WHEN BINDING CAN USE AGREE: COMMENTS ON PREMINGER 2019a

Eric Reuland
Utrecht institute of Linguistics OTS
e.reuland@uu.nl

Preminger (2019a) argues that the Anaphor Agreement Effect (AAE) (Rizzi 1990, Woolford 1999) shows that what he refers to as binding-as-agreement is untenable. Preminger uses the definition of the AAE in (1), his (9):

(1) **Anaphor Agreement Effect**

\[
* H^0 \ldots \text{DP}_{\text{anaph}}, \text{where } R \text{ is a nontrivial } \varphi\text{-agreement relation}
\]

Given the aim to provide evidence against binding-as-agreement it is important to be precise as to which particular approaches it is actually an argument against. Most of the time it seems that his argument goes against approaches claiming that all binding proceeds by the syntactic Agree operation. Let’s call this the full reductionist approach. I am not sure anyone ever claimed that. Kratzer (2009), for instance, did not, so it seems to me more a straw man than an approach that has been actually defended.

What is confusing in the exposition is that it refers to proposals such as Reuland (2011) as if they were full reductionist. However, contra to what Preminger appears to suggest (p.1), Reuland (2011) never proposed a reduction of binding per se to \(\varphi\)-agreement. In fact, his analysis is crucially based on a distinction between the two. Nor does Reuland (2011) propose a general reduction of anaphor binding to \(\varphi\)-agreement. In his analysis binding by Agree is limited to SE-anaphors, and for complex anaphors consisting of a SE-anaphor and a SELF-type element for those binding relations based on the SE-anaphor component. The additional elements of complex anaphors have their own contribution to establishing the dependency. In this approach the complex binding patterns observed result from the interaction between different modules of the language system. Let’s call this the modular approach.

So, the following questions arise: i) Do Preminger’s arguments show that any approach in which binding of SE-anaphors is syntactically encoded by Agree is untenable? ii) What does Preminger’s own proposal achieve?

To my mind the arguments against the full reductionist approach are convincing, including the ‘final point’ in section 3, that "\(\varphi\) -feature matching is neither a necessary nor sufficient condition for coreference" and the further considerations in sections 4 and 5. Preminger’s alternative is largely underspecified, though. His premise is that the AAE is a unified phenomenon. I will argue that this cannot be maintained. Upon proper consideration, it seems, then, that the facts he discusses favor a modular approach to binding and to the AAE as well.

---

1 I would like to thank Omer Preminger and Peter Zubkov for their very helpful comments. Preminger (2019) is a revised and rewritten version of a draft I was
1. Preminger's arguments

1.1 General arguments

Preminger claims that "there is plenty of evidence unrelated to the AAE suggesting that reductionist approaches are off track" and mentions two of them. One is an argument from directionality. Preminger argues that it is well established that $\phi$-agreement transmits values upward in the structure: from a c-commanded goal to a c-commanding probe. Anaphoric binding, in contrast, appears to transmit values from a c-commanding antecedent to a c-commanded bindee. This leads him to the statement in (2):

(2) The idea that anaphoric binding is underpinned by $\phi$-agreement is antithetical to the apparent structural properties of the two relations.

This criticism, however, does not apply to the approach in Reuland (2011), as Preminger acknowledges in a footnote. The same applies to a recent approach by Zubkov (2018) (see the appendix for a brief overview). But, then, as an argument against a general type of approach, the argument lacks force.2

Another source of evidence against reductionist theories, Preminger argues, concerns the ban on morpho-phonologically null agreement (Preminger 2019a), saying that "It is a truism, however, that anaphoric binding exists even in languages that lack overt $\phi$-agreement—like Japanese—which, by hypothesis, lack syntactic $\phi$-agreement as well. Japanese, after all, does have anaphors. Thus, the idea that anaphoric binding is underpinned by $\phi$-agreement is antithetical to the ban on null agreement, which is required for the AAE to even be statable with any generality." But note that this argument is significantly weakened by the observation by a reviewer, which Preminger reports in a footnote, namely that there are languages that lack morpho-phonologically overt $\phi$-agreement and yet manifest AAE or AAE-like effects (e.g. Mainland Scandinavian; Woolford 1999:283n31). The fact that "a full retreat from the ban on null agreement" would land us "back in the conundrum [...] concerning why the AAE applies in, e.g., Icelandic but not in, e.g., Japanese" is not an argument.

Even so, the argument based on Japanese is only relevant as an argument against the position that all binding is done by syntactic $\phi$-agreement, but not against a position that limits binding by Agree to SE-anaphors and their kin (unless Japanese zibun would be shown to be a SE-anaphor in all other relevant respects).

In an earlier version of this paper Preminger noted that the inferences he draws only go through if 'anaphor' is a natural class. If not, the AAE would be an epiphenomenon. I agree. The working definition he provided there was for good reasons omitted from the revised version. The resulting version, however, does not provide any definition of anaphor. So, that anaphor is a natural class is now silently presupposed. But the assumption itself is just wrong. The interpretively dependent elements that are commonly called anaphors are subject to rather varying conditions. Some always require a local antecedent, like Dutch zichzelf; others allow a non-local antecedent, like Russian sebja; others may allow a non-local antecedent depending on the environment, like English himself, or

---

2 In fact it does not apply to the approach by Rooryck and Vanden Wyngaerd (2011) either.
sometimes don’t need a linguistic antecedent at all, like Icelandic sig, or Indonesian dirinya, and yet others are just licensers of a reflexive interpretation, like Indonesian diri (Kartono 2013, 2019).

Reuland (2011: 239) concludes that the notion of an anaphor as it is used in the CBT lacks theoretical significance. There is no intrinsic property that can characterize elements as anaphors. What can be reconstructed is the relative property of being used as an anaphor as in (3):

(3) A particular element is used as an anaphor iff it is linked to its antecedent by a syntactic operation.

This derives one property traditionally taken to distinguishing anaphors from pronominals, namely that anaphors don’t allow split antecedents, but pronominals do. This property, however only obtains when (3) is met. For instance for English himself it only applies when it is syntactically bound by SELF-movement (“used as an anaphor”), not when it is in exempt position.

So far, there is every reason to assume, as argued originally in Reinhart and Reuland (1993), and subsequently in Reuland (2011) and more recently in Reuland et al. (2019), that there are different types of processes involved in the interpretation of what the tradition qualifies as anaphors (minimally four, such as chain formation, lexical reflexivization, syntactic reflexivization, and assigning a value in discourse). Consider then the AAE again:

(4)

\[
\text{Anaphor Agreement Effect} \\
* H^0 \ldots \text{DP}_{\text{ANAPH}}, \text{where } \mathcal{R} \text{ is a nontrivial } \varphi\text{-agreement relation}
\]

Whereas the original formulation by Rizzi forbade for an anaphor to enter any agreement relation, the formulation here restricts its application to non-trivial \(\varphi\)-agreement relations. This is to accommodate the fact that some languages (such as Georgian or Modern Greek) allow a complex anaphor in this position with default agreement. But languages like Icelandic with a SE-anaphor have a full ban.\(^3\) Hence, on closer examination the AAE in Preminger’s formulation is less unified than one might initially think.

Thus, the AAE, in so far as it applies to very different types of elements may well result from different interactions, although still reflecting a universal requirement of agreement.

---

\(^3\) Unlike what Preminger argues, Albanian vetja is not a simplex anaphor. It is morpho-syntactically a noun, optionally modified by an overt possessive, not much unlike in Georgian, Basque or Greek. This is in line with Franks (2013) who writes that “Albanian has a simple reflexive, either vetë/vetja or vetvetë/vetvetja ‘self’, which declines like any other feminine noun and shows the same definite/indefinite morphological behavior as normal nouns (definite is used for argument positions, indefinite is triggered by accusative-taking prepositions unless there is a modifier attached to the reflexive noun).”
So, Preminger’s general arguments against binding by agree don’t apply to the approach in Reuland (2011). Preminger’s general line, which uses the universal validity of the AAE as a premise and requires it to have a universal explanation, is either doubtful, or even in its general form unfalsifiable in the absence of a definition of anaphor.

1.2 Binding by Agree and the AAE.
Despite these general considerations about the status of the AAE, interesting empirical issues may arise in particular languages that show an AAE.

Preminger gives an interesting and detailed analysis of the AAE in Basque, but the general line is quite simple and general. In a nutshell, Preminger argues that Binding by Agree gives rise to a timing problem. If the anaphor is valued by Agree on time for the dependency to be seen at the point where compliance with the AAE to be checked and or it is handed over to the to the interpretation system, then, at the same time, the anaphor will be fully valued, spelled out as a pronominal. As such it should no longer be able to trigger the AAE. But it is.

In fact this reflects a more general issue: If Agree values a φ-deficient expression why does this expression show up as φ-deficient at all? (See Reuland 2010 and Rooryck and Vanden Wyngaerd 2011 for discussion from different perspectives.) The question is, then, whether or not a deficient φ-feature bundle will retain some characteristic of an anaphor after valuation. This is precisely what the approach proposed by Zubkov (2018) and Reuland and Zubkov (2019/2020) as summarized in the appendix achieves:

(5) A binding dependency is established with a feature chain based on a single probe, Person or #. Hence φ-deficient goal (anaphor) may end up being valued for Person or # but not for both.

Such an element will still be deficient after binding. It is distinct from a pronominal and will spelled-out by a different vocabulary item. As such it will also be visible for whatever principle that underlies the AAE in its relevant manifestation.

All in all it seems fair to conclude that Preminger’s arguments against full reductionism don’t carry over to a modular approach.

In section 9 of his contribution Preminger presents an analysis of the AAE based on the notion of encapsulation, which I will discuss in the next section.

3 Encapsulation and reductionism
As I noted earlier, I fully agree with Preminger that a strong version of binding by Agree is untenable. Here, I will just briefly evaluate his approach to the AAE in terms of encapsulation.

As is extensively discussed in, for instance, Reinhart and Reuland (1993) and subsequent work Reuland (2011, 2017) and reference cited there, there is a pervasive contrast between simplex and complex anaphors, not only in form, but also in contribution to interpretation. Complex anaphors generally consist of a φ-feature bundle (pronominal or SE-anaphor) as a core with one or more additional components (intensifier, body-part expression, etc.) attached.

Preminger proposes, along the lines of Middleton (2018), that anaphors universally have the structure in (6), his (1):
AnaphP stands for the outermost layer for the anaphor; P stands for the projection within the anaphor that hosts valued $\varphi$-features. Preminger proposes that 'anaphoricity' is encoded in the outermost layer of this expression, but keeps silent on how precisely binding is effected and – where applicable - enforced.4

Preminger suggests that the AAE arises because $\varphi$-agreement comes upon a structural layer (AnaphP) that prevents access to the feature-values hosted on the $\varphi$-bearing portion of the reflexive (P). Thus, the AAE arises since the bearer of $\varphi$-features is contained in larger structure that is opaque to

---

4 On p. 25 Preminger provides an argument that in English anaphoricity is encoded on the outermost layer of a reflexive anaphor (it bears' the relevant binding index'). He argues that under the 'reductionist' position the outermost layer of a reflexive anaphor should behave as if it is not the bearer of the relevant binding index. He then considers the case in (i) (his (53a–b), attributed to [Norvin Richards, p.c.]):

(i)  
   a. John expects Mary to outdo him
   b. John expects himself to outdo him.

Indeed (ia) does not give rise to a disjoint-reference effect between John and the pronoun him. As he argues, the cause of the disjoint-reference effect observed in (ib) must be the anaphor himself. But this could only be the case if the binding index resided on the outermost projection on the anaphor. However, as may happen with straw-man arguments it is problematic if applied to actual theories. For instance, (ib) independently violates the chain condition in Reinhart and Reuland (1993). However, Reinhart and Reuland (1993: 118) provide a pattern from Dutch that does support this particular instance of Preminger’s claim:

(ii)  
   a. *zich zich critiseren.
   b. zichzelf zichcritiseren.
   c. zelf *[self] SE criticize

Here the SELF element licenses the reflexive interpretation both in object and subject position. So, if we equate 'licensing reflexivity' with 'bearing a binding index' this might be a relevant example.
probing. In fact I agree that such a configuration may well be what underlies the intricate pattern observed in Basque, which is based on a possessive structure, and may carry over to languages with a similar structure. However, Preminger does not offer a real proposal as to what precisely makes AnaphP opaque. In order to make the argument convincing, more than a stipulation that AnaphP might be phasal is needed.

It is indeed uncontroversial that languages have 'anaphors' of varying degrees of complexity. Indonesian, with diri, dirinya and dirinya sendiri is a case in point (see Schadler 2014). And one might argue, like Middleton and Preminger would, that the pair dirinya and dirinya sendiri nicely fits the pattern in (6) since dirinya sendiri, which is local is clearly dirinya plus something. That one finds such patterns is not really surprising from the perspective of Reuland, Wong and Everaert (2019) who argue for a distinction between simplex, complex and 'super complex' anaphors on the basis of the anaphoric system of Mandarin (see also Volkova 2014, 2017 on Meadow Mari, and Rudnev 2017 on Avar). In fact these all require a further layer and each of the layers contributes a particular aspect of the interpretation. However, it is important to note that the correspondence with the structure in (6) is far from straightforward. The φ-component of dirinya sendiri is the possessive component of nya of dirinya. This necessitates one further degree of complexity than envisaged in (6). There is no evidence that the diri part of dirinya realizes the φ-component. In fact it has been shown that diri is not an argument, even though it seems to appear in argument position (Kartono 2013, 2019). The element sendiri in the outermost layer is best characterized as being a reflexivizer rather than just contributing anaphoricity. All in all, the structure in (6) presents an overly simplified picture of a more complex reality.

There is an even more crucial problem. The structure in (6) implies that universally the structure of 'anaphors' is complex, in fact more complex than the structure of pronominals. However, there is not the slightest evidence that SE-anaphors (Icelandic sig, Dutch zich, etc.), or for instance reflexive clitics in Romance, have such a complex structure. The same holds true of possessive anaphors, such as Norwegian sin or Russian svoj. 6

---

5 See for instance the German paradigm mich 'me', dich 'you', sich 'SE', or the corresponding Romance clitics me, te, se.

6 Basing himself on Middleton (2018), Preminger (p.c.) argues that there is fact such evidence. Middleton argues for a distinction between pronouns, diaphors (elements taking long-distance antecedents) and anaphors, and a cross-linguistic ban on an ABA form of syncretism, a syncretism between anaphors and pronouns where diaphors retain a separate form, which is argued to hold without exception. Preminger argues that if SE anaphors are not internally complex in the manner suggested by Middleton, the explanation for the ban on discontinuous syncretism is lost. However, this argument loses force if SE-anaphors are just φ-deficient pronouns, and bound through a different mechanism than complex anaphors. In a nutshell, if diaphors require marking, also local anaphors do. Preminger rightly notes that one should be wary of reasoning from morphological complexity or lack thereof, to syntactic complexity or lack thereof. As noted in Preminger (2019b), interface
It is well-known that in many languages SE-anaphors, and reflexive clitics serve not only as markers of reflexivity, but also mark reciprocals, middles or passives. All such facts would be quite hard to reconcile with a complexity as in (6).

Preminger further claims that the $\varphi$-feature matching we see in anaphor binding cannot be used as an argument in favor of $\varphi$-features being involved in anaphor binding, since matching also occurs in environments where anaphor binding plays no role. So, he argues, we need it anyway. This, however, assumes that $\varphi$-feature matching is uniform. The difference in matching requirements between local and 'not so local' binding relations between the expressions nós and a gente in Brazilian Portuguese shows that it is not. Both pronouns are interpreted as 1st person plural. The former is syntactically 1st person plural indeed, the latter is syntactically 3rd person singular. The former can bind the latter and vice versa in not-so local environments. However, a strict match in morphosyntactic features is required in local environments, coinciding with the domain of syntactic chain formation, see Menuzzi (1999) and Reuland (2011).

4 By way of conclusion.

In general, Preminger’s arguments against full reductionism are sound. However it is not clear that this is more than a straw man. The timing of valuing and visibility for the AAE/spell-out is a potential problem for an approach that relates binding of $\varphi$-deficient anaphors to $\varphi$-feature sharing with an antecedent, but only in so far as ‘anaphoricity’ is not preserved. In an approach along the lines of Zubkov (2018) this issue is resolved.

While there is evidence for an outer layer in complex anaphors, there is none for an outer layer in SE-anaphors such as Icelandic sig, Dutch zich, etc. Consequently, Preminger’s explanation of the AAE in Basque cannot be generalized to all cases of the AAE.

Furthermore, while in Preminger’s analysis the anaphoric property of the outer layer is just stipulated, in Reinhart and Reuland (1991), Reuland (2011), etc. the reflexivizing property of SELF-type elements in the outer layer is derived from more elementary properties.

transparency is an important methodological heuristic, not a law. So, clearly, establishing a relation between morphological complexity of complex anaphors and the complexity of their syntactic structure, or between the morphological non-complexity of SE-type anaphors and the simplicity of their syntactic structure cannot be done without a theory of what the element contributing complexity in fact does, both syntactically and semantically, and why such an element must be absent in SE-anaphors. Such a theory has been provided in Reinhart and Reuland (1993) and much subsequent work. This does not deny that also prima facie simplex elements may have more structure than meets the eye. This applies to pronominals as well as to SE-anaphors. However, there is no reason whatsoever to assume that Dutch zich and Dutch zichzelf share the structure – the outer layer in (6) - that makes zichzelf a local anaphor /reflexivizer, since this is precisely what zich, although occurring as an anaphor is not.
All in all, Preminger’s arguments don’t apply to modular approaches to binding, such as Reuland (2011), Zubkov (2018) and subsequent work.

**Appendix Binding by Multiple Agree**

Zubkov (2018) provides an Agree-based analysis of the argumental anaphor *sebja* and the possessive anaphor *svoj* in Russian (see also Reuland and Zubkov, 2019/2020). Both can be bound locally and non-locally (roughly with a ‘subject’ intervening). Crucially, as Zubkov shows, non-local binding goes with an animacy effect; that is, the antecedent must be animate, in fact aware of the dependency involved.

Let me briefly summarize. Zubkov’s analysis is based on Multiple Agree (Hiraiwa 2001, 2005, see also Boeckx 2003, Chomsky 2008) with dependencies expressed via probe-goal relations as in (7):

(7) **Probes and goals**
- nodes on the verbal or nominal spine representing unvalued Person and Number serve as probes
- anaphors (*sebja, svoj*) are inserted into the derivation with unvalued Person (uP) and unvalued Number (u#) and serve as goals

**Positions of the probes:**
- A High probe with unvalued Person (P) and Number (#) is located in the C-domain
  - A High probe can only be valued by a Nominative (‘privileged’) NP-goal
    (only the Nominative is eligible)
- Low probes with unvalued # are located in:
  - a position in the extended V-projection, with the vP in its domain
  - within the NP in a position below the possessor, but with all thematic arguments, including the ‘author’ – POSS phrase, in its domain.

**Valuation**
- a probe can value a multiplicity of goals (but not vice versa)
- a probe is valued by the nearest eligible goal in its c-command domain
- a goal is valued by the nearest probe

Informally: Once a probe is valued this value spreads to goals (anaphors) in the probe’s c-command domain, subject to a minimality restriction:

(8) A goal can be valued by a probe across another probe only if the probes differ in features → A Person probe can value a goal across Number

---

7 The status of features and the notion of valuation raise interesting problems, which go beyond our present concerns. See Reuland (2020) for discussion.

8 In order to capture the fact that in languages like Icelandic also NP’s with quirky case are able to binding anaphors, the notion of ‘privileged’ goal will have to be generalized. A more detailed proposal will have to wait for further research.
probes (different → no intervention) → non-local binding is enabled. A Number probe cannot value a goal across a Number probe.

- Chain formation takes place in line with Pesetsky and Torrego (2007)
- Person is interpreted as reflecting animacy and awareness.

An important feature of the system is that, in contrast to Reuland (2011) and subsequent work, a binding dependency is established with a feature chain based on a single probe. Due to the distribution of the probes, and the sensitivity to intervention an anaphor may end up being valued for Person or for Number but not for both.

Given the role of single-probe chains, as discussed in Reuland and Zubkov (2019/2020), the chain condition of Reuland (2011) is modified as follows:

(9) **Relativized Chain Condition (RCC)** - based on feature chains involving a single probe:
A pronominal with a fully specified φ-feature bundle cannot be bound by an antecedent if the latter (or other goals referentially indistinct from it):
values all probes that can attempt to value the former
- closest number, closest person
- If the goals that value the probes are positionally distinct but referentially identified with each other, complementarity is preserved.
- If any of the probes diverge, the pronominal is not excluded and complementarity breaks down.

This allows the system to explain quite intricate patterns of complementarity, such as the sensitivity of complementarity to animacy of the antecedent.9

Zubkov’s approach has been worked out in detail for Russian. If extended to, for instance, Germanic, it would entail that also SE-anaphors in Germanic enter a single-feature based chain, leaving one feature unvalued. Although more detailed work would be needed to establish this more firmly, prima facie its basic idea seems quite generalizable. For the benefit of the reader I include some sample derivations for Russian.

**Some sample derivations**
Consider the case in (10) with an animate subject Vanja:

(10) Vanja protivorečit sebe
     'V contradicts himself'

The relevant structure is given in (11):

(11)

---

9 Of course, an <uval P, uval #> element IN the privileged position will not lead to valuation of the probes.
Vanja protivorečit sebe

The unvalued Person probe in the left periphery probes and finds a privileged nominative goal *Vanja*. *Vanja* values the high Person probe, and simultaneously values the unvalued person feature of *sebe*.

As is important to note, *sebe*’s # feature remains unvalued. As we will see in the next derivation, the alternative is binding by #-sharing. In this case the Person feature remains unvalued. Thus, in all cases of anaphor binding in Russian, either # or Person will stay unvalued. This entails that after binding, anaphors will remain distinct from pronominals.

The alternative route of valuing and binding by # is illustrated in (12)
For the low unvalued # probe Vanja, which happens to be nominative (but which is not essential here) is the nearest goal. Hence the low # probe is valued by Vanja. The result is that sebe’s # feature is valued, and sebe ends up bound by Vanja.

Thus, with animate antecedents there are two routes toward binding of an object anaphor in local configurations. However, if the antecedent is inanimate, as in (13), binding is only possible via the # probe, since the Person probe is incompatible with inanimacy.

(13) Èta kniga protivorečit sebe
‘That book contradicts itself’

This gives rise to interesting animacy effects that are discussed in Zubkov (2018) and Reuland and Zubkov (2019; in preparation), to which we refer.

Thus, crucially, even after valuation the φ-feature structure will not match the insertion conditions for the full pronominal. Consequently, the special form sebja will be selected. It is easily seen that this derivation does not distinguish between 1st and 2nd person antecedents on the one hand and 3rd person antecedents on the other. An anaphor bound by 1st person singular ja ’I’ or 2nd person plural vy ’you’ will also end up valued only partially, and hence fit sebja/sebe.

References


Reuland, Eric, Sally Wong and Martin Everaert (2019). Mandarin zi-ji is complex. Linguistic Inquiry, corrected proof https://doi.org/10.1162/ling_a_00355


