

---

# Anaphora Resolution

P. A. M. SEUREN

*Filosofisch Instituut  
Nijmegen University*

It is a remarkable fact in anaphora research that sentence-internal anaphora is dealt with almost exclusively in the linguistic literature, whereas sentence-external anaphora seems hunting ground reserved for psycholinguists. This curious division of labour is no doubt due to the fact, accidental or not, that linguists feel ill at ease outside the sentence while psycholinguists, at least until recently, tended to concentrate on experiments, eschewing the abstract formal structures needed to explain internal anaphora. In this study an attempt is made to bridge the gap by proposing an integrated theory of anaphora, at least in outline. The proposal is rooted in a psychological theory of linguistic understanding where factors of context and discourse play a decisive role. An algorithm is proposed according to which pronouns occurring in sentences can be interpreted. For internal anaphora the algorithm or Assignment Procedure (AP) can be applied to sentence-types; for external anaphora it applies to sentences in context (i.e., utterance types).

AP presupposes a grammar (i.e., a parser and a generator) relating surface structures to semantic analyses (SA) via shallow structures and syntactic deep structures. The calculus of semantic interpretation takes SA as input and delivers, in a cumulative fashion, bits of cognitive discourse representation. AP is part of the semantic calculus. It applies to fully lexicalized NP's and to pronouns. The pronominal part of AP differs from the rest of the semantics in that it can draw information from all derivational stages of sentences between SA and surface structure, though it still takes SA as input.

Pronouns are treated by the parser as any other NP: only morphological markings of number, gender, and case play a role there. Their specification in terms of very general lexical categories ('male', 'thing', etc.) will be largely neglected here. In SA they are represented as pronouns (the symbol "x" will be used here); when they are morphologically marked for 'reflexive' one way or another, this is indicated in SA by a subscript R: "x<sub>R</sub>". NP's are either quantified or definite. Definite NP's are either variables (pronouns) or denoting NP's. The relation of denoting holds between denoting NP's and cognitive entities or addresses. The term "referring" is used for the relation between NP's or their corresponding addresses on the one hand and real entities on the other. Denoting NP's in false assertions or in subdomains of the cognitive discourse representation need not refer.

Three kinds of (personal or possessive) pronouns are distinguished: reflexive pronouns, bound variable pronouns, and denoting pronouns. The first and second categories are obligatorily pronominal. The third category may be represented, *salva veritate*, by a full noun phrase containing a predicate denoting either a very general property or expressing an evaluative judgement. Such NP's will be called *epithetic NP's*. They must always be unaccented. Only denoting pronouns (or epithetic NP's) can have a sentence-external antecedent.

The category of *reflexive pronouns* does not coincide with the class of pronouns with reflexive morphological marking (the latter will be called *marked reflexives*). A pronoun is reflexive just in case it is co-denotational with (or is the same variable as) the subject of either the same clause or of the higher clause to which its own clause is a complement, at some stage of analysis. If the pronoun is related to the subject of its own clause, it is a *directive reflexive*. If it is related to the subject of the higher clause it is an *indirect reflexive*. Languages differ considerably in the morphological marking of reflexives. In Attic Greek, for example, there is a special morphological set of indirect reflexives. Latin uses ordinary reflexive marking also for indirect reflexives:

- (1) a. *Sibi Asterigem mitti Caesar iussit.*  
(*Caesar ordered that Asterix be sent to him*)
- b. *Caesar Asterigi imperavit ut sibi nuntium mitteret.*  
(*Caesar ordered Asterix that he send him a messenger*)

Latin (like Swedish) also has a separate third person reflexive possessive pronoun *suus* (Swedish: *sin*), besides the non-reflexive personal pronoun genitive *eius* (*hans*). Dutch has three different forms for the third person reflexive: *zich*, *zichzelf*, and *hemzelf* (*haarzelf*, *henzelf*) for English *himself* (*herself*, *itself*, *themselves*). The conditions of use for the three Dutch marked reflexives differ from each other (and from English) in fairly complex ways. A few examples will suffice here:

- (2) a. *Ben toonde Josef een foto van zichzelf.*  
 (*Ben showed Josef a picture of himself*)  
 b. *Ben toonde Josef een foto van hemzelf.*  
 (*Ben showed Josef a picture of himself*)  
 c. *Karel liet zich het pakje bezorgen.*  
 (*Karel let the parcel be delivered to him*)  
 d. *Karel liet Ben hem het pakje bezorgen.*  
 (*Karel let Ben deliver the parcel to him*)  
 e. *Karel liet voor zich werken.*  
 (*Karel allowed (people) to work for him*)  
 f. *Karel liet Ben voor zich werken.*  
 (*Karel let Ben work for him*)  
 g. *Karel liet Ben voor zichzelf werken.*  
 (*Karel let Ben work for himself*)

The obvious differences between Dutch and English in this respect come about as a result of differences in the grammar of complementation (Dutch *laten* ("let") takes obligatory Predicate Raising (see Seuren, 1972), and of differences in reflexive morphology and the conditions of use of the morphology. Restrictions of space forbid an adequate discussion of the rules of Dutch grammar in this respect, so we must be content with the conclusion that, apparently, languages differ considerably in the morphological marking of reflexives. We will assume that reflexives are often not marked morphologically. Unmarked reflexives differ from denoting pronouns in their conditions of use, and also because they cannot be replaced by epithetic NP's *salva veritate*:

- (3) a. In *his* office *Graham* writes novels.  
 b. !In *the man's* office *Graham* writes novels.<sup>1</sup>  
 c. !In *Graham's* office *he* writes novels.  
 (4) a. That *she* was clever, *Mary* knew very well.  
 b. !That *the old girl* was clever, *Mary* knew very well.  
 c. !That *Mary* was clever, *she* knew very well.

If no account were taken of reflexivity it would be difficult to explain why co-denotation is disturbed in the (b) sentences by the epithetic NP's, or why there is no co-denotation in the (c) sentences. The latter point is remarkable (see, e.g., Akmajian & Jackendoff, 1970) since one of the outstanding features of denoting pronouns is that they have no difficulty in taking *preceding* NP's as their antecedent. In fact, however, pronouns are subject to the *Reflexive Principle*:

Whenever a *subject* is involved, at any stage of analysis, in a co-denotation (or co-variable) relation between two NP's, the subject must be the antecedent and the other NP will be reflexive,

marked if the grammar of the language requires marking, and otherwise unmarked or null, depending on the grammar of the language.

By “subject” is meant either the subject of the same clause, or the subject of the higher clause to which the clause in which the other NP occurs is a complement.<sup>2</sup> Furthermore, if an NP is antecedent to a pronoun, it may, of course, itself be a pronoun of any kind needing its antecedent or binder.

The phrase “at any stage of analysis” needs some comment. It has been said that a grammar establishes an identity relation between sentence representations at different levels: semantic analysis (SA), Syntactic Deep Structure, Shallow Structure, and Surface Structure. These are related by top-down (i.e., from SA to Surface Structure) transformations or bottom-up parsing rules. The parsing rules are the inverse of the transformations, so that the stages of bottom-up analysis are identical with those of the topdown generation. The Reflexive Principle applies at any stage of generation or parsing between SA and Surface Structure, not necessarily at one of the four defined levels of representation.<sup>3</sup> (These are defined by the kind of rules to which they are input or output.) It is thus possible to have marked reflexives that are not co-denotational with the surface subject, but with the semantic subject:

- (5) a. Ben had wanted *Graham* for a long time to look after *himself*.  
 b. Karel gelastte *Ben* voor *zichzelf* te werken.  
 (Karel ordered *Ben* to work for *himself*)  
 c. Ben told *Graham* a story about *himself*.

In (5a) the semantic subject *Graham* of the lower clause has been raised into the main clause by a well-known transformation. In (5b) *Ben* is object to *gelasten* (“order”), causing the deletion (by the rule usually called “Equi”) of the co-denotational subject of *werken* in the lower clause. In (5c) *Graham* is the subject, in SA, of the clause “*Graham* know a story about *himself*” in the overall structure “*Ben* caused<sub>s</sub>[*Graham* know a story about *himself*]”—with the semantic presupposition that the causing was mediated by verbal communication.

It is to be noted that (5a) does not allow for a reading in which *Ben*, i.e. the surface subject, is co-denotational with *himself*. This shows that the assignment of morphological reflexive marking takes place top-down: marking is assigned at the ‘earliest’ opportunity in the transformational process, not in the parsing process. This has the interesting consequence for AP that a considerable amount of top-down checking is required. This is not unnatural if we consider that AP takes as input SA’s of sentences plus their entire derivations to surface structures.

*Bound variable pronouns* are so called because they function logically and semantically as the well-known variables bound by quantifiers in predicate

calculus. For a pronoun to belong to this category it is necessary that it be preceded, in its surface structure, by a quantified NP. Furthermore, it must be commanded by it, according to the traditional definition of the relation “command”:

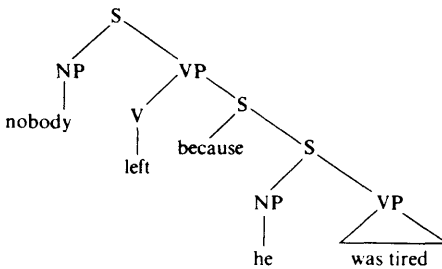
In a constituent tree structure, a node *A* *commands* a node *B* just in case the first S-node in the tree up from *A* also dominates *B*.

If a pronoun is both preceded and commanded by a quantified NP then its preferred interpretation is that of a variable bound by the quantified NP in question. This interpretation can be overruled by context or discourse factors, in which case the pronoun will take a sentence-external antecedent and belong to the category of denoting pronouns. Consider the following examples:

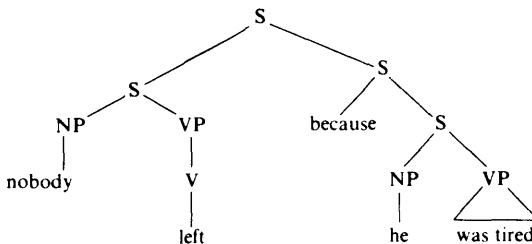
- (6) a. *Nobody* left because *he* was tired.  
 b. !*Nobody* left, because *he* was tired.  
 c. !Because *he* was tired, *nobody* left.  
 d. !Because *nobody* was tired, *he* left.

Assuming the following approximate surface structures for (6a) and (6b), respectively:

(6) a'



(6) b'



we see that only (6a) fulfills the conditions for the pronoun *he* to be interpreted as a variable bound by *nobody* in the sense of the predicate calculus analysis:

(6) a''  $\sim \exists x[x \text{ left because } x \text{ was tired}]$

It is, in fact, assumed that the SA of (6a) corresponds to (6a''), with the bracketed part analysed in a suitably abstract way. The first occurrence of the variable *x* is incorporated, by the rules of the grammar, into the quantified NP *nobody*; the second occurrence emerges in the surface structure as *he*. It is clear that in (6b-c) the pronoun *he* does not fulfill the conditions for a bound variable pronoun interpretation. In these sentences the pronoun *he* cannot but be interpreted as a denoting pronoun. Note that replacement of *he* by an epithetic NP leaves the interpretation intact:

- (7) a. Nobody left, because the kid was tired.  
 b. Because the little runt was tired, nobody left.  
 c. Because nobody was tired, the great man left.

The third category of pronouns distinguished, the *denoting pronouns*, has attracted most attention in anaphora research. In fact, since reflexives were recognized only in so far as they are morphologically marked, and since, at least in certain quarters, there was a reluctance to recognize bound variable pronouns as a separate category, the existing literature is characterized by the phenomenon that many pronoun occurrences were taken to be of the denoting kind whereas they are not. It is hardly surprising, therefore, that attempts to account for (sentence-internal) anaphora relations with the help of only one principle had to fail. Now that the category of denoting pronouns has been severely restricted we may expect that the problem of sentence-internal anaphora resolution will be less refractory.

The 'primordial' principle (Langacker, 1969) for the assignment of antecedents to internally anaphoric pronouns was that the antecedent must either precede or command the pronoun. This principle could account for a large number of cases, such as:

- (8) a. Since *Harry* felt unwell, *he* stayed at home.  
 b. Since *he* felt unwell, *Harry* stayed at home.  
 c. !*He* stayed at home, since *Harry* felt unwell.  
 d. *Harry* stayed at home, since *he* felt unwell.

It was, however, unequal to substantial classes of counterexamples, such as (3) or (4c). We have now seen that these are explained by the Reflexive Principle. Likewise, apparent complications arising in connection with bound variable pronouns, as in (6), are eliminated by the recognition of that category as a separate class. Note, in particular, that the Reflexive Principle solves seemingly difficult cases like:

- (9) a. !That *Mary* was clever, was something *she* knew very well.  
 b. That *she* was clever, was something *Mary* knew very well.

If *she* in these sentences were indeed a denoting pronoun, then, according to the primordial principle, anaphora as indicated should be all right in (9a) and not in (9b). In fact, however, *she* is not a denoting pronoun, as appears from:

(9) c. !That *the old girl* was clever, was something *Mary* knew very well.

The strict analogy with (4) above suggests very strongly that *she* in these cases must be treated as an indirect reflexive. All that is needed to let the Reflexive Principle apply is a level or stage of representation where (9b) shares with (4a) the property that the *that*-clause is structurally a complement of the clause *Mary knew very well*. Even though present-day grammatical theory does not provide a clear account of the grammatical relations between sets of sentences like (4) and (9), it is clear that such an account must eventually be found, given the semantic properties of the sentences in question. Not all counterexamples to the primordial principle disappear, however, with the separating out of reflexives and bound variable pronouns. Consider, for example:

- (10) a. The fact that *Mary* was clever made us suspect that *she* would come out first.  
 b. The fact that *she* was clever made us suspect that *Mary* would come out first.  
 c. The fact that *the old girl* was clever made us suspect that *Mary* would come out first.  
 d. The fact that *Mary* was clever made us suspect that *the old girl* would come out first.

Examples (10c) and (10d) show that the pronoun *she* in (10b) and (10a), respectively, is a true denoting pronoun. Although (10a) is unproblematic, in that the antecedent precedes the pronoun, (10b) should be impossible, since the antecedent *Mary* neither precedes nor commands the pronoun *she*. Quite a few more cases could be cited showing that Langacker's primordial principle is inadequate for denoting pronouns, even after the reflexives and the bound variable pronouns have been split off. How the old principle could be amended so that it covers all cases adequately, is not entirely clear.<sup>4</sup> It does seem, however, that internal anaphora resolution for denoting pronouns is perhaps not to be accounted for exclusively in structural terms, but also in terms of discourse phenomena such as topic and focus. We might venture the following principle:

- (11) For denoting pronouns carrying sentence-internal anaphora, the antecedent may follow the pronoun only if the pronoun is in a subordinate clause and if the denotation of the antecedent is the topic of the discourse in process. Otherwise, the antecedent must precede the pronoun.

This takes care of a significant number of counterexamples, and perhaps all. It is, however, still incomplete because it fails to cover a number of relevant cases. Thus, it has been observed (e.g. Akamajian & Jackendoff, 1970) that NP's under contrastive or emphatic accent are prevented from being antecedents for any pronouns not commanded by them:

- (12) a. !That *JIM* was going to be Harry's supervisor surprised *him*.  
 b. That *JIM* was going to be *Harry's* supervisor surprised *him*.  
 c. !That *Jim* was going to be *HARRY's* supervisor surprised *him*.  
 d. That *Jim* was going to be *HARRY's* supervisor surprised *him*.  
 e. People dislike *JIM* because *he* is abrasive, not *HARRY*.

Note that *he* in (12e) is not a denoting pronoun:

- (12) f. !People dislike *JIM* because *the fellow* is abrasive, not *HARRY*.

It is however, in (12b) and (12d):

- (12) g. That *JIM* was going to be *Harry's* supervisor surprised *the creep*.  
 h. That *Jim* was going to be *HARRY's* supervisor surprised *the old maniac*.

The answer to these cases would be given if it could be established that NP's under emphatic or contrastive accent function the same way the quantifiers do: they bind variables and require bound variable pronouns. This would explain why (12a) and (12c) are ruled out (the binder precedes but does not command the pronoun), while (12e) passes muster (the binder both precedes and commands the pronoun). In (12b) and (12d), on the other hand, we have to do with denoting pronouns, and here there is no problem, since the antecedent precedes the pronoun. In fact, we can have the antecedent and the pronoun swap places (provided the antecedent is topic), since the pronoun will then occur in a subordinate clause:

- (12) i. That *JIM* was going to be *his* supervisor surprised *Harry*.  
 j. That *he* was going to be *HARRY's* supervisor surprised *Jim*.

What is needed, therefore, is an analysis whereby emphatic or contrasted NP's function as quantified NP's. Such an analysis will not be attempted here.

Another class of difficult cases not covered by our analysis is provided by sentences containing anaphoric reference to an antecedent NP—not, however, taking over its denoting and/or referring function but making reference to its mere occurrence as a word:



- (13) You've lived in *Dnepropetrovsk* for twenty years, and you still find *it* an impossible name to pronounce!

Note that the inverse is not possible, at least not in the same way:

- (14) !“*Volapük*” means “world speak” in *it*.

Yet we do find:

- (15) It's called “*punch*” because it has *it*.

I prefer not to venture an explanation for these cases. A feeble excuse might be that our main interest here is in establishing antecedent–pronoun relations, not in determining the semantic role fulfilled by the pronoun once it has found its antecedent. (The excuse is feeble since, even though that might be so, the question still remains there to be solved. Moreover, even under the excuse (14) will remain unexplained.)

The same feeble excuse might help us through the problem area of so-called ‘sloppy identity’, or other forms of vicarious identity:

- (16) a. Harry drank *the milk*, but Leo threw *it* away.  
 b. *The temperature* was 50 degrees yesterday, but now *it* is much higher.  
 c. I used to be able to remember *the names of my students*, but nowadays I keep forgetting *them*. (De Rijk, 1974)

It is clear what the (external) antecedent is in these cases, but it is not clear how the pronoun is related semantically to the antecedent, since referential (or denotational) identity is absent (though there is some role analogy).

A much-quoted problem in philosophical circles is that of so-called ‘donkey-sentences’:<sup>5</sup>

- (17) a. If Harry owns *a donkey* he beats *it*.  
 b. Either Harry owns *no donkey* or he beats *it*.

The problem resides in the fact that there is no straightforward relation between what must count as the antecedent on the one hand and the pronoun (*it*) on the other. The pronoun, in these cases, is clearly of the denoting kind:

- (18) a. If Harry owns *a donkey* he beats *the poor animal*.  
 b. Either Harry owns *no donkey* or he beats *the creature*.

We should note that in our theory the question of the grammatical or external relation with an antecedent is of lesser importance than the determination of the right cognitive discourse ‘address’ for the pronoun to ‘land at’, i.e., the determination of the correct denotation relation between the denoting pronoun and an address in the discourse representation. This is

especially true for external anaphora, where no (or hardly any) grammatical restrictions exist. What is needed here is a theory of discourse representations that sets up suitable addresses for the pronouns in question. It seems that this is best done by treating these cases of anaphora as being sentence-external.

Denoting pronouns, as has been said, are the only pronoun category that can take external antecedents. These are, in principle, of two kinds. First, the antecedent may be a (salient) definite NP in a preceding sentence (conjunctions counting as successive sentences). Here the pronoun (if it is of the right gender and number) takes over the denoting function of the antecedent NP, —due allowance being made for 'sloppiness' or other denoting complications. Under certain still partially unknown<sup>6</sup> conditions the required antecedent NP may not be overtly present in the preceding discourse but retrievable on grounds of either background or situational knowledge.<sup>7</sup> Secondly, the antecedent may however have just been set up in the discourse representation by an existentially quantified NP in the preceding sentence (or conjunct), as in:

(19) Harry owns *a donkey*, and he beats *it*.

In these cases the semantics must be such that a new discourse address is set up for the donkey mentioned in the first conjunct, so that the pronoun *it* in the second conjunct can denote that address.

It is not too difficult to analyse sentences of the types (18a) and (18b) as either directly involving a conjunction like (19) or invoking the first conjunct on cognitive grounds. Both (18a) and (18b) would in fact be reducible (though with minor differences) to an abstract representation of:

(20) Either Harry owns no donkey, or he owns a donkey and he beats it.

The semantics will then set up two subdomains, one for each disjunct. In the first, Harry owns no donkey, but in the second Harry both owns a donkey and beats it. It is clearly not possible to elaborate here the semantics, the lexicon and the grammar needed for such an analysis.

A final and highly relevant problem to be discussed here is that of the so-called 'Bach-Peters paradox.' This is about sentences containing one or more denoting pronouns that are distributed in such a way that a *grammatical* characterization of the antecedents results in an infinite regress:

(21) a. The girl who got them liked the flowers that were sent ~~her~~.  
 b. The man for whom Ann bought the car he liked was not grateful.

Owing to their occurrence in relative clauses, the pronouns cannot be replaced by their antecedent NP's without recurring themselves. This appar-

ent paradox is solved if (a) it is accepted that pronouns occur in semantic analysis as pronouns, and not as their antecedents, and (b) a procedure of tentative denotation is developed, whereby tentative or provisional denotations are confirmed if consistency is achieved.

This last point is in need of some clarification. We assume that lexical definite NP's denote addresses that have either been set up in previous discourse or can be supplied on grounds of background or situational knowledge. They select the right address on the basis of their lexical material. Thus, taking (21a) as an example, there must be a discourse address  $d_1$  characterized by at least the predicates 'girl' and 'got the flowers that were sent her'. Likewise, there must be an address  $d_2$  characterized at least by the predicates 'flowers' and 'were sent to  $d_1$ '. More precisely:

$d_1$	$d_2$
girl( $x$ )	flowers( $x$ )
be sent( $d_2$ , to $x$ )	be sent( $x$ , to $d_1$ )
get( $x$ , $d_2$ )	get( $d_1$ , $x$ )

One can imagine these addresses having been set up on the basis of some amount of preceding discourse such as: "There was a girl. She was sent flowers. She got the flowers." Neither of the two main NP's in (21a), however, matches either of the addresses completely: *the girl who got them* matches  $d_1$  but for the pronoun *them*, which does not match  $d_2$  in any lexical sense. However, if the pronoun can be made to denote  $d_2$ , then indeed *the girl who got them* will denote  $d_1$ . So we provisionally assign  $d_1$  to *the girl who got them*, and likewise  $d_2$  is provisionally assigned to *them*. Now take the other main NP, *the flowers that were sent her*. Here the same complication arises: *her* does not match  $d_1$  lexically but could still denote  $d_1$ . Now we are in a quandary, since neither of the two NP's is assigned a definitive address. But if *them* denotes  $d_2$  and *her* denotes  $d_1$ , all is well. Note that they may well do so, since *the flowers that were sent her* fulfills the conditions for being antecedent to *them* (though it follows the pronoun, the pronoun is in a subordinate clause), and *the girl who got them* is allowed to act as antecedent to *her*. Moreover, the assignments required have already been made provisionally, so we cut the tie and confirm the provisional assignments, a confirmation to be overruled only on grounds of discourse plausibility, in which case one or both of the pronouns take external antecedents.<sup>8</sup>

Before an assignment procedure (AP) can be formulated, a summary specification of the syntax of SA must be given (for further analysis, see

Seuren, 1985). SA structures can be considered as constrained by the following formation rules:

$$\begin{aligned}
 (1) \quad S &\rightarrow \left\{ \begin{array}{l} V^1 + \left\{ \begin{array}{l} \text{NP}_s \\ S_s \end{array} \right\} \\ V^2 + \left\{ \begin{array}{l} \text{NP}_s \\ S_s \end{array} \right\} + \left\{ \begin{array}{l} \text{NP}_o \\ S_o \end{array} \right\} \\ V^3 + \left\{ \begin{array}{l} \text{NP}_s \\ S_s \end{array} \right\} + \text{NP}_i + \left\{ \begin{array}{l} \text{NP}_o \\ S_o \end{array} \right\} \end{array} \right\} \\
 (2) \quad \text{NP}_s &\rightarrow \left\{ \begin{array}{l} x \\ x: + S_x^{**} \\ \hat{x} + S_x \end{array} \right\} \\
 (3) \quad \text{NP}_i &\rightarrow \left\{ \begin{array}{l} x^{(R)} \\ x: + S_x^{**} \end{array} \right\} \\
 (4) \quad \text{NP}_o &\rightarrow \left\{ \begin{array}{l} x^{(R)} \\ x: + S_x^{**} \\ \hat{x} + S_x^{**} \end{array} \right\}
 \end{aligned}$$

Subscript s stands for 'subject'; o for 'object'; i for 'indirect object'.  $S_x$  stands for S with at least one argument  $x$ .  $S^{**}$  stands for any number of successive S's. ' $x: + S_x$ ' is to be read as 'the  $x$  such that  $s[\dots x \dots]$ '. Definite NP's are analysed (parsed) into this form. Thus, *the rose* is analysed as ' $x: \text{Rose}(x)$ ' or 'the  $x$  such that  $x$  is a rose'. Subsequent S's after the first end up as relative clauses or adjectives in surface structure. The operator  $\hat{x}$  denotes sets of potential addresses in any cognitive discourse representation: ' $\hat{x}(\text{Rose}(x))$ ' reads 'the set of potential addresses characterized by the predicate "Rose"'.

Such set denotations are used in the grammatical and semantic analysis of quantified NP's (see Seuren, 1985 for details). The quantifiers are analysed as two-term predicates ( $V^2$ ) which take set denotations as terms. The first term corresponds to the nuclear sentence, the second (object) term to the quantified NP. For example, (22a) is analysed as (22b) and reads as (22c):

- (22) a. *The man drives a new car.*  
 b.  $\exists 1 [\hat{x}(\text{Drive}(x:\text{Man}(x), x)), \hat{x}(\text{Car}(x), \text{New}(x))]$   
 c. "The set of things the man drives overlaps (with an overlap of at least one element) with the set of things that are cars and are new."

The discourse increment brought about by (22) consists in the setting up of a new address characterized by the lexical material 'car(x)', 'new(x)', and 'drive( $d_n, x$ )', where  $d_n$  stands for the address denoted by the NP 'the man' and its semantic analysis ' $x:\text{Man}(x)$ '. This increment value is indicated in the

assignment procedure to follow as  $i[S]$ . Finally, the subscript R indicates morphologically marked reflexivity (the brackets indicate optionality).

Having said this, we now proceed towards a formulation of the assignment procedure (AP) proper. It starts with an initial instruction used to decide whether the lexical or the pronominal procedure is to be followed. It ends with the final instruction F assigning the proper discourse address, unless AP has blocked somewhere in mid-course, in which case no assignment is made and the term in question remains uninterpreted. Surface pronouns always end up in SA as  $x$  (or as  $x_R$  if they are marked for reflexivity). A bound variable pronoun is considered 'fixed' in SA by the operator  $\hat{x}$  just in case it is an argument to the predicate of an  $S_x$  specified in the NP-rules for  $x$ . An occurrence of  $x$  in an  $S_x$  specified for the definite NP operator  $x$ : is also considered fixed by  $x$ :. Such occurrences of  $x$  only appear overtly in surface structure as reflexives (*the man who likes himself*) or as relative pronouns (*the car that he bought*). Otherwise, they are not overtly present, but 'hidden' in the grammar of NP's. If more than one  $x$  occurs in an  $S_x$ , the preferred procedure is to let  $AP[x_1] = AP[x_2]$ . However, due to overriding cognitive factors, one  $x$  may be treated as not fixed. The notation  $t_f$  is used to indicate the NP headed by the fixer of a fixed  $x$ .

Whenever the symbol  $t$  is used in AP, it refers to the definite NP-term under treatment. The symbol  $d$  ranges over addresses, while  $P(t)$  stands for the lexical (predicate) material characterizing terms. The expression  $d[t]$  is the denotation of a term  $t$  (i.e., its address).  $D$  stands for the domain in question, and  $G$  for the grammar of the language in question.  $t_s$  is the clause-mate subject term. The instruction  $\rightarrow$  is a traffic direction of the 'go-to' kind. Provisional address assignments are marked as  $d[t] =_p d_n$  (Provisional address assignments are marked thus:  $d[t] =_p d_n$ ). Provisional assignments to an address  $d_n$  are made when  $P(t)$  contains one or more terms  $t_u$  which do not match  $d_n$ , though all other material does. In such a case,  $d_n$  contains either an address-name  $d_k$  in the appropriate position, or  $x$ . If the latter, AP must assign to  $x$  a defined semantic role, as in (23). If the former, AP applied to  $t_u$  may now yield  $d[t_u] = d_k$ . If it does, the provisional assignment of  $d_n$  to  $t$  is confirmed. If neither  $d_k$  nor any other available address either fits  $t_u$  lexically or can be expanded (because of cognitive backing) so that it will fit, and if there is no cognitive backing for the post hoc insertion of a new address, then AP blocks. If  $t_u$  is pronominal (it follows from the system that it must be a denoting pronoun) AP looks for an internal antecedent NP. AP is then let loose on the internal antecedent term  $t_i$ . If  $d[t_i] = d_k$  or  $d[t_i] =_p d_k$  all previous provisional assignments are confirmed and  $d_k$  is assigned to  $t_i$ .

AP applies to the terms of a sentence in a hierarchical order. That is, AP starts with the terms of the highest predicate in the order subject-object-indirect object. Track is kept of assignments made so that terms that have

already been interpreted in the course of a complex-term interpretation or a pronominal interpretation can be skipped.

The *Assignment Procedure* (AP) now runs as follows:<sup>9</sup>

- I: Check if  $t$  is *Lexical*  $\rightarrow$  L-1  
       *Pronominal*  $\rightarrow$  P-1
- L-1: Check if some  $d_n$  (or the provisional  $d_n$ ) is matched by  $\mathbf{P}(t)$ .  
       *Yes*  $\rightarrow$  F  
       *No*  $\rightarrow$  L-2  
       *Tentatively*  $\rightarrow$  L-4
- L-2: Check if for some  $d_n$  (or the provisional  $d_n$ ) cognitive backing of  $\mathbf{D}$  allows for further characterization of  $d_n$  by whatever lexical material in  $\mathbf{P}(t)$  fails to match  $d_n$ .  
       *Yes*  $\rightarrow$  F  
       *No*  $\rightarrow$  L-3
- L-3: Check if cognitive backing of  $\mathbf{D}$  allows for the post hoc insertion of a  $d_n$  matching  $\mathbf{P}(t)$ .  
       *Yes*  $\rightarrow$  F  
       *No*  $\rightarrow$  BLOCK
- L-4: Provisionally assign  $d_n$  to  $t$ . Take the uninterpreted term  $t_u$  and provisionally assign any corresponding address  $d_k$  to  $t_u$  (i.e.,  $d[t_u] = {}_p d_k$ ). Apply  $I$  to  $t_u$ .
- P-1: Check if there is a  $d_n$  such that  $d[t] = {}_p d_n$ .  
       *Yes*  $\rightarrow$  P-7  
       *No*  $\rightarrow$  P-2
- P-2: Check if  $t$  is reflexive. (A pronominal term  $t$  is reflexive just in case (a)  $t$  is marked for reflexivity and required to be thus marked by  $G$  if P-2 yields 'yes', or (b)  $t$  is not marked for reflexivity and not required by  $G$  to be thus marked if P-2 yields 'Yes', and if morphological, lexical,<sup>10</sup> and cognitive factors allow for the selection of a (direct or indirect) subject.)  
       If *Yes*, select the appropriate  $t_s$ .  $\text{AP}[t] = \text{AP}[t_s]$ .<sup>11</sup> Confirm any provisional assignments made. CLOSE.  
       *No*  $\rightarrow$  P-3
- P-3: Check if  $t$  is fixed (i.e., if  $t$  is argument to a predicate of an  $S_x$  specified for  $x$ : or  $\hat{x}$ , and if there are no overriding cognitive or lexical factors blocking this path).  
       *Yes*  $\rightarrow$  P-4  
       *No*  $\rightarrow$  P-7
- P-4: If  $t$  is fixed by  $x$ : (i.e., a denoting NP)  $\rightarrow$  P-5  
       If  $t$  is fixed by  $\hat{x}$  (i.e., a set denotation)  $\rightarrow$  P-6
- P-5: For any  $d_n$  such that  $d[t_f] = d_n$  or  $d[t_f] = {}_p d_n \rightarrow$  F

P-6: → semantics of quantification<sup>12</sup>

P-7: Check if there is an internal antecedent term  $t_i$ . (A pronominal term  $t$  takes a term  $u$  as internal antecedent ( $t_i$ ) just in case (a)  $u$  is of the form  $x$  or  $x: + S_x^{**}$ , (b)  $u$  either precedes  $t$  or  $t$  is in a surface structure subordinate clause and  $u$  is topic, and (c) there are no cognitive or lexical factors blocking this path.)

If *Yes*, apply  $I$  to  $t_i$  if necessary. If there is a  $d_n$  such that  $d[t] = {}_p d_n \rightarrow$  P-10. Otherwise, for any  $d_n$  such that  $d[t_i] = d_n$  or  $d[t_i] = {}_p d_n \rightarrow$  F

If *No* → P-8

P-8: Check if there is an external antecedent term  $t_e$ . (A pronominal term  $t$  takes a term  $u$  as external antecedent ( $t_e$ ) just in case (a)  $u$  is of the form  $x$  or  $x: + S_x^{**}$ ,  $u$  occurs in a recent (preferably the preceding) sentence,  $t$  matches  $d[u]$  lexically, and  $d[u]$  is a plausible address for  $t$ .<sup>13</sup>)

If *Yes*, then if there is a  $d_n$  such that  $d[t] = {}_p d_n \rightarrow$  P-10. Otherwise, for any  $d_n$  such that  $d[t_e] = d_n \rightarrow$  F

*No* → P-9

P-9: Check if the preceding sentence or the preceding conjunct  $S$  has a non-negated existential quantifier.

If *Yes*, then if there is a  $d_n$  such that  $d[t] = {}_p d_n \rightarrow$  P-10. Otherwise, for any  $d_n$  such that  $i[S] = d_n \rightarrow$  F

*No* → L-3

P-10: Check if  $d[t_i] = d_n$  or  $d[t_i] = {}_p d_n$  or  $d[t_e] = d_n$  or  $i[S] = d_n$

*Yes* → F

*No* → BLOCK

F: Assign  $d_n$  to  $t$ . Confirm any previous provisional assignments.  
CLOSE.

Finally, let us illustrate AP with a few examples. Consider (23a) with the SA (disregarding tense) (23b), presupposing  $d_1$  and  $d_2$  as given above except that at this point in the imaginary discourse,  $d_1$  still lacks the information 'get ( $x, d_2$ )' and  $d_2$  'get ( $d_1, x$ )':

(23) a. *The girl got the flowers that were sent to her.*

b. Get [ $x_1$ : Girl( $x$ ),  $x_2$ : Flowers ( $x$ ), Be sent ( $x$ , to  $x_3$ )]<sup>14</sup>

AP:  $t = x_1$     I:    Lexical → L-1  
                   L-1: Yes ( $d_1$ ) → F  
                   F:     $d[x_1] = d_1$     CLOSE  
 $t = x_2$     I:    Lexical → L-1  
                   L-1: Tentatively ( $d_2$ ) → L-4  
                   L-4:  $d[x_2] = {}_p d_2$   
                        $d[x_3] = {}_p d_1$

$t = x_3$	I:	Pronominal $\rightarrow$ P-1
	P-1:	Yes $\rightarrow$ P-7
	P-7:	Yes: $t_i = x_1$ ; $d[t_i] = d_1 \rightarrow$ P-10
	P-10:	Yes $\rightarrow$ F
	F:	$d[x_3] = d_1$
		$d[x_2] = d_2$ CLOSE

Note that we have skipped the  $x$  in 'Girl( $x$ )', in 'Flowers( $x$ )', and in 'Be sent( $x$ , to  $x_3$ )'. Clearly, these would be immediately revealed to be pronominal, without provisional assignment, and fixed by their fixing  $x_1$ ; or  $x_2$ :. In virtue of P-5, they thus simply take over the assignment made for the whole term. Since, however, the whole term has been assigned an address already by AP, we are fully justified in ignoring these vacuous  $x$ 's. Having now assigned all definite NP's in (23b) to their proper discourse addresses, the semantics now increments  $d_1$  and  $d_2$  as shown above. That is,  $d_1$  is incremented with 'get( $x$ ,  $d_2$ )', and  $d_2$  with 'get( $d_1$ ,  $x$ )'.

Let us now try the Bach-Peters paradox case (21a), based on the addresses  $d_1$  and  $d_2$  as given above. The SA of (21a) is

(21a)'. Like [ $x_1$ : Girl( $x$ ), Get( $x$ ,  $x_2$ ),  $x_3$ : Flowers( $x$ ), Be sent( $x$ , to  $x_4$ )]

AP: $t = x_1$	I:	Lexical $\rightarrow$ L-1
	L-1:	Tentatively ( $d_1$ ) $\rightarrow$ L-4
	L-4:	$d[x_1] = {}_p d_1$
		$d[x_2] = {}_p d_2$
$t = x_2$	I:	Pronominal $\rightarrow$ P-1
	P-1:	Yes $\rightarrow$ P-7
	P-7:	Yes: $t_i = x_3$
$t = x_3$	I:	Lexical $\rightarrow$ L-1
	L-1:	Tentatively ( $d_2$ ) $\rightarrow$ L-4
	L-4:	$d[x_3] = {}_p d_2$
		$d[x_4] = {}_p d_1$
$t = x_4$	I:	Pronominal $\rightarrow$ P-1
	P-1:	Yes $\rightarrow$ P-7:
	P-7:	Yes: $t_i = x_1$
		$d[t_i] = d[x_1] = {}_p d[x_4] = {}_p d_1 \rightarrow$ P-10
	P-10:	Yes $\rightarrow$ F
	F:	$d[x_4] = d_1$ $d[x_3] = d_2$
		$d[x_2] = d_2$ $d[x_1] = d_1$ CLOSE



We try (21b) next. It presupposes the following three addresses:

$d_3$	$d_4$	$d_5$
man(x)	'Ann'(x)	car(x)
like(x, $d_5$ )	buy(x, $d_3$ , $d_5$ )	like( $d_3$ , x)
buy( $d_4$ , x, $d_5$ )		buy( $d_4$ , $d_3$ , x)

The SA of (21b) is:

(21)b'.  $\sim$ Grateful[ $x_1$ :Man(x), Buy( $x_2$ :'Ann'(x), x,  $x_3$ :Car(x), Like( $x_4$ , x))]

AP: $t = x_1$	I:	Lexical $\rightarrow$ L-1
	L-1:	Tentatively ( $d_3$ ) $\rightarrow$ L-4
	L-4:	$d[x_1] =_p d_3$
		$d[x_2] =_p d_4$
$t = x_2$	I:	Lexical $\rightarrow$ L-1
	L-1:	Yes $\rightarrow$ F
	F:	$d[x_2] = d_4$
$t = x_3$	I:	Lexical $\rightarrow$ L-1
	L-1:	Tentatively ( $d_5$ ) $\rightarrow$ L-4
	L-4:	$d[x_3] =_p d_5$
		$d[x_4] =_p d_3$
$t = x_4$	I:	Pronominal $\rightarrow$ P-1
	P-1:	Yes $\rightarrow$ P-7
	P-7:	Yes: $t_i = x_1$
		$d[t_i] = d[x_1] =_p d[x_4] =_p d_3 \rightarrow$ P-10
	P-10:	Yes $\rightarrow$ F
	F:	$d[x_4] = d_3$
		$d[x_3] = d_5$
	$d[x_1] = d_3$	CLOSE

The following sentence contains the reflexive *himself*:

- (24) a. The man who loved himself was vain.
- b. Vain[ $x_1$ :Man(x), Love( $x_2$ ,  $x_R$ )]

It presupposes an address of at least the following form:

$d_6$
man(x)
love(x, x)

AP: $t = x_1$	I: Lexical $\rightarrow$ L-1
	L-1: Tentatively ( $d_6$ ) $\rightarrow$ L-4
	L-4: $d[x_1] = {}_p d_6$
$t = x_R$	I: Pronominal $\rightarrow$ P-1
	P-1: No $\rightarrow$ P-2
	P-2: Yes: $t_s = x_2$ ; $d[x_1] = d_6$
	AP[ $x_R$ ] = AP[ $x_2$ ]
	= address-bound $x$ in $d_6$ .      CLOSE

Finally, let us consider a simple case of variable binding:

- (25) a. Some American expects that he will win.  
 b.  $\exists x [\hat{x}(\text{Expect}(x_1 \text{ Win}(x_2))), \hat{x}(\text{American}(x))]$

No addresses are presupposed. The semantics for set-denotations is not given here, but for a summary indication see footnote 12. From this it follows that the semantic result of (25) consists in the setting up of an address which is characterized by the lexical material in both the subject-term and the object-term. This is unproblematic, provided the  $x$ 's are fixed by the set-denoting operator. The only term which is not so fixed is  $x_2$ . We therefore need to apply AP to  $x_2$ :

AP: $t = x_2$	I: Pronominal $\rightarrow$ P-1
	P-1: No $\rightarrow$ P-2
	P-2: Yes (ind. refl.): $t_s = x_1$
	AP[ $x_2$ ] = AP[ $x_1$ ] = address-bound $x$
	in new address $d_7$ CLOSE

This enables us to set up the new address:

$d_7$
American( $x$ )
Expect( $x$ , Win( $x$ ))

We observe that a sentence like

- (26) a. Some American wants to win.

has an SA which runs exactly parallel to (25b), except that 'Expect' is replaced by 'Want', as in (26b). The difference in the surface structure is brought about by the fact that the lower subject (*he* in (25a)) has been deleted in virtue of obligatory like-subject deletion for the English verb *want*. There has been a lack of clarity in the literature on the conditions to be fulfilled by the two subjects for them to be called 'like'. The answer is now, in principle, given: the higher subject and the lower subject are 'like' just in case both have the same AP-value.

That is, either they denote the same address or they are both added as  $x$  to the same address or addresses. The parser will then have to attach to  $x_2$  in:

(26) b.  $\exists ! [\hat{x}(\text{Want}(x_1, \text{Win}(x_2))), \hat{x}(\text{American}(x))]$

the condition that  $\text{AP}[x_2]$  must be identical to  $\text{AP}[x_1]$ , the latter being the higher subject under the verb *want*. This means that AP is to be extended with an extra proviso for P-2 to the effect that when the pronoun in question is not overtly present in surface structure, it must be an indirect reflexive. That is, condition (b) must then be fulfilled:  $t$  is not marked for reflexivity and not required by  $G$  to be thus marked, though it is a reflexive, and no morphological, lexical, or cognitive factors prevent AP-identity with a (higher) subject. P-2, therefore, has to be amended in that the answer 'No' will lead to 'BLOCK' when the pronoun is zero in surface structure, but to the instruction to go to P-3 in the other cases. These and other refinements, however, are left to a more comprehensive study.

## NOTES

<sup>1</sup> The exclamation mark is used to indicate the impossibility of a co-denotation (or co-variable) relation between the NP's in italics.

<sup>2</sup> Note that morphological marking of reflexive pronouns in English is not simply conditional on a co-denotation (or co-variable) relation with the subject of the same clause. The pronoun in question must also not be in a secondary VP, as in:

(i) *Ben* wanted the parcel to be delivered to *him* before noon.

When the secondary VP is 'weakened' by the deletion of the infinitive, as in:

(ii) *Ben* had the parcel delivered to *him(self)* before noon.

morphological marking seems optional. In any case, however, the pronouns in (i) and (ii) are reflexive. Witness the impossibility of, for example,

(iii) ! *Ben* wanted the parcel to be delivered to *the old rascal* before noon.

<sup>3</sup> Thus, for example, in the Dutch sentence:

(1) Ik wilde *Karel* gelasten *Ben* voor *zich* te laten werken.  
(I wanted to order *Karel* to make *Ben* work for *him*).

*Karel* is not a clause-mate subject with *zich* at any of the four levels of representation. It does, however, fulfill the condition at the stage just before the NP which is the subject of the clause  $s_v[\text{[laten werken]} \text{ NP, Ben, voor zich}]$  is deleted by Equi under identity with *Karel*, which is the subject of *gelasten* ("order").

<sup>4</sup> What is clear is that the introduction of the new notion of C-command ('a node A C-commands a node B just in case A's first dominating node also dominates B'), as in Reinhart (1976), is also insufficient, as is clearly demonstrated by example (10).

<sup>5</sup> This problem was mooted by Geach (1962, pp. 128ff.), who gives the examples:

- (i) If *any man* owns a donkey, *he* beats it.
- (ii) If *Smith* owns a donkey, *he* beats it.

<sup>6</sup> See Sanford & Garrod (1981). There seems, in any case, to be an overriding principle that external antecedents are assigned according to the greatest cognitive plausibility. Other factors, however, also play a role, such as similar syntactic function and intonation:

- (i) Harry kicked *Sam* and then Bill kicked *him*.

'*Sam*' and '*him*' can only be co-denotational if '*him*' is unaccented. With accent on '*him*', '*Harry*' must be the antecedent (see also Akmajian & Jackendoff, 1970).

<sup>7</sup> It seems (see Tasmowski-De Rijck & Verluysen, 1982) that in these cases the suppletion is mediated through a tacit linguistic form. The evidence adduced is from French: when a person A is, for example, trying to get a heavy desk (masculine: *le bureau*) through a door, another person B might say:

- (1) Tu ne *le* feras jamais passer par là.  
'You will never get *it* (masc.) through there.'

However, if the object is a table (feminine: *la table*), B would say:

- (2) Tu ne *la* feras jamais passer par là.

With the feminine pronoun *la* (it).

<sup>8</sup> In (21b) the pronoun 'he' is part of its antecedent NP. Cases of this kind must be accounted for by a refined definition of the notion 'antecedent' for denoting pronouns. Note that the pronoun is of the denoting kind:

- (1) *The man* for whom Ann bought the car *the pig* liked was not grateful.

<sup>9</sup> I am indebted to Wietske Vonk, Leo Noordman, and Ton Weyters for a thorough discussion of AP.

<sup>10</sup> By 'lexical factors' is meant the very general lexical characterization of pronouns which they inevitably seem to have.

<sup>11</sup> The expression AP[*t*] is used to indicate the final semantic result of AP. This need not be an address assignment in all (successful) cases. AP[*t*] may consist in treatment by the semantics of quantification, as mentioned in P-6 and in note 12. It may also be the placement of a term *x* under a predicate *P* in an address *d<sub>n</sub>*, as in (24).

<sup>12</sup> The semantics of quantification (i.e., the increment values of sentences containing quantified terms) is left out of our account. Essentially, the quantifying predicate '∃' is an instruction to set up one or more new addresses characterized by the lexical material in both of its set-denoting terms. The quantifying predicate '∀' is an instruction to add to all addresses answering the description of the object term the description of the subject term. Pronominal *x*'s fixed by the set denotator 'x' are taken care of by the semantics of quantification.

<sup>13</sup> The full set of conditions for a term to be an external antecedent is not yet known, as has been said before (note 6). 'Plausibility' includes lexical (note 10) and morphological matching, as well as cognitive congruence. As for the latter, cp:

- (1) Audrey asked Sue to drive Helen to the station.
  - a. *She* didn't want her daughter to miss the train.
  - b. But if *she* didn't have time, Jeeves could do it.
  - c. Otherwise *she* would miss her train.

<sup>14</sup> The subscripts are added merely for ease of reference. They have no semantic significance.  $x_1$  stands for the whole of  $x_1$ : Girl( $x$ ), etc.:  $x_n$ ; stands for the whole denoting term fixed by  $x_n$ . The term  $x_3$  in the  $x_2$  term of (23b) is singled out for separate indexing because the application of AP to  $x_2$  yields a provisional assignment for  $x_3$ .

## REFERENCES

- Akmajian, A., & Jackendoff, R. Coreferentiality and stress. *Linguistic Inquiry*, 1970, 1(1), 124-126.
- De Rijk, R. P. G. A note on prelexical predicate raising. In P. A. M. Seuren (Ed.), *Semantic syntax*. Oxford: Oxford University Press, 1974.
- Geach, P. T. *Reference and generality: An examination of some medieval and modern theories*. Ithaca, NY: Cornell University Press, 1962.
- Langacker, R. W. On pronominalization and the chain of command. In D. A. Reibel, S. A. Schane (Eds.), *Modern studies in English*. Englewood Cliffs, NJ: Prentice-Hall, 1969.
- Reinhart, T. *The syntactic domain of anaphora*. Unpublished doctoral dissertation, MIT, Cambridge, MA. 1976.
- Sanford, A. J., & Garrod, S. *Understanding written language: Explorations in comprehension beyond the sentence*. Chichester: Wiley, 1981.
- Seuren, P. A. M. *Predicate raising and dative in French and sundry languages*. Trier: Linguistic Agency University Trier, 1972.
- Seuren, P. A. M. *Discourse semantics*. Oxford: Blackwell, 1985.
- Tasmowski-De Ryck, L., & Verluyten, P. Linguistic control of pronouns. *Journal of Semantics*, 1982 1(4), 323-346.