PERCEPTUAL STRATEGY AND ANAPHORA

MASAHIRO KATO

0. Introduction

Up to this time three principal approaches have been proposed to untangle the mysteries of the cataphoric use of pronouns. One is the transformational approach put forward mainly by Ross (1967a), Postal (1969) and Langacker (1969). Another is the interpretive approach advocated by Jackendoff (1972), Wasow (1972), Culicover (1976), Lasnik (1976) and Reinhart (1976). The third is the functional approach proposed by Kuno (1975a).

Consider examples (1) and (2):

(1) *He choked when John swallowed the bone.
(2) When he swallowed the bone, John choked.

Ross accounts for the difference of acceptability between (1) and (2) by the following condition:

(3) Condition on backward pronominalization

If one element precedes another, the second can only pronominalize the first if the first is dominated by a subordinate clause which does not dominate the second. Ross (1967: 195)

This condition can correctly predict the acceptability of the sentences in the following paradigm:

(4) a. Jim will go if he feels good.
    b. *He will go if Jim feels good.
    c. If Jim feels good, he will go.
    d. If he feels good, Jim will go.

So far, condition (3) seems to adequately reflect the linguistic facts. That is, we may say, in the structure \([S [\text{sub. cl.} \ldots \text{pron} \ldots] \ldots \text{ant} \ldots]\) pronouns are, in general, permitted to appear to the left of antecedents.
This paper deals, primarily, with the notion that all pronouns are, fundamentally, used anaphorically, and can be accounted for if we incorporate into our treatment the Perceptual Strategy (PS) which has been proposed in Townsend and Bever (1977). I will propose the Principle of Anaphora and Anaphora Restriction I which incorporate PS and will attempt to show that these interpretive rules can account for the cases which neither the restriction put forward by Lasnik (1976) nor that proposed by Reinhart (1976) can explain.

1. Perceptual strategy and its psychological reality

A major question in the study of language is how much of linguistic data should be accounted for by purely linguistic theory. In this sense the standard theory (Chomsky (1965)) appears to lack formal explanation for a number of problems, such as how to account for the varying degrees of acceptability of sentences. As a result, many linguists, including Chomsky himself, have proposed various modifications of the standard theory. Chomsky has proposed that grammar is divided into Sentence Grammar and Semantic rules interacting with other cognitive structures, giving fuller representations of meaning. (Chomsky (1975))

Williams (1977) has also suggested that grammar is composed of two distinct subgrammars, Sentence Grammar and Discourse Grammar, each governed by its own laws. It is well known that Chomsky assumes linguistics to be one part of cognitive psychology and expects the results of psychological research to contribute to resolving the problems of language:

... there is in principle an interplay between the study of the structure of language (that part of psychology called "linguistics") and experimental psycholinguistics, which is largely concerned with models of perception and production. I am personally interested in the possibility of testing linguistic hypotheses. Certain questions cannot be resolved by sole reliance on the customary methods of linguistics. For example: the study of temporal processes, constraints on memory, the interactions between cognitive systems. Furthermore, the abstract study of grammar, and the kinds of data utilized by linguists,
are simply insufficient to resolve certain questions concerning language . . . And if the linguist is interested in the real nature of human beings—which is what I suppose—then he will seek to discover the system that is really utilized. The data of linguistics are not rich enough to answer these fascinating questions beyond a certain point. Therefore, the linguist must hope for further insight from the study of process models and neurological structures. (Chomsky (1979 : 47-8))

In spite of the remarks cited above, few linguists have ever tried to pay serious attention to relevant observations offered by psycholinguists. Thus, the integration of the study of language and that of cognitive systems, which could be of considerable significance for linguistic researcher has largely been ignored until recent years. This is one of the motivations that drives me to inquire into perceptual strategies and their influence on the surface forms of English sentences.

Listeners apply various strategies while hearing a clause in order to determine the functional relations between the words of the clause. The most prominent of the strategies is based on the assumption that the sequence of elements in the clause corresponds to the order of elements in underlying structure. It is generally accepted that when the listener hears a clause, he applies functional labelling operations, and by the time the next clause begins, he has computed an interpretation of the previous clause, stored its meaning, and lost from the memory specific details about its form. This loss of details about the previous clause, then, make immediate memory free for application of comprehension operations to the following clause. (Townsend and Bever (1977)) The evidence commonly cited for the strategy which erases from the memory the form of the preceding clause once the next one begins the probe latency experiment in Caplan (1972). According to Townsend and Bever (1977), however, there are semantic reasons for thinking that Caplan’s clause by clause comprehension model may be incorrect in some details.

They claim that clauses often differ in the types of information they convey. For example, they suggest that the main clause of a sentence gen-
erally contains the assertion which the speaker wishes to convey, while the subordinate clause conveys information to which the speaker attaches less importance. They also note that the main clause and the subordinate clause differ in that the former conveys new, important information, (analogous to visual figures), while the latter usually conveys old, less important information which proves a context or setting for new information, (like visual grounds). Since Townsend and Bever suspect that the semantic difference between main and subordinate clauses is reflected in the processing the two types of clauses, they have devised three experiments designed to reveal that the structural differences among sentences correspond to the different perceptual strategies.

In the structure [main cl.—subordinate cl.], the hearer computes an interpretation of the main clause, then, at the end of the main clause, stores its meaning, erases its verbatim form, which has until that point been held in immediate memory. As a result, active memory is freed for analysis of the second clause. On the other hand, in the structure [sub. cl.—main cl.], the initial subordinate clause is interpreted in relation to the following main clause. Thus, the hearer retains the initial subordinate clause in a relatively verbatim form as he analyzes the following main clause. This retention of the unanalyzed verbatim form of the initial subordinate clause should therefore produce a great load on active memory. Based on the observations above, Townsend and Bever (1977) propose the following perceptual strategy:

(5) Perceptual Strategy (PS)
Subordinate clause information is retained relatively longer in unanalyzed form in immediate memory. (Townsend and Bever (1977 : 18))

They refer to Ross’s Penthouse Principle as the reflection of PS (5) in the real linguistic world and as the supporting evidence for its psychological reality:

(6) The Penthouse Principle
All syntactic processes that apply in prenominal clauses also apply in other subordinate clauses. All syntactic processes that
apply in subordinate clauses also apply in main clauses. In other words, where ‘\(\geq\)' means 'is as syntactically active as, or more syntactically active than', the following inequality obtains: main clauses \(\geq\) subordinate clauses \(\geq\) prenominal clauses (Ross (1973 : 411))

This principle explains the acceptability of the following sentences:

(7) a. The alarm went off.
   b. Off went the alarm.
   c. *I know that off went the alarm.

(8) a. We will hire a new person tomorrow.
   b. Tomorrow we will hire a new person.
   c. *That tomorrow we will hire a new person is unlikely.

The Penthouse Principle (6) means, roughly, word order in prenominal clauses tends to be more narrowly constrained than in subordinate clauses, and word order in subordinate clauses, in turn, tends to be more narrowly constrained rather than in main clauses. The fact that movement rules are more constrained in subordinate clauses (e. g., (7c), (8c)) is explained by principle (6), which lies behind the PS (5), suggested in Townsend and Bever (1977). That is, since initial subordinate clauses are retained longer in unanalyzed verbatim form in immediate memory, this retention will produce a great load on immediate memory. As a result, further application of movement rules to initial subordinate clause increases the structural complexity of a sentence, and consequently, produces much greater load on immediate memory. As the reflection of this perceptual strategy, we have the grammatical constraint.

2. On the viability of "cataphoric use" of pronouns

2.1 Non-existence of cataphora. Considering the perceptual strategy (5) introduced in the previous section, let us reexamine so called "cataphoric use" of pronouns.

Kuno (1975a) proposes the following exclusively intrasentential constraint on Pronominalization:
(9) *Intra-Sentential Pronominalization*

If Discourse pronominalization is not applicable, apply Intra-Sentential Pronominalization, which works only from left to right. (Kuno (1975a: 280))

Intra-Sentential Pronominalization in (9) roughly means that Pronominalization applies, basically, "from left to right". Reconsider the example mentioned above:

(2) When he swallowed the bone, John choked.

In Kuno (1975a), (2) should be used in a situation where John, which is the antecedent of he, has already been established as a topic in the previous discourse. In relation to cases like (2), Halliday and Hasan (1976) suggest the following fundamental ideas about coreference:

In the text it is natural for the element occurring second to depend for its interpretation on the one occurring first; hence, anaphora is the unmarked and cataphora is the marked term in the opposition. Cataphora occurs only as an EXPLICIT relation, with the first element always being one that is inherently presupposing. (Halliday and Hasan (1976: 19))

Now let us consider what kind of linguistic discovery about the common notion, "cataphoric use" of pronouns we can make, if we take PS (5) into consideration.

In (2), the surface order is [...pron...ant...]. But on the level of perceptual processing, it seems reasonable to assume that the sentence initial subordinate clause is retained in its uninterpreted verbatim form, and the following main clause is interpreted first. The initial subordinate clause is then interpreted with reference to the meaning of the main clause. Consequently, he is identified with John and we arrive at the interpretation of the sentence as a whole. In short, if we adopt PS (5), the interpretation of the antecedent of the main clause in (2) precedes the interpretation of the pronoun. In this sense, it is highly reasonable to assume that pronouns like the one in (2) are anaphoric rather than cataphoric, as has been
commonly accepted so far. In fact, considering the level of perceptual processing, we may say that coreference between two NP's is possible so far as the order [. . . antecedent . . . anaphor . . . ] is kept on that level.

2.2 Further evidence for Perceptual Strategy (5). If we take Ross's principle in (6) as overt supporting evidence for the existence of PS (5), the following constraint may be assumed as the covert evidence for it:

(10) The Predictability Requirement of Backward Pronominalization
Backward Pronominalization requires that the referent of the pronoun be 'determinable' or 'predictable' from the preceding context. (Kuno (1975a : 280))

Novelists and journalists, however, sometimes intentionally violate this requirement in order to give a dramatic effect. (Kuno (1975a : 287-8))

(11) a. At 2 or 3 or 4 a. m., somewhere along in there, on August 25, 1966, his forty-eight birthday, in fact, Leonard Bernstein woke up in the dark in a state of wild alarm.

b. Leonard Bernstein woke up in the dark in a state of wild alarm at 2 or 3 or 4 a. m., somewhere along in there, on August 25, 1966, his forty-eight birthday, in fact.

Comparing (11a) with (11b), it can readily be seen that (11a) rather than (11b) holds the reader in suspense. In this sense, (11a) has much more literary flavor than (11b). This effect can be attributed to the feelings of anxiety or expectation caused by lack of the information necessary for the identification of the referents of pronouns. These feelings will disappear at the moment proper antecedents came up. Viewed from a different perspective, the fact that the reader retains such feelings until the moment when he can successfully identify the pronoun with the proper antecedent supports the hypothesis that the clause which contains a pronoun followed by its antecedent in surface structure is retained uninterpreted and thus stored temporarily in immediate memory. This means that the existence of such feelings may reflect the psychological reality of PS (5).
According to Bolinger (1977), the sentences in (12) below are characteristically offhand:

(12) a. Hey Gus! Tell him to come in if you see Tom out there, will you?
    b. Hey Gus! If you see him out there tell Tom to come in, will you?

(13) a. Hey Gus! Tell Tom to come in if you see him out there, will you?
    b. Hey Gus! If you see Tom out there, tell him to come in, will you?

The difference in tone between the sentences in (12) and (13) suggests that the intentional violation of the constraint in (10) is the source of the offhand flavor associated with (12). Thus, in formal expressions, the structure [. . . pron . . . ant . . .] is, usually, to be avoided. (Bolinger (1977:9))

(14) a. If the general comes back, put him to death.
    b. *If he comes back, put the general to death.

(15) a. If Tommy comes back, put him to bed.
    b. If he comes back, put Tommy to bed.

It is clear from the sentences in (11)-(15) that the literary flavor or dramatic effects are caused by the intentional violation of the constraint in (9).

This seems to imply that coteference relations in sentences with the structure [. . . ant . . . anap . . .] are, in general, most easily understood, for it has less information dependent on the structure. In this sense, the surface structure [. . . ant . . . anap . . .] is the unmarked case, while the surface structure [. . . anap . . . ant . . .] is the marked case because more information adds to it which is inherently dependent on the structure. As for the latter case, it may well be said that the fact that the increase of the extra information reveals the psychological reality of PS (5).

In summary, if we take PS (5) into consideration it may be reasonable to claim that pronouns are, fundamentally, used anaphorically.
3. Principle and restriction on anaphora

3.1 Principle of Anaphora. The following principle can be induced from the considerations presented in the previous section:

(16) Principle of Anaphora (PA)

NP₁ and NP₂ are noncoreferential unless the perceptual order [··· [antecedent] NP₁] ... [anaphor NP₂] ...] is kept on the level of perceptual processing.

In what follows, I will show how the PA in (16) can handle cases which are inexplicable on the basis of the constraints proposed in Lasnik (1976) and Reinhart (1976) as well as cases which these two treatments can account for. For ease of exposition, I cite their constraints in (17) and (18):

(17) Noncoreferential constraint

If NP₁ precedes and commands NP₂ and NP₂ is not a pronoun, then NP₁ and NP₂ are noncoreferential. (Lasnik (1976 : 15))

(18) Coreference restriction

Two NP's cannot be coreferential if one is in the syntactic domain of the other and is not a pronoun. (where the domain is defined by the c-command relation) (Reinhart (1976 : 125))

Consider the following sentences:

(19) Near him, John saw a snake.

(20) In her bed, Zelda spent her sweetest hours.

(21) For his wife, Ben would give his life.

(22) *Near John, he saw a snake.

(23) *In Zelda's bed, she spent her sweetest hours.

(24) *For Ben's wife, he would give his life.

Since pronouns both precede and command full NP's in (19)-(21), (17) incorrectly blocks them. On the other hand, in (22)-(24), the full NP's both precede and command the pronouns, while (17) incorrectly marks them
as acceptable. (18) can account for all the cases above. In (19)–(21), all the NP's in the domain of the full NP's are pronouns, so (18) correctly marks them as acceptable. In (22)–(24), on the other hand, since all the NP's commanded by the pronouns are full NP's, (18) correctly marks them as unacceptable.

Now let us see whether the Principle of Anaphora in (16) can assign correct interpretations to the sentences in (19)–(24). It will be shown that apparent counter-examples to the PA in (16) are not real counterexamples to it, given the restriction proposed in the following section.

3.2 Anaphora Restriction I. It is reasonable to assume that we may have the following perceptual strategy because of the existence of the similar strategy cited above as (5):

(25) Perceptual Strategy (PS)
In processing a sentence with the surface form [. . . anaphor . . . antecedent . . . ], the constituents containing anaphors are, in general, temporarily kept uninterpreted and stored in immediate memory.

Given PS (25), sentences with the structure [. . . anaph . . . ant. . . . ] should not be permitted coreferential relation between the anaphor and the antecedent under a certain restriction. This restriction is supposed to be formalized explicitly as follows:

(26) Anaphora Restriction I (AR I)
In the structure [ S . . . [ a1 . . . NP2 . . . ] . . . NP1 . . . [ a1 e] . . . ], NP1 and NP2 are noncoreferential unless NP1 is the antecedent of NP2. (where [ a1 e] is a trace left by preposing transformations)

In the rest of this section, I will show how the Principle of Anaphora in (16) and the Anaphora Restriction I in (26) can account not only the sentences for which (17) and (18) can account but also for the sentences for which they cannot. Let us reconsider the sentences in (19) – (24). Their S-structures are as follows:
(19)' Near him, John saw a snake [pp e]

(20)' In her bed, Zelda spent her sweetest hours [pp e]

(21), For his wife, Ben would give his life [pp e]

(22), Near John, he saw a snake [pp e]

(23)' In Zelda's bed, she spent her sweetest hours [pp e]

(24)' For Ben's wife, he would give his life [pp e]

In (19)', (20)', and (21)', since the pronouns correspond to NP₂ in AR I (26) and the full NP's to NP₁ and the latter stand as the antecedents of the former, AR I (26) correctly marks them as acceptable. On the other hand, in (22)', (23)', and (24)', since the full NP's correspond to NP₂ in AR I (26) and the pronouns to NP₁ and, in these cases, the latter do not stand as the antecedents of the former, AR I correctly marks them unacceptable.

The sentences in (27) - (30) are examples that Lasnik's constraint (17) and AR I, but not Reinhart's (18), can explain. In each case the (b) structure underlies the corresponding (a) from:

(27) a. *To him, Mary spoke in Ben's office.
   b. To him, Mary spoke [pp e] in Ben's office.

(28) a. *To her, John talked about Mary's husband.
   b. To her, John talked [pp e] about Mary's husband.


(30) a. *For her picture, Ben searched in Jane's room.
   b. For her picture, Ben searched [pp e] in Jane's room.

Since the pronouns correspond to NP₂ in AR I and the full NP's to the left of the traces correspond to NP₁, and NP₁ is not the antecedent of NP₂, AR I rightly marks the sentences in (27) - (30) unacceptable.
3.3 The relation between Principle of Anaphora and Anaphora Restriction I

As for the relation between PA (16) and AR I (26), Principle of Anaphora, where the notion of "the level of perceptual processing" is explicitly specified, is a constraint assigned to a component to be called Discourse Grammar, in the extended sense defined in Williams (1977). AR I, on the other hand, is a restriction assigned to a component to be called Sentence Grammar. For evidence that AR I is assigned to Sentence Grammar and indispensable to it, consider the following sentences:

(31) a. John couldn't find Bill. He searched for him at the edge of the forest. He called Bill's name but there was no answer. Suddenly he came upon a figure sprawled on the ground. It was Bill lying there. Near him, John saw a snake.
   b. Near him, John saw a snake.

(32) a. Zelda had never spent much time in her mother's room. But now in her bed, Zelda was recovering from the fatigue of the trip.
   b. In her bed, Zelda was recovering from the fatigue of the trip.

In (31a), it is most natural to interpret the referent of him as Bill, for this possible antecedent has already appeared as a topic in the previous discourse. In other words, the force of the preceding discourse, which encourages the association of him with the established topic Bill is stronger than that created by the sentence which encourages the association him with John. Consider sentence (31b). In isolation, it is most natural to interpret the referent of him as John. In this case, it is necessary to resort to an interpretive rule in Sentence Grammar that will give an appropriate interpretation to him, i.e., ARI should be assigned to Sentence Grammar. The same reasoning holds in the case of (32). In (32), the dominant interpretation of the referent of her is her mother, while in (32b), the referent of her is Zelda. Thus, we need AR I in order to interpret (32b) correctly.
One more comment, in passing, is in order with respect to the application of AR I. It follows naturally from the structural description of the formulation in (26) that AR I applies exclusively to sentences where certain preposing rules have been applied. Keeping this in mind, consider the following examples:

(33) *I talked to him about John.
(34) *Mary always talks to him in Bill's apartment.
(35) *I spoke to him in Ben's office.
(36) *It didn't occur to her that Rosa has failed the exam.

The sentences in (33) – (36) are relevant counterexamples to Reinhart's restriction in (18) (but not to Lasnik's (17)). Since neither pronoun nor antecedent in each sentence does c-commands each other, the restriction in (18) incorrectly marks the sentences in (33) – (36) fully acceptable. The question arises; can the Principle of Anaphora in (16) or Anaphora Restriction I in (26) reasonably account for (33)–(36)? The answer is that PA (16), not AR (26), governs these cases, for AR I is not applicable because no preposing rules have been applied to them. Since all these sentences have pronouns to the left of full NP's on the level of perceptual processing, PA correctly marks them as unacceptable.

4. Summary

The discussion in this paper has shown that pronouns can be seen as being fundamentally anaphoric devices, if we accept the idea that the perceptual strategy in (5) is used in processing sentences with the structure [subordinate clause. . .main clause]. As a reflection of this “monolithic” use of pronouns in the grammar, I have proposed the Principle of Anaphora in (16) as a constraint on the level of perceptual processing and the Anaphora Restriction I in (26) as a constraint on the surface structure of sentences in which preposing rules have applied. I have also mentioned that the former is assigned to Discourse Grammar, the latter to Sentence Grammar.
I have also demonstrated that these constraints can account for not only the sentences for which the restrictions proposed in Lasnik (1976) and Reinhart (1976) can account, but also the sentences for which they cannot.

*I am especially grateful to Professors Minoru Yasui, Minoru Nakau, and Shosuke Haraguchi for invaluable comments and suggestions. All errors and inadequacies in the paper, needless to say, are my own.

Footnotes

1. Since I would like to avoid the theoretical inconsistency caused by the term “backward pronominalization” which may give the illusion that such a transformation actually is tenable, I will use the expression “cataphora” or “cataphoric use of a pronoun”, in the sense defined in Halliday and Hasan (1976:17), instead.

2. Williams (1977) claims that Grammar should be subcategorized into two components; Discourse Grammar and Sentence Grammar. His claim is mainly based on the following two criteria:

   (i) Those rules which obey the Coordinate Structure Constraint (CSC) and the Complex Noun Phrase Constraint (CNPC) are assigned to Sentence Grammar. Those rules which do not obey these constraints are assigned to Discourse Grammar.

   (ii) Those rules which do not apply across sentences in a discourse are assigned to Sentence Grammar. Those rules which apply across sentences in a discourse are assigned to Discourse Grammar.

(1) The man who didn’t leave knows the man who did.
(2) John didn’t immediately open the door --- first he shut the window, and then he did.

In (1), the missing VP is contained in a complex NP that does not contain its antecedent (leave). Thus the rule relating the VP to its antecedent does not obey the CNPC. In (2), the missing VP is contained in a coordinate structure that does not contain its antecedent; the rule relating the missing
VP to its antecedent does not obey the CSC. Hence, Williams claims, the rule usually referred to as VP Deletion differs from such rules as Gapping and Comparative Deletion, which obey these constraints.

He also illustrates that VP Deletion can operate "across utterance boundaries" while Gapping cannot:

(3) A: Did John leave?
    B: Yes, he did.
(4) A: Did Sam go to the store?
    B: *No, Bill to the supermarket.

In (3), the missing VP in B is anaphoric to the VP in A, while in (4), which involve Gapping, such anaphoric relationships are impossible. Williams presents additional examples which show that Comparative Deletion, like Gapping, requires that the antecedent of the deleted term appear in the sentence involving occur in the deletion:

(5) A: Did John see cows?
    B: *Yes, but Sam saw more horses than John saw.

As a result, Williams has proposed the following:

Those rules like Comparative Deletion and Gapping that are sentence-bound and that obey Ross's constraints we will call Sentence Grammar rules, following Chomsky (1975). These rules define the form and meaning of sentences. Those rules which, like VP Deletion, apply across sentences in a discourse and which do not obey Ross's constraints we will assign to a component to be called Discourse Grammar. These rules specify the relationship of a sentence to its linguistic context—that is, its relationship to other sentences in a discourse. (Williams (1977: 102-3))

In passing, Hankamer and Sag (1976) and Sag (1976) claim that VP Deletion cannot be pragmatically controlled; for it falls into the category "Surface Anaphora", which prohibits pragmatic control:
(6) (Hankamer attempts to stuff a 9-inch ball through a 6-inch hoop)

Sag : #It's not clear that you'll be able to φ.

(where # is used to mark a sentence which is grammatical in isolation, but unacceptable in a specific context)

Although this discrepancy between the two arguments regarding pragmatic controllability requires further investigation, it is outside the scope of the present study.

3. Williams defines the rules of Discourse Grammar as:

The rules of Discourse Grammar are rules whose relevant terms, such as deletion site and antecedent, are not in general contained within a single sentence. (Williams (1977: 192))

4. Since Kuno (1975a) does not formulate Pronominalization explicitly in his own terms, it is difficult to reach a satisfactory understanding of his ideas about the relation a pronoun and its antecedent. Considering the following statements, I think that Kuno assumes Pronominalization is a transformation and applies after all movement rules:

Pronominalization, as I view it, applies after all movement transformations have applied. (Kuno (1975a: 280))

5. Consider the following sentence:

(1) Although he doesn't know it yet, John has just won $1,000.

Kuno (1975a) refers, in fn. 13, to Kartunen's observation that (1) is perfectly acceptable without any assumption that the antecedent of he has been mentioned in the immediately preceding discourse. Then, he claims that sentences such as in (1) can be used only when the speaker assumes that the hearer can determine the referent of the pronoun from the preceding or from the nonlinguistic context. I think, however, that his claim is too strong, for (1) is perfectly acceptable even it is uttered discourse initially without
the help of nonlinguistic factors. On the other hand, Principle of Anaphora (16) can account for (1).

An observation in Lasnik (1976) also seems to argue against Kuno's proposal:

(2) I spoke to Oscar yesterday. *He finally realized that Oscar is unpopular.

Lasnik points out that even the existence of prior identifying information in the previous discourse is not sufficient to permit coreference between He and Oscar.

6. Although the term \([a]\) is used to represent a trace left by the moved constituent in the formulation of AR I (26), I would at present refrain from committing myself to Trace Theory as developed in Chomsky (1975, 1976, 1977, and 1979), and in Fiengo (1974, 1977).

It has been proposed that "...when a transformation moves a phrase P from position X to position Y it leaves in position X a trace bound by P." (Chomsky (1975: 95)), and that "...the movement of any element, not necessarily NP, leaves trace." (Fiengo (1974: 61)). We can find in Chomsky (1976: 114-5) supporting evidence for the hypothesis that elements other than NP's may possibly leave a trace in the position from which they are moved:

(1) a. John saw \([\text{NP a picture [PP of who]}]\)
    b. John saw \([\text{NP a picture t] [PP of who]}]\)
    c. who John saw \([\text{NP a picture t] [PP of t]}]\)
    d. who did John see a picture of?

(2) a. he destroyed \([\text{NP a picture of who]}\)
    b. *he destroyed \([\text{NP a picture t] [PP of who]}]\)
    c. *Who did John destroy a picture of?

Chomsky explains that the difference of acceptability between (1) and (2) is due to the fact that a lexically governed rule of extraposition from NP
can apply to the structure in (1a), but it cannot apply to the structure in (2a). This explanation depends crucially on the premise that PP's leave traces in the position from which they are moved.

Prof. Nakau has pointed out to me that the following categories must leave traces in the original position, within the framework of Trace Theory:

(3) A review t came out yesterday of this book. (Prepositional Phrase)
(4) A room t was discovered full of ancient artifacts. (Adjectival Phrase)
(5) None t can be found speaking well of him (Participal Phrase)
(6) The time t had come to decorate the house for Christmas. (Infinitival Phrase)
(7) A gun t went off which I had cleaned. (Relative Clause)
(8) Word t spread that a serpent had washed up on the beach. (Appositive Clause)
(9) The problem t arose what contribution the public should pay. (Embedded Question)

All these examples show that it is not unreasonable to assume that other constituents as well as NP's are supposed to leave traces after the application of movement rules. Unfortunately, there are few papers which deal with this assumption. Thus, for the moment, in this present paper, I assume that \[^{\alpha i} \varepsilon\] is the symbol which merely indicates the original positions of moved elements.

7. See fn. 2.

REFERENCES


(Received April 30, 1982)