Pronouns in a Derivational Syntax

FREDRIK HEINAT

1   Introduction

Binding theory has been one of the central modules in syntactic theory since the early eighties. However, research on syntax in the last ten years has given binding theory less attention. One problem is that even though core concepts used in binding theory, e.g. governing category and binding domain, have been abandoned in all other areas of syntax, syntactic theory still makes an informal use of the binding principles as they were formulated in Government and Binding theory. Another problem is the fact that in a derivational approach to syntax, such as Chomsky (2001a, b) and Epstein et al. (1998), there does not appear to be a place for conditions that apply on representations. Since the binding conditions are exactly these kinds of conditions they pose a real problem to any derivational approach to syntax. From a theoretical point of view that means we must find a way to incorporate the empirical predictions of the binding conditions into a theory that does not make use of representations but instead uses merge and move as the only ways to form syntactic relations. In the last few years there have been proposals how we can solve at least some of the problems that binding theory used to take care of.

In this paper I will outline and modify two of those proposals, namely Kayne (2002) and Zwart (2002). They both have in common that they view anaphors and their antecedent as one syntactic element. There are several reasons why this is an attractive idea. First, we can dispense with the notion of co-indexing. Second, we do not have to stipulate that the binding conditions apply at a certain level of representation, since there are no binding conditions. Third, it makes it possible to incorporate binding theory in a derivational syntactic theory. The outline of the paper is as follows. In the second section I give a very brief overview of the problems the Minimalist Program (Chomsky 1995) introduces for binding theory. In the third section I will briefly review and point out some of the problems with Kayne's and Zwart's analyses. The fourth section contains my analysis which makes use of ideas from both Kayne's and Zwart's papers. The fifth and last section is a concluding discussion where areas of future research are outlined.

2   The Minimalist Program

This short exposé of the Minimalist Program (Chomsky 1995) and Derivation by Phase (Chomsky 2001) will deal only with what is essential to the study of binding relations. The standard definitions of the binding conditions A, B, and C, as in (1), rely on concepts such as governing category, coindexing and complete functional complex.

(1)   The binding principles (from Chomsky 1986):
         Principle A: An anaphor is bound in a local domain
         Principle B: A pronominal is free in a local domain
         Principle C: An R-expression is free

In these definitions *bound* involves c-command and coindexing and *local domain* involves governing category (which in its turn uses the notion of barrier) and complete functional complex, i.e. the requirement that a local domain has a subject. This makes the local domain more or less equivalent to a clause and a noun phrase with a specifier.

Chomsky (1995) gets rid of the syntactic relation government and replaces it with the head-complement relation and head-specifier relation that come for free in the syntactic representation.

In the Minimalist Program Chomsky also introduces the inclusiveness condition (1995:228). The inclusiveness condition states that any structure formed by the computational system must be formed by the elements present in the lexical items selected. In other words, no new elements, e.g. indices, can be added to the computation when an item is selected for the lexical array. Since coindexing is a notion of coreference it obvious that lexical elements cannot bear their index in the lexicon.

With the introduction of Derivation by Phase (Chomsky 1999, 2001a,b), the systems that interface with narrow syntax have more than one interface point. Spell-out applies at each strong phase, vP and CP (and possibly DP). The consequence of this is that there is no 'final' LF where reconstruction or checking of binding can take place (Chomsky 2001a:4). Consequently, the question is where does the checking of the binding principles take place. If we retain the notion of coindexing, there are several possible solutions: the binding principles are all checked at one and the same strong phase, they are checked at different phases or they are checked at more than one phase. Note that Chomsky (2001a) still posits some 'final' LF where binding relations can be checked.
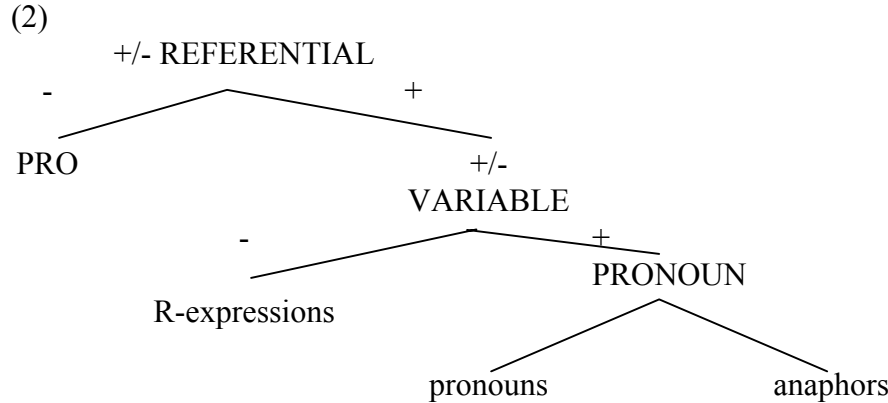
This phase-based approach to syntactic computation obviously requires that the binding theory be reformulated to accommodate the theory. In the following section I will look at two similar approaches to binding of pronouns and anaphors that simply do away with binding and let the syntactic derivation do the work of establishing binding relations.

3   Previous work

In this section I will look at two similar proposals how we can get rid of binding theory simply by letting the syntactic derivation build the relations between pronouns/anaphors and their antecedents. They all more or less adhere to the tenet of Epstein et al (1998:3): "the structure-building rules Merge and Move (Chomsky 1994) naturally expresses all syntactically significant relations; that is, if X and Y are concatenated, then X and Y naturally enter into syntactically significant relations". The consequence of this is that Kayne (2002) and Zwart (2002) both have as a basic structure a complex DP with antecedent and pronoun or anaphor merged together. Their approaches differ as to how the movement of the antecedent is triggered and how far the antecedent can move. The structure of this section is as follows. The first part examines Zwart's (2002) analysis of pronouns and their antecedents. The second part takes a look at Kayne's (2002) analysis of pronouns and antecedents.

3.1 Zwart (2002)

Zwart starts out with questioning the classic (Chomsky 1981) division of NPs[1] into four categories based on the distinction +/- anaphor and +/- pronominal. Zwart (2002:273) suggests instead the following hierarchical order of NP types (2).

(2)
```
                +/- REFERENTIAL
       -                        +
     PRO                      +/-
                            VARIABLE
              -                     +
                                PRONOUN
          R-expressions
                          pronouns        anaphors
```

Based on especially body-parts anaphors as in (3) (Fulani) (from Sylla (1993)), Zwart's argument is that anaphoricity is not a lexical property, rather it is a syntactic feature acquired in the syntax. Zwart deals only with locally bound anaphors and disregard long-distance anaphors and what is usually called logophors.

(3) a. en   tooñ-ii   koye   men
         we   harm-ASP   heads   our
         We have harmed ourselves
     b. koye   men kell-ii
         heads   our hurt-ASP
         Our heads hurt

Rather, he says, syntax recognises only one category with the features [+variable, +referential]. This category is PRONOUN. In syntax this is a root that may acquire features in the syntactic derivation. The features acquired in the syntactic derivation have repercussions on morphology in the PF component. This particular feature, [+coreferential] can only be acquired in one way, by merging the PRONOUN to its antecedent. The result is the structure shown in (4).

(4)   [$_{XP}$ [Antecedent] [PRONOUN]]

This merge takes place before the complex XP is merged as complement to the verb. Zwart assumes that the PRONOUN is a head, the antecedent which is a DP presumably occupies the spec-position of the XP.

A feature [+coreferential] invariably leads to the spell-out of an anaphor (in English and some other languages). Zwart points out the fact that there is no implication that it is the PRONOUN that gets the [+coreferential] feature. Languages differ in the ways the feature is morphologically marked. Some languages mark the pronoun (e.g. English); some languages mark the antecedent (e.g. Ponapean (5a)); some languages mark both the antecedent and the

---

[1] I will use DP and NP interchangeably, without any specific assumptions.

PRONOUN (e.g. Lezgian (5b)); whereas others do not mark neither of them (e.g. Tiri (5c)). (Data cited in Zwart 2002:297)

(5) a. Irail   pein   duhp-irail   (Rehg 1981)
          They self   bath-them
          'They bathed themselves.'
      b. Ada   wiči   wič   alzurar-zawa   (Haspelmath 1993)
          He-erg self-erg self-abs   deceive-impf
          'He is deceiving himself.'
      c. Nrâ   dreghe nrî   fadre   rroto   (Osumi 1995)
          He-subj injure He-obj   with   car
          'He injured himself in a car.'

It is crucial to morphology that this feature is present, if it is absent the underspecified PRONOUN, i.e. a pronoun, will be used in PF. The derivations with (6) and without (7) coreferential DPs look the following ways:

(6)   likes [[Mary] [PRONOUN]] $\rightarrow$ Mary likes [[Mary$_{copy}$] [PRONOUN]]
        PF: Mary likes herself.
(7)   likes [PRONOUN] $\rightarrow$ Mary likes [PRONOUN]
        PF: Mary likes her.

In (6) the merging of the antecedent and the root pronoun leads to a [+coreferential] feature on the pronoun and consequently it will be spelled out as an anaphor, *herself*. In (7) there is no merge of the two and the root pronoun will be spelled out without the feature [+coref.]. The consequence is that PF will assign the root the form of *her* (or another form that fits the pronoun's features). This appears to be a violation of the inclusiveness condition, since it is not plausible to think that neither the antecedent nor the PRONOUN would have such a feature in the lexicon. It appears as if this feature is actually introduced in the actual merging of the two elements.

    The reasons why the antecedent in (6) moves out of the complex phrase are its need to get a theta role and check case. The positions where the antecedent can do this are by definition A-positions. As Zwart points out, A-movement is usually clause bound. The exception is ECM constructions. In order to explain them, Zwart makes use of Chomsky and Lasnik's (1993) notion of *L-domain*. Zwart (2002:279) then defines the local domain for A-movement of α as "… the maximal projection of the highest functional head *f* L-related to a verb *V* L-related to a functional head *f'* licensing α." In other words, if the subject argument of a verb in an embedded clause is licensed by a functional projection associated to a verb in a matrix clause. This is a regular ECM-construction. As a result the antecedent can move from the PRONOUN and get a theta role and case from the matrix verb within the local domain, illustrated in (8).

(8)   Mary saw [[Mary$_{copy}$] [PRONOUN]] kiss John
        PF: Mary saw herself kiss John.

The presence of a CP invariably blocks this movement of the antecedent since C is not an L-related head according to Chomsky and Lasnik (1993). As a consequence the coreference between *Mary* and *her* (if there is one) in (9) must be accidental.

(9)   Mary saw that she kissed John.

Zwart gives several reasons why coreference of *Mary* and *her* is the preferred interpretation (2002:288-290). One of the reasons is that (9) is ambiguous and there is a possibility for the pronoun to refer to someone other than *Mary*. I would argue that this line of reasoning looks at the ambiguity from a parsing perspective. From a production perspective there is no reason why the coreference not be marked. This must be crucial information to the C-I interface and LF. But assuming that there is accidental coreference, there are no obvious reasons why there could not be accidental coreference in (10), a classic principle C violation.

(10)   She saw that Mary kissed John.

In conclusion this is a very interesting approach to trying to account for the empirical facts that the binding principles took care of in GB, by using the two operations that come for free in syntax, merge and move. There are several problems to this approach but since Kayne's has the same problems I will postpone the discussion of them until after the outline of his paper.

3.2   Kayne (2002)

Kayne (2002) has a similar approach to coreference as Zwart but he goes one step further and claims that:

> Antecedent –pronoun relations as in [John thinks he is smart] REQUIRE movement out of a constituent of the form [*John-he*]. That is the ONLY way to express an antecedent-pronoun relation.(2002:138).

The structure of this complex phrase is the same as in Zwart's paper. One crucial difference between their approaches is that Kayne assumes that the head of the XP is either a pronoun (11a), an anaphor (11b) or PRO (11c).

(11)   a.   Mary thinks [[Mary$_{copy}$] she]] is smart
          b.   Mary likes [D^0 [Mary$_{copy}$] [her] ('s) self]
          c.   Mary tried to [[Mary$_{copy}$] PRO]] solve the problem

Kayne, too, assumes that the movement of the antecedent is driven by its need for a theta-role and case. Kayne claims that the antecedent-pronoun relation holds across sentences, too. In (12) *Mary* starts out in the specifier-position of the complex phrase [[Mary] [she]].

(12)   Mary is a skier. She is a runner, too.

This is a clear step away from what Hornstein (2001) claims to be one of the unquestionable features of language, namely that sentences are the basic linguistic unit. In addition, this way of looking at coreference across sentences seems to be incompatible with the notion of phase (Chomsky 2000, 2001a and b) (see discussion of (22)). Also, the first phrase containing the antecedent in a discourse must contain all the pronouns that will be used in that discourse. It is not plausible to think that the working memory could store that much information.
      Central in Kayne's and Zwart's analyses is that extraction is only possible from the spec-position of the complex DP. The question Kayne sets out to answer is why there cannot be any coreference in a sentence like (13) in contrast to the sentence in (11).

(13) Mary praises her.

With reference to Romance clitic movement and the assumption that unstressed pronouns must move, Kayne suggests that the pronoun, i.e. the doubling constituent, must move to an A-bar position above the subject theta position. The reason for this movement is very unclear and Kayne's (2002:145) suggestion that the doubling constituent must move because it is a doubling constituent is circular. Kayne claims that there is no licensing position for the pronoun in the VP and it has to move and the antecedent is pied-piped so the whole constituent has to move. In (11) the doubling constituent can move to spec-IP, and thereby license the pronoun *her*. The antecedent, *Mary*, can then move up to the theta position of *think*. In (13), on the other hand there is no way that the antecedent can get its theta role once the doubling constituent has moved above the subject theta position, since downward movement is not allowed. This analysis seems problematic since the pronoun in (13), according to Kayne, must have an antecedent in the specifier position, be it *Mary* or someone else. In any case the antecedent will be trapped in the doubling constituent and it will fail to get a theta role and the derivation should crash, which it obviously does not if *her* refers to someone else than *Mary*.

Kayne (2002:147) claims that this is the reason language has reflexive pronouns; there must be some way to express a relation between *Mary* and the object in (13). The addition of *self* introduces a position the doubling construction can move to without moving above the theta-position for the subject. The structure Kayne proposes is the one in (11). The specifier position in the possessive type DP is the intermediate position for [Mary she]. Kayne notes the problem that the antecedent can move long-distance out of possessives (14), but not out of reflexives (15). Another problem to this analysis is the fact that some languages do not have reflexive objects as Zwart points out (2002:297).

(14) Mary wants me to take care of her cat. Her = [[Mary she] 's]
(15) *Mary wants me to take care of herself. Herself = [[Mary she] 's self]

The fact that the antecedent can move out of a construction like (16) is explained as movement of *Mary* from the spec of the lowest position to the specs of the two pronouns and finally the antecedent gets its theta role in the subject position of the main clause verb (2002:160).

(16) Mary$_i$ said that [t$_i$ she] thinks [t$_i$ she]'ll be late.
(17) Mary said that she wants me to take care of her

It is still unclear how the antecedent can move out of the doubling constituent in (17), since the doubling constituent is trapped below the subject theta-position of the lowest verb, see the discussion of (13).

In his discussion of backwards pronominalization, (18), Kayne's solution is extensive remnant movement. He also claims that since these constructions are not possible in many languages, e.g. Swedish, there are reasons why they are so complicated to derive.

(18) Her mother loves Mary

In this approach to coreference it is necessary to allow sideward movement. Kayne states that the availability of sentences such as (19) can only be derived if we allow sideward

movement in the sense of e.g. Nunes (2001) and Bobaljik and Brown (1997). Though he stresses the fact that such movement can only take place to the left.

(19) [The book [that Mary$_i$ is reading]] annoys her$_i$.

In the preceding two sections I have given a short summary of the main points of two similar analyses of coreference in a derivational approach to syntax. The effect of both analyses is that the former binding conditions can be abandoned; all configurations that used to be filtered out by the binding conditions can be explained (at least to some extent) by the way move and merge work. I will conclude this section by listing some of the advantages that the analyses have, and some of the problems that they do not solve.

In the two analyses presented above there is no use of the concepts government, barrier, and complete functional complex. By establishing coreference by merger (at least in Kayne's analysis) there is no need for something like co-indexing. By using the (only) syntactic structure building rules, merge and move, it is possible to derive the effects of the binding principles and thereby dispense with them. With no binding principles, the problem of where and when they apply does no longer exist. Since they do not exist we do not have to posit a final LF in the sense of Chomsky (2001a).

However, in my view there are still some problems with the above analyses. I will end this section by focussing on some of the problems that I think a modified and unified approach of Kayne's and Zwart's approaches can take care of. If we want to explain coreference by means of move and merge, Zwart's analysis comes short since all coreferentiality between a pronoun and its antecedent has to be coreferential. In this respect Kayne's analysis is more appealing since it takes care of all coreference. The drawback of his analysis on the other hand is that coreference is a notion that spans across sentences. Another problem is that there has to be a lot of (remnant) movement without any apparent triggers. Yet another problem with Kayne's analysis is the fact that the antecedent merges with a morphologically specified pronoun. Zwart on the other hand makes use of the root pronoun PRONOUN, which lets the morphological/PF-component take care of the actual form. In the following section I will present a short sketch of Distributed Morphology. In section 4 I will present an analysis that makes use of both Zwart's and Kayne's analyses and which also explores the notions of phases. This analysis views the morphology of pronouns as something that can be accounted for within the framework of Distributed Morphology.

3.3   Distributed Morphology

Distributed Morphology (DM) gets its name from the structure of the grammar; morphology is not concentrated to a single component of the grammar but distributed over several different components (Halle & Marantz 1993:111-112). According to Harley and Noyer (1999:3) there are three core properties of DM that makes it different from other morphological theories: Late Insertion, Underspecification and Syntactic Hierarchical Structure all the way down.

Late Insertion is the hypothesis that the phonological expression is in all cases provided in the mapping to phonological form. According to them syntactic categories are purely abstract. Phonological expressions, Vocabulary items, are inserted in a process called Spell Out after the syntactic derivation.

Underspecification of Vocabulary Items means that there is no need for phonological expressions to be fully specified for the syntactic positions where they can be inserted. "Vocabulary Items are in many cases default signals, inserted where no more specific forms are available" (1999:3).

Syntactic Hierarchical Structure all the way down means that the elements in the syntactic derivation and in morphology enter into the same type of constituent structures, i.e. binary trees.

Harley and Noyer (1999) describes the grammar in DM in the following way (figure 1). The syntax generates structure by manipulating (merging and moving) morphological features, taken from the Lexicon's List A. At Spell-Out the features are sent to LF and to a morphological component where they undergo morphological operations and get phonological form. Here Vocabulary Items are inserted from Lexicon, List B. The structures from phonology and LF converge at the Conceptual Interface, which also has access to List C, Encyclopedia. The important thing for my analysis is that in syntax the phonological form of a pronoun does not matter. The actual morphology comes after spell-out.
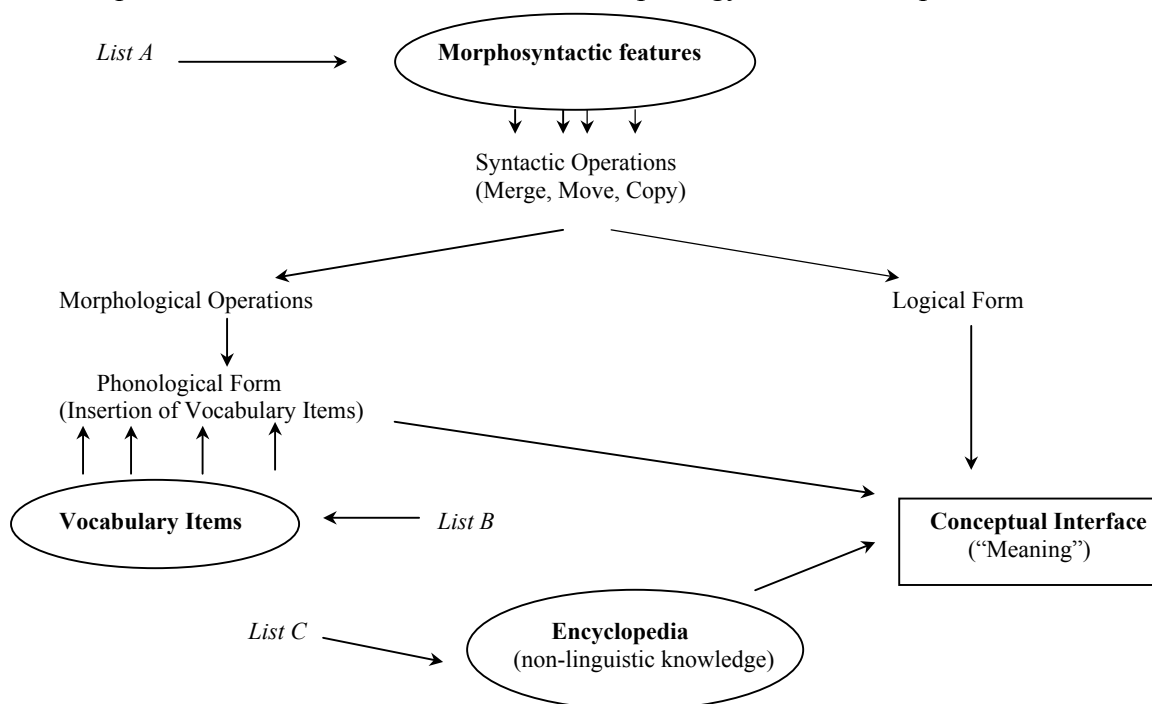

Figure 1. A DM-grammar (from Harley and Noyer 1999)

4   The proposal

First, I think it is important to restrict the domain of syntactic coreference to the sentence. Even though there is no limit to the length of sentences, other cognitive abilities, such as memory, impose constraints on sentence length. Such constraints are not imposed on discourse. Therefore I argue that the use of pronouns with an antecedent in the same sentence is very different from the use of pronouns that take their antecedent from a context outside their sentence.

In line with Zwart (2002), I would like to suggest a different categorization of DPs from Chomsky (1981). I would like to suggest that there are reasons for viewing PRO as part of the general category PRONOUNS[2]. If PRO is part of the root category pronoun, we have a tripartite division instead of a binary division as Zwart has. The consequences are that there can be no default choice for a pronoun if the pronoun lacks the feature [+coref.].

---

[2] It is not within the scope of this paper to elaborate on the reasons.

I also follow Kayne's analysis in the sense that coreference has to be established by merging of the antecedent and the pronoun/anaphor/PRO. Contrary to his analysis, I take this to hold at the sentence level. The implication of this is that the merging of antecedent and PRONOUN cannot be the only source for distinguishing between pronouns and anaphors. Another implication is that a pronoun does not need an antecedent merged in its specifier position in order to be well-formed, as Kayne suggests. The derivations of sentences involving the three types of pronouns would look as in (20).

(20)   *Syntactic derivation*   *Phonological form:*
   a.   Mary behaved [[Mary$_{copy}$] [PRONOUN]]   Mary behaved herself.
   b.   Mary said [[Mary$_{copy}$] [PRONOUN]] was tired.   Mary said she was tired
   c.   Mary tried [[Mary$_{copy}$] [PRONOUN]] to go.   Mary tried PRO to go.
   d.   Mary considered [[Mary$_{copy}$] [PRONOUN]] a genius.   Mary considered herself a genius.

The obvious question is then what it is that gives rise to the different morphological forms in the PF-component of the root pronoun. Based on Zwart's analysis my suggestion is that the pronoun picks up/checks features in syntax, thereby giving rise to different morphological words. As Zwart points out (2002:297), languages differ as to whether they have reflexive pronouns and how reflexivity is marked, i.e. by pronouns or by semantically bleached DPs representing for example body parts as in (3). In principle it should be possible to find a language that makes use of one morphological form to cover for the four cases in (20). A problematic language would be one where there are no principle C effects.

I assume that when the lexical array is assembled the lexical item PRONOUN is selected from the lexicon or List A in DM, with relevant case and Φ features. First, there is a distinction between PRO and anaphors and pronouns. This distinction is case. If we assume, in line with Hornstein (2001:167), that some ECM verbs are inserted in the lexical array with accusative case feature and others are inserted without them, we get the distinction between anaphors and PRO. This contrast is illustrated in (21). I will come back to the question why there has to be an anaphor, not a pronoun in ECM constructions.

(21) a.   Mary considers herself (*PRO) clever.
      b.   Mary tried PRO (*herself) to be clever.
      c.   Mary wanted herself/PRO to be clever
   d.   Mary washed herself/PRO.

In (21) the verb *consider* has accusative case that needs to be checked. Consequently the root pronoun must have accusative case features for the derivation to converge. The verb *try* on the other hand does not have accusative case features when it takes a clause as complement. The result is that PRONOUN must be sent to PF without case features. The verbs *want* and *wash* can optionally enter the array with or without accusative case features. When the derivation is sent to List B in spell-out a root pronoun, PRONOUN, without case features will get PRO as vocabulary item.

Let us next consider the distinction between pronouns and anaphors. Since the assumption is that both types are merged together with their antecedent, this merge alone is not enough to separate the two as in Zwart's analysis. Crucially we do not want to make use of a feature [+anaphor] since that would be a mere description without any explanatory force.

My suggestion for one contributing factor for the morphological distinction between pronouns and anaphors relies on the notion of phase (Chomsky 2000, 2001a,b). According to

Chomsky (ibid.) the phases are vP and CP. Following Chomsky (ibid.) the structure of a sentence is as in (22).

(22)
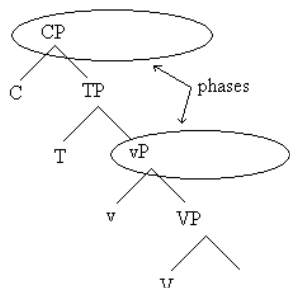
Once vP is formed, whatever is in the complement of v, i.e. VP, is sent to spell-out. Only the spec(s) of v and the head v are available in the next phase, This is called the phase impenetrability condition, PIC (Chomsky 2000:108). They will be sent to LF and PF when the next higher phase is merged, in (22), when C is merged to TP. Crucially the object has to move to spec-v to avoid being spelled out. There are reasons to believe that this is the case, at least in English (Chomsky 2001b (42a)), (23) and in Swedish (24).

(23)   (guess) what$_{Obj}$ John$_{Subj}$ T [$_{vP}$ t$_{Obj}$ [t$_{Subj}$ read t$_{Obj}$]]]
(24)   Johan såg henne   inte
       John saw her   not
       'John didn't see her'.

The suggestion is as follows. The fundamental factor determining if a PRONOUN will get the morphological form of an anaphor or a pronoun is whether the antecedent and the pronoun are sent to spell-out at the same time. If the object moves to the specifier position of vP it will not get spelled out until C is merged when the whole sentence is spelled, obviously that includes the antecedent.

4.1   Main clauses and embedded finite clauses

The derivation of a simple clause will look like (25).

(25)   [$_{CP}$ Mary likes [$_{vP}$ [[Mary$_{copy}$] PRONOUN] [$_{vP}$ …
       PF: Mary likes herself.

The antecedent *Mary* and the PRONOUN are spelled out at the same time and morphology will give the root pronoun the form of an anaphor.
    If there is an embedded clause as in (26), the PRONOUN and the antecedent will be spelled out at different phases. Consequently, lexical insertion will choose a pronoun for the PRONOUN.

(26)   [$_{CP}$ Mary$_i$ said [$_{CP}$ that [Mary$_{copy}$] PRONOUN$_i$] loves [$_{vP}$ John…
       PF: Mary said that she loves John

This also explains why we do not get an anaphor in sentences like (27).

(27) [CP Mary_i said [CP that John loves [vP PRONOUN_i ...
      PF: Mary said that John loves her/*herself.

The PRONOUN is spelled out in the lower CP-phase and the antecedent in the higher phase. This means that morphology accesses the pronoun without its antecedent and it will be spelled out as a pronoun. If all DPs in (27) are coreferential as in *Mary said that she likes herself* they will be merged together in the lowest argument position in some kind of embedded complex DP, where the second pronoun enters into a relation with the lowest one and the antecedent enters a relation with the second pronoun. This assumption is not very controversial since it usually is the head in a DP that enters into relations. The structure of a DP of that kind will look something like (28). The numbers are only for an expository reason and have no theoretical status.

(28) [DP [DP Mary] PRON_2.] PRON_1.]]

This structure is inserted in the object position of the verb *love* and PRONOUN_1 gets its theta role and case checked (29).

(29) loves [DP [DP Mary] PRON_2.] PRON_1.]]

For lack of a theta role PRONOUN_2 moves to the subject theta position of *love*. Since PRONOUN is a head (by assumption), the smallest constituent that can move is the complex DP [DP Mary] PRON_2.]. This DP moves up to check case of PRONOUN_2 and PRONOUN_1 moves up to the spec-position of vP. The DP *Mary* still does not have a theta role and must be available for the next phase so it moves to Spec-CP in the embedded clause (30).

(30) [CP Mary [TP PRONOUN_2 loves [vP PRONOUN_1 ...

Now everything in the complement of CP is sent to spell-out. This means that PRONOUN_1 and PRONOUN_2 will be spelled out at the same time. That means that PRONOUN_1 will get the morphological form of an anaphor. The exact reasons why it is PRONOUN_1 and not PRONOUN_2 that gets the form of an anaphor is still unclear. One plausible reason is that the antecedent values/checks the features of PRONOUN_2. PRONOUN_2 in turn values PRONOUN_1 and functions as the antecedent. *Mary* is spelled out at the final CP-phase (31).

(31) [CP ... Mary said [CP that...[PRONOUN_2 loves PRONOUN_1 ...
      PF: Mary said that she likes herself.

4.1.1   Several antecedents

So far we have looked at derivations where the antecedent is a DP in the singular. DPs in the plural are subject to the same derivation as above. Things get a little trickier when we have several different antecedents. Here we find a difference between anaphors and pronouns. Anaphors do not allow split antecedents, whereas pronouns do (32).

(32) Liza said that John drives them /*themselves.

Since the pronoun *them* can be coreferential with the two DPs they must have been merged together. Remember that the lexicon (list A) does not specify the form of the PRONOUN; the form is determined by the syntactic derivation. The crucial thing is how the antecedents are

merged to the root pronoun. It cannot be in the same way as in (28). Only PRONOUNS have a specifier position available for DPs[3]. The two DPs must occupy one specifier position each, unless they are coordinated. But if they are coordinated they can never split up (coordinated NP island). This is probably because of theta theoretical reasons, since they both get their theta role from the same verb. One DP, *John*, checks case and get a theta role from the verb, *drive*, in (32). The second DP, *Mary*, has to move to get a theta role and check case in the matrix clause. The consequence is that the two antecedents are in different phases and the PRONOUN must get the morphological form of a pronoun. A similar example is (33).

(33)   Mary told John that they had to behave themselves.

The difference is that the PRONOUN *they* gets its values from the two DPs. In parallel to the derivation of (30), it is the top PRONOUN that values the lowest one.

    The derivations of simplex main clauses and embedded finite clauses are pretty straightforward. In the next section we will look at raising and ECM-constructions, which are slightly more complicated.

4.2   Raising and ECM constructions

The problem with ECM constructions is that the subject in the embedded clause appears to be bound by the subject in the matrix clause. The problem with raising constructions is that the raised subject appears to be too far away to bind an object in the embedded clause. First let us look at the structure of an ECM construction (34) and a raising construction (35).

(34)   Mary considers herself (to be) smart.
(35)   Mary seems to admire herself.

I argue that the two constructions lack certain structure. It is generally accepted that the embedded clause in the ECM construction lacks at least a CP (see e.g. Chomsky 2001b:9), as in (36).

(36)   [$_{CP}$ [$_{TP}$ [$_{vP}$ [$_{VP}$ ([$_{TP}$)[$_{vP}$ [$_{VP}$]]]]]]]

The problem with the ECM construction is that there appears to be a phase, the circled vP, between the matrix clause and the embedded clause that will block spell-out of the antecedent in TP and the PRONOUN at the same time as in (37).

(37)   [$_{CP}$ Mary considers [$_{vP}$ [$_{vP}$ … herself …

Data from Swedish, (38), show that the object undergoes object shift and has moved up into the matrix clause.

(38)   Mary såg sig inte springa.
       Mary saw refl not run
       'She did not see herself run'.

---

[3]This is why there can be no coreference between referring DP and why a pronoun/anaphor cannot be coreferential with a DP it c-commands.

In (38) the subject of the embedded clause, *sig*, raises above the negation in the main clause. Suppose this position is (at least as high as) spec-vP. The result is that the PRONOUN will be spelled out together with the antecedent at CP in the matrix clause.

A problem which I do not have a solution to at the moment is how it is possible for the pronoun to bind a reflexive that is in the object position of the embedded verb as in (39).

(39)   ?Mary wanted herself to give herself a present
(40)   Mary wanted Liza$_i$ to give herself$_i$ a present.
(41)   Mary wanted Liza to give her a present.

From the data in (39-41), it appears that the subject in the embedded clause has a relation to both phases. If we assume that the ECM subject is spelled out at the higher clause it is a mystery how the object in the ECM can be an anaphor. If we say that the ECM subject is spelled out in the lower clause, as (40) and (41) indicate, it is a mystery why it is rendered as an anaphor in (39). The third way of looking at it would be to say that for unclear reasons all DPs are spelled out at the same time. This would explain (39), but not the fact that *herself* in (40) must be coreferential with *Liza* not *Mary*. I will leave this problem open for future investigation.

5   Conclusion

In this paper I have unified Kayne's and Zwart's derivational analyses of anaphors and pronouns. By using the notion of root pronoun and distributed morphology from Zwart and the notion of coreference from Kayne, I think the analysis I present has more explanatory value. The benefit of my analysis is that we get rid of the binding theory completely, the different forms on pronominal elements is only a matter of morphology. Crucially, there is never any need for an anaphor in syntax, which explains the fact that not all languages have anaphors. In addition the genitive reflexive can be analysed with exactly the same tools as other pronouns, namely, move, merge and the notion of phase. Tools that are independently motivated in a minimalist syntax. More importantly the difference between different languages boils down to differences in morphology, a most welcome result given the premises of the theory. Obviously this analysis is not without problems, problems it shares with other derivational approaches to binding and control. The requirements on movement have to be loosened and the distinction between internal and external merge (Chomsky 2001a) is unclear in these analyses.

References

Bobaljik, J. and S. Brown 1997. Interarboreal Operations: head movement and the extension requirement. *Linguistic Inquiry* 28: 345-356.

Chomsky, N. 1986. *Knowledge of Language*. New York: Praeger

Chomsky, N. 1995. The Minimalist Program. Cambridge, Mass.: MIT Press

Chomsky, N. 2000. Minimalist Inquiries: The Framework. *In Step by Step*, eds. R. Martin, D. Michaels & J. Uriagereka, 89-155. Cambridge, Mass.: MIT Press

Chomsky, N. 2001a. Beyond Explanatory Adequacy. MIT Occasional papers in linguistics 20

Chomsky, N. 2001b. Derivation by phase. *In Ken Hale: A life in language*, ed. M. Kenstowicz, 1-52. Cambridge, Mass.: MIT Press

Epstein et al. 1998. A Derivational Approach to Syntactic Relations. Oxford: Oxford University Press

Halle, M and A. Marantz. 1993. Distributed Morphology and the pieces of inflection. *In the view from Building 20*, eds. K. Hale & S.J. Keyser, 111-176. Cambridge, Mass.:MIT Press

Harley, H and R. Noyer. 1999. *Distributed Morphology*. Glot international 4:4 3-9.

Haspelmath, M. 1993. *A Grammar of Lezgian*. Berlin: Mouton de Gruyter

Hornstein, N. 2001 Move! A Minimalist Theory of Construal. Oxford: Blackwell

Kayne, R. 2002. *Pronouns and their antecedents* in Derivation and Explanation in the Minimalist Program. Eds. S. Epstein & D. Seely. Oxford: Blackwell

Nunes, J. 2001. Sideward movement. *Linguistic Inquiry* 32: 303-344.

Osumi, M. 1995. *Tinrin Grammar*. Honolulu: University of Hawaii Press

Rehg, K. 1981. *Ponapean Reference Grammar*. Honolulu: University of Hawaii Press

Zwart, J-W. 2002. *Issues relating to a derivational theory of binding* in Derivation and Explanation in the Minimalist Program. Eds. S. Epstein & D. Seely. Oxford: Blackwell