagusameta ga, Bill}_{2}-wa [_{NP} \text{e }] \text{semeta.}
\]

b. **Copy:**

\[
\text{LF}_{j}: \text{John}_{1}-wa [_{NP} \text{zibun}_{1-o }] \text{nagusameta ga, Bill}_{2}-wa [_{NP} \text{zibun}_{1-o }] \text{semeta.}
\]

The Affect Approach incorporating covert Copy can be further motivated when we examine various types of interpretive restrictions observed in English — the violation of the "i-within-i" filter in (154), "referential circularity" in (155), and the violation of BT (D) in (156).

(154) *\[\text{his, friend}\]
(155) *[His, wife] admires [her, husband]

(156) *He, saw John.

Consider first the so-called "i-within-i" violations exemplified by (154). The interpretive restriction imposed on the representation here — that his and his friend cannot share an identical referent — has been captured by the "i-within-i" condition, which takes the form of a filter prohibiting coindexation as schematized in (157) (Chomsky (1981, p. 212)).

(157) *[βi ... αi ... ]
In the Affect Approach, (154) will be represented as in (158) after Identify applies at LF.

(158) *[his friend]

In (158), the saturation of his is to be achieved by its dependency upon the entire NP his friend. The interpretation of this NP, however, is contingent upon the saturation of his. Intuitively, what goes wrong here is clear — the interpretation of an item is forced to be given in terms of that item itself because of the involvement of dependency and containment at the same time. As a result, the saturation of his cannot be achieved. The "i-within-i" violation in (154) therefore can be regarded as a case of the violation of FI involving failure of referential saturation. In the Affect Approach, we can appeal to the notions "identification", "dependency" and "saturation", and depict this FI violation as a problem that arises when a syntactic state described in (159) below holds at LF, which we will refer to as "Infinitely Regressive Saturation (henceforth IRS)".

(159) Infinitely Regressive Saturation (IRS):

If β is dependent upon α, and α contains β, the saturation based upon the identification of α and β is infinitely regressive.

Informally, IRS depicts the failure of saturation when the identification of an unsaturated item based upon c-command circulates back to that item itself due to containment. I believe
that it is quite understandable in the perspective of compositional semantics that structural containment plays a crucial role in inducing IRS since the interpretation of the whole remains incomplete when the referential content of its part is undetermined.47

In the characterization of IRS, the definition of c-command we have adopted from Reinhart (1983) (repeated here as (160)) plays an important role, reflecting Reinhart's original intention behind this particular definition.

(160) C-command: \((= (17a))\)

\[ \text{Node } \alpha \text{ c-commands } \beta \text{ iff the branching node immediately dominating } \alpha \text{ also dominates } \beta, \text{ where } \alpha \neq \beta. \]

The definition in (160) permits \(\alpha\) to c-command \(\beta\) even when the former dominates the latter. \(\alpha\)'s containment of \(\beta\), therefore, permits \(\beta\) to be dependent upon \(\alpha\) when they are identified.

In (158), for instance, the entire NP \textit{his friend} c-commands \textit{his}, and hence \textit{his} is dependent upon \textit{his friend} when they are identified. The saturation intended by the application of Identify in (158), therefore, gives rise to IRS, and the "i-within-i" violation is reduced to FI.48

With IRS in mind, let us now turn to 'referential circularity' observed in a sentence like (155) (repeated below as (161)), in which it cannot be the case that \textit{his} is anaphoric to \textit{her husband} and \textit{her} is anaphoric to \textit{his wife} simultaneously (Brody (1981), Higginbotham and May (1981), et al.).

(161) *\([j \textit{His, wife } \textit{admires } [i \textit{her, husband }]\]

Providing the analysis in (162) below in his linking approach, Higginbotham (1983, 404) attempts to capture the interpretive constraint here in terms of the transitivity of some anaphoric dependence established by the "containment" of an anaphoric item in the antecedent of another anaphoric item.49
In (162), the pronoun *his* is contained in *his wife*, the antecedent of *her*, and at the same time the pronoun *her* is contained in *her husband*, the antecedent of *his*. As a result, *his* comes to be anaphorically dependent upon *his* itself and *her* also comes to be anaphorically dependent upon *her* itself, forcing these items to be interpreted in terms of themselves. In this approach, the essence of the problem is correctly captured — that a pronoun cannot depend upon itself to establish its reference. I would like to point out, however, that this linking analysis does not characterize the problem precisely enough, involving certain redundancy. I would like to claim further that the Affect Approach incorporating covert Copy permits us to completely assimilate "referential circularity" in (161) to the "i-within-i" violation in (158) above by eliminating this redundancy.

Let us now reanalyze (161), starting with (163) below, in the Affect Approach.

In (163), both *his* and *her* underwent Identify, and *her* comes to be dependent upon *his wife* due to the c-commanding relation holding between the two. *His*, on the other hand, cannot be dependent upon *her husband* since the latter does not c-command the former. As a result, *his* undergoes Copy for its saturation, and *her husband* is duplicated onto it, as illustrated in (164).

Note that the referential circularity involving *his* has disappeared at this point, unlike in Higginbotham's linking approach. A problem, however, still persists. Since *her* was to be
saturated by his wife in (163), and his wife now has turned into her husband's wife in (164),
the saturation of her in (164) (= \(\beta\)) is dependent upon her husband's wife (= \(\alpha\)), which
contains her (= \(\beta\)). That is, IRS defined in (159) (repeated here as (165)) arises in (164) due
to a basically identical dependency-containment pair with that we observed in the "i-within-i"
violation in (158).

(165) Infinitely Regressive Saturation (IRS):

\[ \text{If } \beta \text{ is dependent upon } \alpha, \text{ and } \alpha \text{ contains } \beta, \text{ the saturation based upon the identification of } \alpha \text{ and } \beta \text{ is infinitely regressive.} \]

Note also that IRS cannot be avoided even when dependency is regarded to be established by
the "original" her (= \(\beta_1\)) as in (166) below, since the saturation of her husband's wife (= \(\alpha\)) still
contains the "duplicate" of her (= \(\beta_2\)), which functions on a par with the "original" her (= \(\beta_i\)).

(166) LF:\[ _\alpha [ \text{her}_{\beta_2} \text{ husband }] \text{'s wife } \text{ admires } [ \text{her}_{\beta_1} \text{ husband } ] \]

Under the Copy analysis, in other words, (163) is characterized as involving not two but one
instance of referential circularity. Because of that simplification, we can completely reduce
"referential circularity" to the "i-within-i" violations in terms of IRS arising in their LF-representations.

Interestingly, referential circularity can be avoided when we interpret his in (163)
above as a deictic pronoun, for example, with the speaker's pointing to its actual referent
and/or placing some emphatic stress, i.e., as an item that is saturated from the outset (as
indicated by capitalization with stress) as in (167).

(167) LF:\[ [ \text{HIS wife } \text{ admires } [ \text{her } \text{ husband } ] \]

This contrast is also a natural consequence in the Copy approach. In (167), HIS is saturated
from the outset, and her is saturated by its dependency upon HIS wife. No further application
of Affect therefore is called for, and no IRS arises. The sentence can be interpreted perhaps with multiple guises, as illustrated in (168).

(168)

\[
\begin{array}{c}
\text{LF}_j: \quad [\textbf{His} \text{ wife }] \text{ admires [her husband ]}
\end{array}
\]

One puzzle related to this phenomenon is that referential circularity seems to persist when we interpret her in (163) (instead of his) as a deictic pronoun, as indicated in (169).

(169) \text{LF}_j: \quad [\textbf{His} \text{ wife }] \text{ admires [\textbf{HÉR} husband ]}

In (169), his is yet to be saturated since it is identified with, but is not dependent upon \textbf{HÉR husband} due to the absence of a c-command relation. It therefore will undergo Copy and the multiple guise interpretation comparable to that in (168) will require an LF-representation as in (170).

(170)

\[
\begin{array}{c}
\text{LF}_j: \quad [ [\textbf{HÉR} \text{ husband }] 's \text{ wife }] \text{ admires [\textbf{HÉR} husband ]}
\end{array}
\]

"Duplicate"

"Original"

In this representation, however, the "duplicate" \textbf{HÉR} is not identified with, and hence is not dependent upon, \textbf{HÉR husband's wife}, and no IRS is expected to arise.\footnote{Yet, referential circularity appears to be unavoidable. One may consider that the crucial difference between (167) and (169) is that only the latter involves covert Copy and comes to yield containment}
relation at LF, as in (170). The intended interpretation in (167), however, is available even when we alter the c-command relation between the identified items as in (171a) and induce containment at LF as in (171b).

(171) a. LF:\ [ Friends of [ HÍS wife ] ] admire [ her husband ]

 b. LF:\ [ Friends of [ HÍS wife ] ] admire [ [ HÍS wife ]'(s) husband ]

Containment relation without identification, therefore, does not seem to cause any problem in (170), as expected in the Affect Approach. Another path we can take is to regard the difficulty in (169) as arising from some discourse problem due to the backward pronominalization, i.e., the linear order involved there. The relevance of linear order in fact can be demonstrated with a sentence involving a psyche verb as in (172).52

(172) LF:\ [ Stories about [ her husband ] ] annoy [ HÍS wife ]

Here, we have reversed the linear order of the identified items in (171a), and the problem as in (169) emerges again. As shown in (173) below, however, the backward quantifier binding is permitted in the same configuration.

(173) [ Stories about [ her husband ] ] annoy every woman.

This suggests that the c-command relation between the identified items in (172) is identical to that in (167), and hence no containment relation should arise at LF. I therefore conceive that not only the referential restriction observed in (172) but also that observed in (169) is caused by some discourse factor involving linear order rather than by IRS. If so, (169) is still compatible with the Affect Approach.

Finally, let us take up BT (D) violation as in (174).

(174) LF:\ He saw John.


He in (174) is identified with, but not c-commanded by, John. He therefore is not dependent upon John, and undergoes covert Copy, as illustrated in (175).

(175) LF₁: John saw John.

At first sight, this representation appears to be totally unrelated to the "i-within-i" violations or "referential circularity" examined above. It, however, can be also regarded as involving IRS (whose definition is repeated below as (176)) when we follow Higginbotham (1983) and assume that the notion 'contains' stated in its definition can be reflexive as in (177).

(176) Infinitely Regressive Saturation (IRS): (= (159))

If β is dependent upon α, and α contains β, the saturation based upon the identification of α and β is infinitely regressive.

(177) The relation "contains" is understood as reflexive: α always contains β if α = β.53

First, let us label he and John in (174) as in (178) below for ease of explanation.

(178) LF₂: [α He] saw [β John]

Note that John as β here is dependent upon he as α although the opposite is not true. The application of Copy as in (175) then has the effect of turning α (he) into β (John) as illustrated in (179).

(179) LF₃: [α=β John] saw [β John]

Since α now is β, α does contain β in accordance with (177). α (he) thus comes to contain β (the "duplicate" John) while β (the "original" John) is still dependent upon α. This gives rise to IRS in accordance with (176). Note that (176) defines IRS whether the item to be saturated is α or β. The crucial aspect of BT (D) violations, in other words, is that if a name is being anaphoric to a pronoun via dependency, the anaphoricity inevitably circulates back to the name itself because of the containment arising from the application of covert Copy at
LF. Note that the reflexive containment in (179) does not give rise to \( \beta \)'s dependency upon \( \alpha \) since the condition "where \( \alpha \neq \beta \)" in the definition of c-command (160) above guarantees that he as \( \alpha \) does not c-command itself (= the "duplicate" John as \( \beta \) in (179)). Rather, it is the "original" John as \( \beta \) which establishes dependency upon \( \alpha \).\textsuperscript{54}

The involvement of dependency established in this fashion plays a crucial role in distinguishing BT (D) violation from successful coreference achieved by covert Copy as illustrated in the derivation (180a-b).

(180) a. LF: \([ [_{\beta}\text{John's} ] \text{ mother } \text{ saw } [_{\alpha}\text{him } ] \]

b. LF: \([ [_{\beta}\text{John's} ] \text{ mother } \text{ saw } [_{\alpha=\beta}\text{John } ] \]

In (180b), while the self-containment holds between \( \alpha \) and \( \beta \), it does not achieve dependency, as has just been discussed. Moreover, unlike in the case of BT (D) violation in (174), the "original" John as \( \beta \) is not c-commanded by, and hence is not dependent upon, \( \alpha \). (180b) therefore does not induce any IRS, and the intended coreference is permitted. The application of cover Attract as illustrated in (181) below does not give rise to IRS, either.

(181) a. LF: \([ [_{\alpha}\text{There } ] \text{ arose a storm } [_{\beta}\text{in the South } ] \]

b. LF: \([ [_{\alpha=\beta}\text{in the South } ] \text{ arose a storm } t \]

Again, the self-containment holds between \( \alpha \) and \( \beta \) in (181b), but does not achieve dependency. Unlike in the cases of Copy, moreover, no referential content of the "original" \( \beta \) is left behind since Attract transfers rather than duplicates them onto \( \alpha \).\textsuperscript{55} Thus, no dependency is achieved between the "original" \( \beta \) and \( \alpha \), and no IRS arises. The legitimate cases of covert Copy in (180) and Attract in (181), in other words, can be correctly permitted.
To sum up, the Affect Approach incorporating covert Copy permits us to unify all of "i-within-i" violations, "referential circularity" and BT (D) violations in terms of IRS holding at LF. This, I believe, is a commendable result since all of these phenomena intuitively share one and the same problem — the failure of saturation when the referential dependency of an item circulates back to that item itself. The Copy analysis permits us to reduce this problem to IRS, inducing simultaneous containment and dependency in all these phenomena alike.\(^56\)

4. Summary and Conclusions

We proposed and argued for the Affect Approach to anaphoric interpretation, in which referentially underspecified and hence "unsaturated" proforms undergo one or more of the options of Affect — Identify, Copy and Attract — at LF to satisfy the Principle of Full Interpretation. We provided empirical and theoretical motivations for the application of each such option of Affect for the syntactic preparation of the proper anaphoric interpretation in the CI-interface, i.e., semantic interpretation. The arguments provided for the support of this approach led us to the theoretical conclusions that the driving force of "Affect" is not limited to the need for the checking of "formal features", and that postulation of a special class of entities called expletives is not necessary.

Notes

*Acknowledgment here.

\(^1\) Throughout this work, we will use the term "deictic" in this general sense.
A similar view can be found in Reinhart and Reuland (1993), in which anaphors are regarded as referentially defective and binding is considered as the procedures assigning the content necessary for their referential interpretation.

Higginbotham in turn ascribes the original observation to Nancy Brown. Throughout this work, labelling with uppercase letters indicates that the example sentences make up a discourse.

Following a growing trend, I will refer to the following generalization offered by Lasnik (1989) as the condition D of the binding theory:

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

In Section 3.3 below, we will attempt to reduce BT (D) to FI.

Fiengo and May's (1994) "σ-sequence" more or less corresponds to the layer of guises in (11). Following Fiengo and May, we will assume that the domain of the layer of guises is a discourse rather than a clause or a single utterance.

One possibility is that G-link is established along with the association of nominal expressions and other types of grammatical information when Numeration is formed. Alternatively, we may consider that it is established in the lexicon as part of the lexical information of R-expressions.

The following example brought to my attention by Juan Uriagereka (p.c.), however, may suggest that (12a) is too strong:

(i) George Washington's mother [vp loved George dearly ], and George Bush's mother did [vp e ], too.
Note that the interpretation of the name *George* in the VP reconstructed at the ellipsis site suggests that this R-expression is not G-linked independently, or alternatively, it is not interpreted as an R-expression. In what follows, I will concentrate on the apparently unmarked use of R-expressions, and leave this interesting phenomenon unexplored.

8 Our definition of referential dependency is distinct from Evans's (1980) "referential dependency" in several crucial respects. For example, while Evans defines it in terms of its semantic consequence, referring to "one term's picking up its reference from another," we regard it as a syntactic state that necessarily induces such a semantic consequence, which may or may not lead to referential saturation. Evans also permits referential dependency to hold even without a c-command relation between two terms, while we do not.

9 If one opts for Chomsky's copy theory of movement, the options of Affect will be Identify, Copy and Delete. Our "Attract" in (21c) then corresponds to Copy followed by Delete, while our "Copy" in (21b) corresponds to Copy without subsequent application of Delete.

10 Recall here that identification already requires the two items to agree in relevant grammatical features.

11 One possibility pointed out to me by Chris Tancredi (p.c.) is that Copy actually duplicates the G-link as part of the content of the antecedent, and the two guises are necessarily co-linked with an identical referent, as illustrated in (i):

(i) Copy for Saturation:
Note that the one-to-many R-linking as in (i) is independently needed for the case of multiple guises as in (11) above, although (i) is crucially distinct from (11) in that the two guises (G₁) are identical.

12 Or by (13c) if the copy theory of movement is adopted.

13 One of the anonymous reviewers presents the derivation in (i) below and claims that the ungrammatical sentence in (ia) would be incorrectly permitted in the Affect Approach:

(i)  a. *[ the fact that he seems that Mary was sick ] didn't bother John
     ____________________________

     b. LF: [ the fact that John seems that Mary was sick ] didn't bother t
          ____________________________

He (or she) claims that this derivation would be incorrectly permitted since Attract involves only one thematic position, the landing site of the attracted John in (ib) being a non-thematic position. Note, however, that this derivation leaves a trace that is not properly bound, which can account for the ungrammaticality of (ia).

14 See Rizzi (1999) for the claim that the basic syntactic relations — chains, binding, ellipsis and head-XP relation — can be exhaustively classified in terms of obligatory involvement of identity, c-command relation and locality (i.e., relativized minimality). If Reinhart and Reuland (1993) are correct in claiming that Binding Conditions A and B are essentially the constraints imposed on the relation between a predicate and its arguments, i.e., on the head-
XP relation, we find all these basic syntactic relations established also for anaphoric interpretation at LF.

To be more precise, the definition of referential saturation in (13a-c) should be further supplemented by (i) below:

(i) \( \beta \) is dependent upon, or is a copy of, an item referentially dependent upon a G-linked item.

This extension captures the referential saturation of himself in (iia) below and that of each other in the reconstructed VP at the ellipsis site of (iib):

(ii)  a. John is not aware that he is allowed to vote for himself.
     b. John and Mary [\( VP \) liked each other's papers], and I did [\( VP \) e], too.

In what follows, however, we will concentrate on the simpler cases of saturation that (13a-c) suffice to capture. See Kitagawa (1991a) and Kitagawa (2000) for detailed examination of cases like (iib).

Chomsky (1995, p. 208), in fact, postulates what he labels as Cliticization\(_{LF}\), which is a covert syntactic process motivated by anaphoric interpretation despite his general stance. See also Lebeaux (1983) and Chomsky (1986) for the details of this covert computational process.

According to Freeze (1992), expletives in English are exceptional in this respect.

A similar observation actually led Chomsky (1991) to propose an analysis in which the associate DP is adjoined to there at LF to derive there as what he calls an "LF affix" as in (i).

(i)  LF: [ there — many pictures] aren't \( t \) on the wall.

Under this analysis, the moved associate DP does not c-command anything in the sentence, and the lack of the scope reading in (39b) is predicted. As den Dikken (1995) points out,
however, this account gives rise to many problems. For instance, the moved associate would not c-command its own trace, failing to create a proper chain.

19 Chomsky (1995, pp. 272-275), somewhat contradictorily, it seems, assumes that FF (associate) has the ability to serve as a controller or binder, and ascribes the BT (A) violation observed in (41b) to the failure of anaphor binding arising within the functional complex to which FF (associate) has been adjoined. He says, "Then the head of the matrix clause of (44) [= our (41b)] at LF, would have the structure (45a) or (45b) [= our (i) and (ii) below], depending on how covert operations are ordered..."

(i) \[ _I \text{each other } [\text{FF } [\text{V-I } ] ] \] \quad (\text{FF-attraction} < \text{Cliticization}_{LF})

(ii) \[ _I \text{FF } [\text{each other } [\text{V-I } ] ] \] \quad (\text{Cliticization}_{LF} < \text{FF-attraction})

"On reasonable assumptions, neither of these structures qualifies as a legitimate binding-theoretic configuration, with An[aphor] taking FF (linguist) as its antecedent". (pp. 275-6)

20 There also are some speakers who do not permit the anaphor binding in (42a) and antecedent contained deletion in (42b).

21 I owe much to the comments made by the anonymous reviewers for sharpening the arguments presented in this section.

22 We will assume that when a there construction does not contain an overt locative, a phonetically empty locative expression is involved as illustrated in (i).

(i) LF\(_i\): **There** arrived a spy \([\text{Loc e }]\).
The insight behind the proposed analysis can be already observed in Kuno's (1973) "Locative Postposing" transformation, which "will move locatives to the sentence-final position, leaving a trace in the form of there in their original position (p. 370)," as illustrated in (ii).

(ii) a. D: On the table are two books.
    b. S: There are two books [on the table].

See also Moro (1991), in which there is analyzed as a locative predicate raised out of a small clause, as in (iii):

(iii) [IP [IP There are two books [on the table].]

I am grateful to Lisa Travis for bringing this phenomenon to my attention. See Chomsky (1981) for a similar observation with respect to the expletive it.

I owe the example in (55c) to one of the anonymous reviewers.

The reviewer who provided the example (55c) also finds the bound variable interpretation of its in (i) below possible, though there seems to exist variations among speakers as to this intuition.

(i) [A book about every city] seems to its residents to be in the Library of Congress.

See Kitahara (1994) and Takano (1996) for the idea that the NP complement of a PP is covertly raised to the specifier position of the functional category accompanying this PP, and the Spec-head agreement permits the entire functional phrase to project the properties of the raised NP complement.
27 The example in (64a) was drawn from Dunlap (1985), who discusses binding both into and out of PPs. See also Reinhart and Reuland (1993) and Epstein, et al. (1998) for relevant discussion.

28 All pictures as an associate DP in the same sentence presumably exhibits the same scope restriction, but is ruled out due to the definiteness effect.

29 It can also take scope lower than negation, which we will not discuss. See, however, footnote 30 below.

30 One of the anonymous reviewers considers that we incorrectly predict the absence of narrow scope of many / all in a sentence like (ia) below in the Attract Locative analysis, in which the locative PP is attracted to the position higher than not at LF, as in (ib).

    (i)  a. There do not seem t to remain Power Ranger dolls [ in { many / all } of our stores ].
        b. LF: [1 in { many / all} of our stores ] do not seem t₁ to remain Power Ranger dolls t₁

This, however, is not necessarily a prediction we make. Suppose that we adopt the view that an interpretive operation may apply at more than a single stage of derivation (Lebeaux (1988), Lebeaux (1991) and Epstein, et al. (1998)), and extend it to covert syntax, as has been done in Kitagawa (1986). In this "multiple level" interpretation approach, the seemingly contradictory binding facts in (i) below is explained by encoding the violation of BT (C/D) (he-2... Bill₁) before wh-movement applies, and by encoding the satisfaction of BT (A) (Mary₁ ... herselfₜ₁) after wh-movement applies.

    (i) Mary₁ wondered [ which claim that pictures of herself₁ disturbed Bill₂ ] he₂ made  t
        (Brody (1995, 134))
Following the spirits of Chomsky (1993) and Fox (1995b), we can also reinterpret May's (1985) and Aoun and Li's (1989) analyses of quantifier lowering. In particular, the scope ambiguity in a sentence like (ii) below is ascribed to the two distinct positions the quantified subject occupies at different points of derivation, as illustrated in (iii).

(ii) **Most nebulas**, seem [ t₁ to gyrate ].  
(SEEM > MOST / MOST > SEEM)

(Mats Rooth (p.c.))

(iii) a. ____ seem [ **most nebulas** to gyrate ].

b. **most nebulas**, seem [ t₁ to gyrate ].

Similarly in the derivation of (ia-b), the quantified locative occupies the subordinate position before it undergoes Attract Locative, which may be regarded as the source of the lower scope reading.

31 Lasnik offers the following three arguments against the QR account (pp. 54-55), each of which can be shown to have some flaw. First, Lasnik points out that, if QR is responsible for the wide scope of *everyone* in (ia) below, it should also resolve ACD in (ib), contrary to the fact.

(i) a. **Who**, do you think [ιp **everyone** saw t₁ at the rally ]?  
(∀ > WH)

b.*John believed (that) [ιp [ΙP **everyone** you did [vP e ] ] was a genius ].

The impossibility of the wide scope of *everyone* in (ii) below, however, suggests that a similar scope reading in (ia) is not made possible by QR to begin with.

(ii) **Who**, t₁ thinks [ιp **everyone** saw her at the rally ]?  
(*∀ > WH)

Second, if QR is to resolve ACD, it must raise the whole DP containing the modifier of the quantified nominal. Lasnik argues that such an analysis would incorrectly leave the contrast in (iii) below concerning the BT (D) violation.
(iii) a. *He\(_1\) liked [\(NP\) every book that \(John\_1\) read ].
    b. [\(NP\) Which book that \(John\_1\) read ]\(_1\) did he\(_1\) like t\(_2\)?

As discussed by Brody (1995, 134), Lebeaux (1991), and Epstein, et al. (1998), however, BT (D) violation is induced by overt Wh-movement as long as the R-expression is not introduced in an adjunct, as illustrated in (iv).

(iv) Mary wondered [ which claim that pictures of herself disturbed Bill\(_2\] he\(_2\) made t
\[_______________________________]\n
Note that in Lasnik's example (iiiib), the name John is introduced in an adjunct clause.

Finally, Lasnik cites Takahashi's (1996) example in (v) below, which allegedly demonstrates that QR fails to resolve ACD, disallowing a sloppy identity interpretation of the reflexive proform zibun 'self'.

(v) John-ga/mo \([VP\) zibun-no hahaoya-ni \([DP\) Mary-ga \([VP\ e p ro_1 t_2]\) okutta\(_2\)
    -NOM/ALSO self's mother-DAT -NOM sent

    \(\text{dono-hon}_1\)-mo t\(_3\) okutta\(_4\)
    whichever.book sent

    'John (also) sent to his mother all the books that Mary sent.'

This example, however, does not seem to either argue for or argue against the QR account of ACD. First, although the easiest and the strongest reading here involves strict identity, this example can be felicitously uttered and interpreted even when John sent to his mother all the books that Mary sent to her mother, perhaps due to what Hoji (1998) calls "sloppy-like" interpretation. This reading becomes easier when we add _mo 'also' to the dative phrase containing zibun, as in (vi) below.
(vi) John-ga zibun-no hahoya-ni-mo [DP Mary-ga e pro₁ okutta
-NOM -ALSO
dono-hon₁-mo ] okutta (koto)

A similar reading seems to be straightforwardly available also in (vii).

(vii) John-ga [ Mary-ga e okutta dono-kaisya]-ni-mo [ zibun-no rirekisyo ]-o
-NOM -NOM sent to.whichever.company self's CV-ACC

okuranakatta wake
did.not.send reason

'the reason why John sent his CV to none of the companies to which Mary sent her CV'

This kind of interpretation, in fact, is possible even when we eliminate the quantificational
element from (v), as in (viii) an (ix).

(viii) John-ga zibun-no hahoya-ni-mo [DP Mary-ga e pro₁ okutta
-NOM -ALSO
hon₁-o ] okutta₃ (koto)
book-ACC

(ix) Mary-ga zibun-no booihurendo-ni-mo [DP Susan-ga e pro₁ okutta
-NOM self's boyfriend-DAT-ALSO -NOM sent

itimanen-no tyokoreeto₁-o] okutta₃ (koto)
¥10,000-GEN chocolate-ACC sent (fact)

'(the fact that) Mary sent to her boyfriend the ¥10,000 chocolate, which Susan sent.'

Finally, Otani and Whitman's (1991) "Disguised VP-Ellipsis" analysis of (v), which
Takahashi adopts, is full of problems, as pointed out by Hoji (1998) and Kitagawa (1999).

I would like to leave open, however, whether the covert raising of there should be regarded
as ACC-raising or not. One of the anonymous reviewers in fact considers that it should not,
since the locative does not require Case-checking. (Note, however, that there may require
Case-checking.) Lasnik (1999, 160, 167) claims that the object shift in general is triggered
by an EPP feature that resides in Agr_o. Once we are freed from the assumption that Case-
checking is the only possible trigger for object shift, we can also imagine some other possible
trigger, for instance, Φ-features of Agr_o which require object agreement in covert syntax.

33 As one of the reviewers predicts, my informants find it difficult to get the extraposed clause
(CP_1) in (i) below to be contained in the deletion part of the sentence. That is, (i) does not
mean something like "Watson always expects there to be something hidden that will surprise
all of us whenever Holmes expects there to be something hidden that will surprise all of us."

(i) Watson always [_{vp} expects [_{ip} there to be something hidden t, whenever Holmes
does [_{vp e }] ] [_{cp1} that will surprise all of us ]]].

I believe that we can regard this restriction on the interpretation of the extraposed clause as
the indication that ACD cannot be resolved in this construction.

34 While my informants generally did not have any trouble, some speakers may find it
difficult to extract any experiencer PP as in (i) below.

(i) a. To which congressman, does the possibility of meltdown seem t, to exist?
   b. To whom, does John seem t, to be clever?

For those speakers, therefore, superiority effect may not be the sole cause of the
ungrammaticality of (108b).

35 See also Epstein, et al. (1998) for relevant discussion on this issue. Analyzing overt
locative inversion in languages like Hindi, Freeze (1992, 562) claims that Infl agrees directly
with the post-verbal DP when a locative P', whose head selects the post-verbal DP as its PP-
internal subject, moves to Spec-IP and turns the PP node into a non-barrier.
Chomsky (1986, pp. 131 ff., 143, 179) proposes LF-replacement of expletive *there* with its DP associate, but seems to abstain from extending this analysis to expletive *it* and its associate CP.

Hoekstra (1984, p. 295, footnote 65) pointed out with the example (ia) below that the expletive *it* can in fact antecede an anaphor, and hence should be regarded as having a referential function.

(i)  a. *It* suggested *itself* [CP] that this solution should be rejected].
    b. [CP That this solution should be rejected] suggested *itself*.

Although there seems to exist some discrepancy among different speakers as to the acceptability of (ia), those speakers who accept (ia) can interpret it synonymously with (ib).

Within the GB-framework, Safir (1985, pp. 65-66) claims that the expletive *it* and the associate CP cannot form a θ-chain since it would yield BT (C) violation. In a sense, the Attract Clause Analysis is the minimalist solution to this problem. See 3.3 below for relevant discussion.

At this point, it is not clear to me how exactly the copied split antecedents should be represented at LF. I tentatively represent the two duplicates on top of each other.

There is one fact, however, that remains unaccounted for in the Attract Clause Analysis — even the speakers who accept (137) with *is* do not seem to permit this sentence to be followed by (i) below, in which *it* is intended to refer to the entire coordinate CP:

(i)  ..., but *it* certainly cannot happen.

This must remain to be an unsolved problem at this moment.
As has been already noted in Kitagawa (1995), one non-trivial question that arises in this approach is why the temporal pro-form *then* may not function as an expletive. A clue to this query may be found, however, in Murasugi’s (1996) work. To make the long story short, Murasugi argues that *wh*-in-*situ* in (i) below is ruled out since *when* is forced to be interpreted as the argument of Tense (or event predicate) present in the argument-taking NP, and this would induce the violation of the θ-Criterion (more generally, that of FI).

(i) *Who witnessed *[NP the falling of the big trees *when* ]?*

Crucially underlying this argument is Miyamoto’s (1992) claim that Larson’s (1990) empty temporal operator, which captures the scope ambiguity of tense observed in (ii) below, be regarded as an obligatorily selected temporal argument of Tense, and is to be distinguished from overt temporal expressions like *when*.

(ii) before [ John said [ Mary would arrive ] ] (Larson (1990))

a. Higher scope: before the time of John’s saying that Mary will arrive

b. Lower scope: before the time of Mary’s arrival as predicted by John

She then contends that *wh*-in-*situ* is permitted in (iii) below with *where* because Tense can select *where* as its NP-argument independently of its empty temporal argument without causing the violation of FI.

(iii) a. Who witnessed (with a remote camera) *[NP the falling of the big trees *where* ]?

b. Who criticized *[NP the shooting of rabbits *where* ]?

Extending Murasugi’s generalization, we may assume that *then* would necessarily yield a conflict with the empty temporal operator when it is introduced under IP (or TP), while there can be introduced without causing such a conflict.
42 Fox (1995a), Kennedy (1997) and Epstein, et al. (1998, 176) also reach a similar conclusion.


44 The same problem arises even if we (alter our assumption and) let the trace of the quantified element every boy G-linked at LF QR applies.

45 See Kitagawa (1991b) for the arguments that sloppy identity in VP-Ellipsis in fact is not derived by λ-abstraction. Hoji (1998) proposes what he calls the "Supplied N Head" analysis of the null object construction. Since the analysis of the null object construction is not the main goal of this work, I will not attempt full comparison of our NP-Copy Approach with the "Disguised VP-Ellipsis" Approach and the "Supplied N Head" Approach in this work.

46 We use coindexation in (153) (and also in some other examples below) simply for ease of presentation. A crucial assumption adopted for the derivation of strict identity in (153) is that after zibun 'self' is legitimately saturated under dependency in the antecedent clause, it may be copied into the second clause, maintaining its saturated status.

47 See Fiengo and May (1994, p. 60) for a different opinion on this matter, although their objection is directed against the appeal to the notion containment in the definition of antecedenthood.

48 Presenting "acceptable and interpretable" examples as in (i) below, one of the anonymous reviewers casts doubt on the claim that "i-within-i" violation is an empirically veritable phenomenon.

(i) a. John, is [, his, own worst enemy ]
b. [, John's story about **itself**, ]

These examples, however, may not actually involve an "i-within-i" violation (or IRS more precisely). First, the use of "equative" copula in (ia) makes it easy for this sentence to be interpreted with multiple guises. This sentence thus can be analyzed as in (ii) below, which does not involve any IRS.

(ii)

```
R₁
  / \  
G₁   G₂
     John is [ **his** own worst enemy ]
```

When we deprive the sentence of the possibility of multiple guises, as in (iii), on the other hand, "i-within-i" violation can never be avoided.

(iii) *John attacked [, his, enemy ].

The dependency involving two referential expressions as in (iv) below, on the other hand, is prohibited by the Last Resort Principle since referential identification can be established only by the application of the computational process "Identity".

(iv) [ **John's** friend ]

The coreference between **John** and **John's friend** in this construction is permitted only when multiple guises as in (ii) is permitted, for instance, as in (v).
The example in (ib) also seems to involve some extra factors. In order to accept this sentence, we must first imagine a somewhat special situation in which John wrote a story, which was about the story he was writing then. (ib) also contrasts with a typical case of "i-within-i" violation as in (vi).

(vi) *[i, John's story of itself, ]
The two sentences are only minimally distinct. In (vi), of introduces itself as the complement of story, while about in (ib) establishes a relation between story and itself. In a sense, then, story and itself in (ib) are equated in a way similar to the two items equated by the copula be, as pointed out to me by Chris Tancredi (p.c.). As a result, the two guises linked to story and itself, respectively, can now be linked to a single referent (or alternatively, itself does not involve reference any longer, being predicative). This way, we can capture the contrast between (ib) and (vi). Simply rejecting "i-within-i" as an empirical phenomenon, on the other hand, would disallow us to capture this contrast, and leave the referential circularity in (iii) and many other cases unaccounted for.

49 Higginbotham (1983) proposes the condition in (i) below together with the definitions of the notions "antecedent" and "dependence" as in (ii) and (iii).

(i) Not: D* (X, X); [where D* denotes the relation of dependence.]
(ii) X is dependent on Y if:
a. Y is contained in an antecedent of X or
b. for some Z, X is dependent on Z, and Z is dependent on Y, where
c. the relation "is contained in" is understood as reflexive: Y is always contained in Z if Y = Z.

(iii) Y is an antecedent of X if X is linked to Y or, for some Z, X is linked to Z and Y is an antecedent of Z.

50 I am grateful to Steven Franks for bringing this fact to my attention. Fiengo and May (1994, pp. 58-60, 74-75) also discuss this phenomenon with the assumption that coindexation for anaphoric dependency has access to the derivational history the traditional phrase structure rules provide, and say "what underlies the problem with circular reference sentences is not a problem with establishing referents for the pronouns". Instead, they regard it as a problem of the unrealizability of embedded dependencies in the derivational history. The plausibility of their approach depends upon the plausibility of maintaining the traditional phrase structure rules, which apply in a top-to-bottom fashion, since there does not exist any way to distinguish the alleged unrealizable dependency (= referential circularity) in (i) below and realizable dependency in (ii) if phrase structures are assumed to be generated in a bottom-up fashion, for example, with the application of Merge (Chomsky (1994)).

(i) His\_1^\beta wife\_2^\alpha loves her\_2^\beta husband\_1^\alpha.
(ii) John\_1^\alpha told his\_1^\beta wife\_2^\alpha that she\_2^\beta is beautiful.

51 Recall that referential identification arises only by way of the application of a computational process "Identify" in the Affect Approach.

52 I am grateful to Chris Tancredi for providing me with this argument.

53 See Higginbotham's (1983) definition of the notion "dependent" (ii) in Footnote 49 above.

While we have chosen the notion "containment" in defining IRS, we could replace it with the
notion "dominance" as well. The assumption that a node dominates itself is quite familiar to generative syntacticians.

54 The IRS observed in (179) is identical to that involved in (166) except that the items saturated under containment are reversed.

55 If one adopts the copy analysis of movement, Attract is followed by the obligatory deletion of the original β while Copy is not.

56 As one of the anonymous reviewers correctly points out, there should be no reason why Copy cannot be applied to a proform that takes a quantificational element as its antecedent, as in (ia-b).

(i)  
   a. Many people, were unhappy, and they, left.
   b. Someone, called. He, didn't leave a message.

The resulting LF, however, certainly is not synonymous with that derived from the surface representations as in (iia-b), in which two distinct sets of individuals are defined by the two instances of quantificational expressions in each sentence.

(ii) 
   a. Many people, were unhappy, and many people, left.
   b. Someone, called. Someone, didn't leave a message.

Kitagawa (2000) proposes to solve this problem making an appeal to a type of economy condition while motivating covert copy of bound variables. In the same work, it is argued that the proposed Copy analysis enables us to capture in a uniform fashion a type of "strict identity" interpretation available in all of VP-Ellipsis, DP-Ellipsis and E-type anaphora in (iii)-(v).
(iii) [ A statement made by the principal of a boys' school ]

In our school, every student [vp respects his teacher], and the parents also expect me to [vp e].

(iv) siritu-daigaku-no dono kyoozyu-ga [dp zibun-no gakusee]-o suisensite-mo, which professor-NOM [dp self-GEN student]-ACC recommend-ever

Monbusyoo-wa [dp e] saiyoosi-nai-daroo.
Ministry.of.Education-TOP [dp e] employ-not-perhaps

'No matter which professor of a private college may recommend self's (= his or her own) student, the Ministry of Education will probably not employ them.'

(v) Few congressmen admire Kennedy(, and) They are very junior.