Condition B effects in two simple steps

Roelofsen, F.

Link to publication

Condition B effects in two simple steps

Floris Roelofsen

Abstract This paper is concerned with constraints on the interpretation of pronominal anaphora, in particular Condition B effects. It aims to contribute to a particular approach, initiated by Reinhart (Anaphora and semantic interpretation, 1983) and further developed elsewhere. It proposes a modification of Reinhart's Interface Rule, and argues that the resulting theory compares favorably with others, while being compatible with independently motivated general hypotheses about the interaction between different interpretive mechanisms.

1 Introduction

The interpretation of anaphoric pronouns is constrained by syntactic structure. One particular approach to capturing such constraints was initiated by Reinhart (1983). Her main thesis was that only one type of anaphoric relation, *binding*, is syntactically encoded and subject to rules of syntactic well-formedness. She assumed that other anaphoric relations, such as *coreference*, are not syntactically encoded and are therefore not regulated by rules of grammar. Instead, she suggested, there is a pragmatic principle which implies that coreference is ruled out whenever it yields exactly the same interpretation as binding.

This approach has been further developed by Heim (1998), Fox (2000), Büring (2005b), Reinhart (2006), and Reuland (2008). The aim of the present paper is to evaluate and refine these analyses. In particular, some previously unnoticed problems for the analyses of Büring (2005b) and Reinhart (2006) will be pointed out,

F. Roelofsen (✉)
ILLC, University of Amsterdam, P.O. Box 94242, 1090 GE Amsterdam, The Netherlands
e-mail: floris.roelofsen@gmail.com

and a modification of the latter will be proposed. The resulting theory will be argued to have certain advantages over the analyses of Heim and Büring. In particular, it avoids certain controversial assumptions about VP ellipsis that Büring's account requires, and it avoids the assumption that anaphoric relations other than binding are syntactically represented and constrained, which is what Heim's account requires.

The idea behind the proposed account can be seen as a particular instance of the more general hypothesis that 'hearers minimize interpretive options' (Reinhart 2006) or of the closely related idea that 'rejection of a certain interpretation by one component of the grammar cannot be overruled by other components' (Reuland 2008). To the extent that the account is empirically correct, it provides support for both these general hypotheses.

2 Reinhart's original approach

The main thesis of Reinhart (1983) is that binding relations are encoded in the syntax and subject to *grammatical* constraints, whereas other kinds of anaphoric relations, such as coreference, are not encoded in the syntax, and therefore not subject to grammatical constraints. Rather, coreference is established contextually, and restrictions on coreference are of a *pragmatic* nature.

In order to discuss Reinhart's proposal in more detail, we must first fix some formal terminology and notation, as well as some basic assumptions about the syntax-semantics interface.

2.1 Basic terminology, notation, and assumptions

The most straightforward way to encode binding relations in the syntax is by means of indices. Let us assume, then, that pronouns enter a syntactic derivation either with or without an index. If a pronoun comes with an index, we will call that index its *binding index*, and we will append the index to the pronoun in subscript (e.g., $[\text{him}_1]$). Pronouns with a binding index will be treated as bound by some other determiner phrase; pronouns without a binding index will be treated as referential.

I will assume that determiner phrases may undergo wh-movement or quantifier raising. If a DP undergoes wh-movement or QR, it receives a *binder index* n, which is adjoined to it in superscript (e.g., $[\text{who}]^3$). It also leaves behind a trace which has that same index n as its binding index (e.g., the trace of $[\text{who}]^3$ would be $[t_3]$).

(1) $[_{TP} \text{ X } [_{DP} \text{ W}] \text{ Y}] \Rightarrow [_{TP} [_{DP} \text{ W}]^n [_{TP} \text{ X } t_n \text{ Y}]]$ (wh-movement)

(2) $[_{TP} \text{ X } [_{DP} \text{ Q}] \text{ Y}] \Rightarrow [_{TP} [_{DP} \text{ Q}]^n [_{TP} \text{ X } t_n \text{ Y}]]$ (quantifier raising)

I will assume a two-stage Montagovian syntax-semantics interface: logical forms are first compositionally mapped to expressions in some type-theoretical language, and these expressions are in turn interpreted model-theoretically. A pronoun or a trace with a binding index n is interpreted as a variable x_n, and a constituent of the form $X^n Y$ is interpreted as:

(3) $X'(\lambda x_n \cdot Y')$

where X′ is the interpretation of X and Y′ is the interpretation of Y. This composition rule embodies what Heim and Kratzer (1998) call *predicate abstraction*. By way of illustration, the logical form in (4) is mapped to the type-theoretical expression in (5) (where I use somewhat sloppy notation):

(4) $[\text{John}]^1[t_1 \text{ loves his}_1 \text{ mother}]$

(5)   JOHN $(\lambda x_1 \cdot x_1 \text{ LOVES } x_1\text{'S MOTHER})$

We will say that a determiner phrase X *binds* a pronoun P iff (i) X c-commands P, (ii) X's binder index coincides with P's binding index, and (iii) X does not c-command any other DP that satisfies (i) and (ii). This notion of binding is what Heim and Kratzer (1998) and Büring (2005a) call *semantic binding* and what Reinhart (2006) calls *A-binding*. Notice that, according to this notion, [John] binds [his] in (4).

   As for referential pronouns, i.e. pronouns without a binding index, I will make a slightly non-standard assumption. I will assume that the formal language contains, besides standard individual constants and variables, a set of individual *proterms*. Moreover, formal expressions are interpreted not only relative to a model and an assignment function, but also relative to a *resolution function*. The interpretation of individual constants is determined by the interpretation function, which is given by the model; variables are interpreted by the assignment function, and proterms are interpreted by the resolution function. If a pronoun does not have a binding index, it is mapped to a proterm, and thus interpreted by the resolution function, which assigns it a contextually salient individual concept.[1]

   Finally, we will say that two referential determiner phrases A and B *corefer* in a particular utterance context C (which furnishes a particular resolution function) iff A and B denote the same individual in every world that is consistent with the speech participants' common assumptions in C. This notion of coreference is often called *presupposed coreference* (cf. Büring 2005a, p. 153).

   If a referential pronoun P is supposed to be resolved in such a way that it corefers with a determiner phrase A, I will write P = A next to the LF in question. For example, I will write:

---

[1] This setup deviates from that of Heim and Kratzer (1998), which is often taken as a starting point. In particular, referential pronouns are not treated as free variables here. Bound and referential pronouns are really dealt with by two separate mechanisms. The main reason for me to take this line here is that it seems to be the most straightforward way of implementing Reinhart's distinction. Incidentally, setting things up in this way has the additional advantage of not having to stipulate that "no logical form may contain both bound occurrences and free occurrences of the same index" (Heim and Kratzer 1998, p. 254). Having said this, it should also be noted that the assumed treatment of referential pronouns, although it will serve our purposes here, becomes problematic as soon as we take 'discourse reference' into account. One general treatment of referential pronouns that is, as far as I can see, compatible with what will be said below, is the one proposed by van der Sandt (1992) and Geurts (1999). However, spelling out the details of such a treatment would bring along several issues that are orthogonal to the ones to be discussed here. Therefore, I adopt the 'toy'-treatment specified above.

(6)   [John]^1 [t$_1$ loves his mother]   his = John

to indicate that the referential pronoun [his] is supposed to be resolved in such a way that it corefers with [John].

2.2 Condition B and the Coreference Rule

We are now ready to spell out Reinhart's proposal in more detail. We will focus on her analysis of Condition B effects. The relevant empirical generalization is that pronouns can normally not be interpreted as bound by—or coreferential with—coarguments.[2] For instance, binding is impossible in the following constructions:

(7)   a.   Every girl loves her.
      b.   Every woman believes her to be a great dancer.
      c.   Mary asked every boy to wash him.

Coreference is equally impossible in the following examples:

(8)   a.   Susan loves her.   ⇒ her ≠ Susan
      b.   Norah believes her to be a great dancer.   ⇒ her ≠ Norah
      c.   Mary asked John to wash him.   ⇒ him ≠ John

However, Reinhart (1983, p. 169) notes that there are at least two kinds of environments in which coreference seems to evade Condition B effects. The first kind of environment involves focus-sensitive operators like *only*. For instance, according to Reinhart, [him] can be interpreted as coreferential with [Max himself] in (9).

(9)   Only Max himself voted for him.

The second kind of environment is one in which, in order to establish appropriate 'contrast' or 'parallelism' between two phrases, a pronoun must be interpreted as coreferential with one of its coarguments:[3,4]

(10)   I know what John and Mary have in common.
       John hates Mary and Mary hates her too.

(11)   If everyone voted for Oscar, then certainly Oscar voted for him.

---

[2] Several versions of this generalization have been proposed in the literature. The present formulation, adapted from (Büring 2005a, pp. 55–56), is relatively theory-neutral and covers the relevant data. Two determiner phrases are coarguments iff their θ-role and/or case are assigned by the same predicate.

[3] Very similar examples were used by Evans (1980) to show that Condition C effects are suppressed in certain environments. Therefore, examples like (10) and (11) are often attributed to Evans, rather than to Reinhart. I will likewise refer to (10) and (11) as Evans' examples from now on.

[4] Examples of this kind occur particularly frequently with first and second person pronouns. Horn (2008, pp. 173–174) provides a list of attested examples.

I must note here that the judgments of my informants do not always confirm those of Reinhart. Many speakers find that coreference is marginal in (9), (10), and (11), and emphasize that there are certainly more natural ways to convey the intended messages. In the recent literature, several authors have pointed out or acknowledged the controversial status of these cases (cf. Schlenker 2005; Jacobson 2007; Grodzinsky 2007; Heim 2007). The theories to be discussed here, however, are in large part designed especially to deal with these cases. Thus, for the purposes of this paper, I will proceed as if Reinhart's observations are undisputed.[5]

So much for the data. Let us now turn to Reinhart's theoretical proposal. The first ingredient of her theory is a syntactic constraint, which only applies to binding:[6]

Def. 1: Condition B

Pronouns cannot be bound by their coarguments.

Besides this syntactic constraint on binding, Reinhart argues that there is a pragmatic constraint on coreference. The rationale behind this constraint is that it is always more risky for a speaker to use syntactic structures which contain referential elements than to use syntactic structures in which anaphoric relations are directly encoded. This is because referential elements always have to be resolved by the hearer, and this can go wrong. If anaphoric relations are syntactically encoded, resolution does not come into play. Reinhart assumes that speakers generally want to avoid any risk of being misinterpreted, and thus always prefer to use syntactic structures which contain *bound* anaphoric elements rather than syntactic structures which contain *referential* anaphoric elements. Only if speakers cannot express the intended meaning using bound anaphora will they use referential elements. This idea can be implemented as follows:

Def. 2: Coreference Rule

A speaker will never use a logical form LF in a context C if LF is semantically indistinguishable from one of its binding alternatives in C.

---

[5] The following example, adapted from Heim (1998), is sometimes taken to instantiate a third set of examples in which coreference is insensitive to Condition B effects.

(i) How can you doubt that the speaker is Zelda? She praises her to the sky.

However, as Heim notes, even though *she* and *her* may be intended to refer to the same person, they are intended to do so *through different guises*. Technically, they are assigned two distinct individual concepts, which may be instantiated by the same individual in the actual world but by distinct individuals in other worlds consistent with the speech participants' common assumptions. Thus, technically speaking, coreference does not obtain here, and the question whether or not it is subject to Condition B effects does not arise.

[6] There is an ongoing debate in the literature about how this constraint should be formulated exactly (cf. Pollard and Sag 1992; Reinhart and Reuland 1993; Büring 2005a), and about whether it can be derived from more general syntactic principles (cf. Reuland 2001). This debate is interesting in its own right, but orthogonal to the discussion here. The present formulation is adapted from Büring (2005a, pp. 55–56).

Def. 3 Binding Alternatives

Let C be a context, let LF be a logical form, and let A and B be two determiner phrases in LF, such that A and B corefer in C and such that A c-commands B in LF. Then the structure obtained from LF by:

(i)   quantifier raising A in case it has not been raised yet, and
(ii) replacing B with a (possibly reflexive) pronoun bound by A

is called a *binding alternative* of LF in C.

Let us see how the Coreference Rule captures the relevant data. First consider an environment where coargument coreference is not available:

(12)   Max voted for him.
       a. $[\text{Max}]^1$ $[t_1$ voted for him]   him = Max
       b. $[\text{Max}]^1$ $[t_1$ voted for himself$_1$]

The Coreference Rule predicts that a speaker will never use (12a), in which [Max] and [him] are intended to corefer, because (12a) is semantically indistinguishable from its binding alternative (12b): both express the proposition that Max voted for Max. A hearer will conclude from this that coreference cannot be intended in (12).

   Now, let us see whether the Coreference Rule can deal with environments in which coargument coreference is exceptionally permitted. First consider a focus construction:

(13)   Only Max himself voted for him.
       a.   [only] $[[\text{Max himself}]^1$ $[t_1$ voted for him]]   him = Max himself
       b.   [only] $[[\text{Max himself}]^1$ $[t_1$ voted for himself$_1$]]

This time, the Coreference Rule does *not* rule out coreference, because the interpretation of (13a) *differs* from the interpretation of its binding alternative (13b): (13a) says that the others did not vote for *Max*, while (13b) says that the others did not vote for *themselves*. Finally, consider one of Evans' examples:

(14)   I know what John and Mary have in common.
       Mary voted for John and John voted for him too.
       a. $[\text{John}]^1$ $[t_1$ voted for him]   him = John
       b. $[\text{John}]^1$ $[t_1$ voted for himself$_1$]

The Coreference Rule does not rule out coreference in (14a), because the interpretation of (14a) differs from the interpretation of its binding alternative (14b): (14a) says that John has the property of having voted for John, and this is indeed the property that John and Mary are supposed to have in common. (14b) on the other hand, says that John has the property of having voted for *himself*, and this is certainly not the property that John and Mary are supposed to have in common.

   So Reinhart's Coreference Rule accounts for standard Condition B effects on coreference, and also for the exceptional cases in which coreference *is* available.

3 Codetermination

Heim (1998) observes that at least three aspects of Reinhart's account need some further consideration. First, the theory does not explicitly state what it means for one LF to be *semantically indistinguishable* from another. One possibility that comes to mind immediately is that two LFs should be regarded as semantically indistinguishable if and only if they express the same proposition. But this would not work: (14a) and (14b) express the same proposition, but intuitively, at least in the context of (14), there is a significant semantic difference between them. So the question of when two LFs are semantically indistinguishable should be addressed more explicitly.

Second, Reinhart's account is only concerned with binding and coreference. There are other kinds of anaphoric relations, which seem to be constrained in similar ways. For instance, in a sentence like (15), the pronouns [he] and [him] cannot be interpreted as *cobound* (i.e., both bound by [every man]).

(15)   Every man said he voted for him.

In general, cobinding is just as unacceptable as coreference in Condition B environments, and we would like to have a rule that embodies these restrictions all in one go, rather than separate rules for coreference, cobinding, and yet other kinds of anaphora.

A third issue that Heim raises is that the Coreference Rule is not compatible with 'baseline theories' of VP ellipsis (Sag 1976; Williams 1977).[7,8] To see the problem, consider the following example:

(16)   Max called his mother and Bob did too.

The pronoun in the source clause can either be bound, as in (17a), or referential, as in (17b). According to Sag and Williams (and according to most of the refined analyses mentioned in footnote 8), (17a) gives rise to the sloppy reading of the target clause (*Bob called Bob's mother*) while (17b) gives rise to the strict reading of the target clause (*Bob called Max's mother*).

(17)   a. $[\text{Max}]^1$ $[t_1$ called his$_1$ mother]
        b. $[\text{Max}]^1$ $[t_1$ called his mother]   his $=$ Max

But the Coreference Rule rules out (17b), because it is semantically indistinguishable from its binding alternative (17a). This means that, on most accounts of VP

---

ellipsis, the Coreference Rule wrongly predicts that (16) does not have a strict reading.[9]

To address these issues, Heim does two things. First, she proposes a new constraint, which preserves all the empirical virtues of Reinhart's Coreference Rule, but is compatible with a wide range of theories of VP ellipsis and applies not only to coreference but also to cobinding and other kinds of anaphora. Second, she refines the notion of semantic indistinguishability. Let us first consider the new constraint.

3.1 The Exceptional Codetermination Rule

Heim's theory is stated in terms of *codetermination*, a notion which embraces that of binding, cobinding, and coreference (and yet other anaphoric relations). It is recursively defined as follows.

Def. 4: Codetermination

Let C be a context, let LF be a logical form, and let A and B be two DPs in LF. We say that A and B are codetermined in LF/C iff

(i)   A binds B in LF, or
(ii)   A and B corefer in C, or
(iii) There is a third DP which is codetermined with A and B in LF/C.

The first ingredient of Heim's theory is a revised version of Condition B, which prohibits codetermination, rather than binding.

Def. 5: Heim's Condition B

Pronouns cannot be codetermined with their coarguments.

The second ingredient of the theory is a rule which states that codetermination is sometimes exceptionally allowed.

Def. 6: Exceptional Codetermination Rule

Let LF be a logical form in which a pronoun is codetermined with, but not bound by one of its coarguments. Then, LF is (marginally) allowed, in violation of Condition B, if it is semantically distinguishable from its binding alternative in the given context.

The reader is invited to check that Heim's Condition B and her Exceptional Codetermination Rule account for regular Condition B effects—not only ones involving coreference, but also ones involving cobinding and other kinds of codetermination. The theory also accounts for the exceptional cases in which

---

[9] It could be objected that binding and coreference *are* semantically distinguishable here, exactly because they give rise to different interpretations of the elided VP. But, as noted by Grodzinsky and Reinhart (1993, p. 81) and Heim (1998, p. 241), this would mean that coreference should also be be allowed in examples like the following:

(i)   John saw him, and Bill did too.

Such examples indicate that the Coreference Rule must operate locally, e.g., on the smallest clause containing the pronoun in question. And under this assumption, examples like (16) are genuinely problematic.

codetermination *is* allowed in Condition B environments. Finally, it is compatible with a wide range of theories of VP ellipsis: (17b) is no longer ruled out.

3.2 Semantic indistinguishability

When should two logical forms be regarded as semantically indistinguishable? For one thing, they should express the same proposition. But Heim notes that Evans' examples, repeated in (18) and (19), show that there is more to it.

(18)   I know what John and Mary have in common.
        Mary voted for John and John voted for him too.

(19)   If everyone voted for Oscar, then certainly Oscar voted for him.

Heim suggests that these are typical cases in which *structured meaning* matters. In (18), there is a certain property P that Mary and John are supposed to have in common, namely, the property $[\lambda x \cdot x$ voted for John] of having voted for John. If such a particular property is under discussion, then an LF which says that John has the property P is to be distinguished from an LF which says that John has the property $[\lambda x \cdot x$ voted for $x]$ (even though these two LFs as a whole denote the same proposition). The same reasoning can be applied to the example in (19), where it is the property of having voted for Oscar that is under discussion. Thus, Heim's general proposal is that, whenever a particular property P is under discussion, and LF and LF′ are two logical forms such that P is denoted by some part of LF but not by any part of LF′, then LF should be distinguished from LF′, even if both express the same proposition. One way of implementing this intuition is to define semantic indistinguishability not only in terms of propositional content, but also in terms of *focus* values (see Roelofsen 2008a for the details of such an implementation).

   The three issues that Heim raised concerning Reinhart's original proposal are now resolved. It must be noted, however, that the general idea behind Reinhart's theory has been lost: Heim's theory cannot be derived from the assumption that speakers generally seek to avoid risks of being misinterpreted. Moreover, Heim assumes that codetermination—and not just binding—is subject to rules of syntactic well-formedness (in particular, her version of Condition B). This can only be effectuated if codetermination is syntactically encoded, and that requires a rather significant complication of the syntactic representation of anaphoric relations. Heim bites the bullet and spells out a specific way of encoding codetermination syntactically. Büring (2005b) proposes an alternative account, which is specifically aimed at avoiding this complication.

4 Locality

Büring appeals to a constraint on binding proposed and independently motivated by Fox (2000). Let us first consider Fox's proposal, and then Büring's application of it.

4.1 Rule H

The general idea pursued by Fox (2000) is that grammar is subject to certain *economy constraints*. Such constraints filter out logical forms whose interpretation coincides with that of alternative logical forms that are in some sense simpler, more economical. Rule H is a specific instance of this kind of economy constraints.

> *Def. 7: Rule H*
>
> A pronoun P can be bound by an antecedent A only if there is no closer antecedent B such that it is possible to bind P to B and get the same semantic interpretation.

Empirical evidence for Rule H mainly comes from a notorious puzzle concerning VP ellipsis, dating back to Dahl (1974). Consider the following sentence:

(20)   Max said that he called his mother and Bob did too.

Notice that the second conjunct contains an elided VP, and that the overt VP in the first conjunct contains two pronouns. We may expect, then, that this sentence has at least four readings: one in which both pronouns are interpreted strictly, one in which they are both interpreted sloppily, and two 'mixed' readings where one of the pronouns is interpreted strictly and the other sloppily. Surprisingly, one of these mixed readings is not available (in neutral contexts):

(21)   a.   ... Bob too said that Max called Max's mother.   [strict-strict]
      b.   ... Bob too said that Bob called Bob's mother.   [sloppy-sloppy]
      c.   ... Bob too said that Bob called Max's mother.   [sloppy-strict]
      d.   #... Bob too said that Max called Bob's mother.   [strict-sloppy]

Rule H accounts for this observation. To see this, consider the possible logical forms of the source clause in (20):

(22)   a. $[\text{Max}]^1$ $[t_1$ said $[he_1]^2$ $[t_2$ called his_2 mother$]]$ (21a)
      b. $[\text{Max}]^1$ $[t_1$ said $[he_1]^2$ $[t_2$ called his_1 mother$]]$ (21a)
      c. $[\text{Max}]^1$ $[t_1$ said $[he]^2$ $[t_2$ called his mother$]]$ $he = his = \text{Max}$ (21b)
      d. $[\text{Max}]^1$ $[t_1$ said $[he]^2$ $[t_2$ called his_2 mother$]]$ $he = \text{Max}$ (21b)
      e. $[\text{Max}]^1$ $[t_1$ said $[he_1]^2$ $[t_2$ called his mother$]]$ $his = \text{Max}$ (21c)
      f. $[\text{Max}]^1$ $[t_1$ said $[he]^2$ $[t_2$ called his_1 mother$]]$ $he = \text{Max}$ (21d)

I have indicated which reading of (20) is associated with each of these logical forms, on most prevailing accounts of VP ellipsis. Now notice that Rule H rules out (22b) and (22f): (22b) is ruled out because it is semantically indistinguishable from the more economical (22a), and (22f) is ruled out because it is semantically indistinguishable from the more economical (22d). Therefore, Rule H correctly predicts that (21a), (21b), and (21c) are possible readings of (20), but (21d) is not.

4.2 Have Local Binding!

Büring (2005b) observes that one of the main reasons for Heim to impose a syntactic constraint on codetermination, rather than binding, was that Reinhart's original account could not deal with Condition B effects on anaphoric relations other than binding and coreference. For instance, Reinhart's account failed to rule out cobinding in (23):

(23)   Every man said that he voted for him.

Büring observes that once Fox's Rule H is adopted, cobinding *is* ruled out in (23). He notes that (24), which exhibits cobinding, is semantically indistinguishable from (25), where binding is established locally.

(24)   [Every man]^1 [t_1 said that [he$_1$]^2 [t_2 voted for him$_1$]]

(25)   [Every man]^1 [t_1 said that [he$_1$]^2 [t_2 voted for him$_2$]]

Thus, Büring argues, the complications proposed by Heim become superfluous. He proposes to return to the simple formulation of Condition B (as in Def. 1) and to adopt Reinhart's original Coreference Rule alongside Fox's Rule H. Finally, he observes that the latter two constraints can actually be collapsed into one:

Def. 8: Have Local Binding! (HLB)

For any two determiner phrases A and B, if A is in a position to bind B (i.e. if it c-commands B and B is not bound in A's c-command domain already), then A must bind B, unless that changes the interpretation.

The reader is invited to check that this constraint does indeed provide a general account of Condition B effects, in particular for those cases in which coreference and other kinds of codetermination are exceptionally permitted.

It must be noted, however, that it is not quite clear what should be regarded as the general idea behind HLB. Reinhart's Coreference Rule and Fox's Rule H were both derived from general ideas about the workings of grammar and the behavior of speakers in discourse. However, these ideas seem to be quite independent, and it is not easy to see how they could be unified. Thus, even though HLB embodies an elegant unification of the Coreference Rule and Rule H at the definitional level, it is not clear whether this unification can also be carried through at the conceptual level.

Moreover, and perhaps more importantly, HLB is not compatible with minimal assumptions about VP ellipsis. That is, in combination with most prevailing theories of VP ellipsis, HLB yields wrong predictions both for cases like (16)—which were already seen to be problematic for Reinhart's account—and for the more intricate pattern found in Dahl's puzzle. To see this, first consider (16), repeated here as (26):

(26)   Max called his mother and Bob did too.

HLB predicts that coreference is impossible in the source clause. Thus, on most accounts of VP ellipsis, HLB wrongly predicts that the target clause does not have a strict reading. As for Dahl's puzzle, HLB allows (22a) but rules out all the other LFs in (22). This means that, again, on most accounts of VP ellipsis, HLB wrongly predicts that (21a) is the only possible reading of (20). So Büring is forced to make very specific assumptions about VP ellipsis. He adopts a proposal originally due to Fox (2000):[10]

Def. 9: DP Parallelism

Corresponding determiner phrases in the antecedent and elided VPs must either:

(i)   have the same referential value, or
(ii)   be bound in parallel in their respective conjuncts.

Def. 10: Referential Value

The referential value of a determiner phrase A is:

(i)   the individual to which A refers, or
(ii)   the referential value of the DP that binds A.

Neither Fox nor Büring says explicitly what it means for two determiner phrases to be 'bound in parallel'. Nevertheless, we could ask whether the generalization that DP Parallelism seems to embody is empirically correct. I think that it is too strong, and arguably also too weak. That it is too strong is illustrated by a well-known example from Rooth (1992):

(27)   First John told Mary that I was bad-mouthing her. Then Sue heard I was.

DP Parallelism erroneously rules out the sloppy reading of (27), which says that Sue heard that I was bad-mouthing Sue. This is because such a sloppy reading would involve 'non-parallel' binding of the pronouns in the elided VP and its antecedent. Rooth's account of examples like (27) essentially amounts to requiring that the elided VP is contained in a constituent that 'contrasts appropriately' with another constituent in the surrounding discourse, where 'appropriate contrast' is formalized in terms of focus values. Many subsequent theories of VP ellipsis have adopted or refined Rooth's proposal (cf. Heim 1997; Tomioka 1997; Fox 1999; Merchant 2001; Hardt 2008). All such theories are compatible with Heim's ECR, but not with HLB.

It can also be argued that DP Parallelism is too weak (although such arguments could always be countered by pointing out that there may be constraints on VP ellipsis *besides* DP Parallelism that rule out the relevant readings). One problematic kind of example is the following:

(28)   Yesterday, only Bob called his mother. Today, only Max did.

---

[10] The present formulation of DP Parallelism is taken from Büring (2005a, p. 132). A slightly different formulation is given by Büring (2005b, p. 267). I assume that the two formulations are intended to be equivalent.

DP Parallelism predicts that there is a reading of (28) which says that yesterday, Bob was the only one with the property [$\lambda x \cdot x$ called x's mother], while today, Max was the only one with the property [$\lambda x \cdot x$ called Bob's mother]. This reading is not attested.

Other cases in which DP Parallelism is too weak are certain instances of *cascaded ellipsis*. Consider, for instance:

(29) Bob called his mother, and Max did too. But Tom didn't.

DP Parallelism wrongly predicts that (29) has a so-called *mixed reading* which can be paraphrased as follows:

(30) Bob called Bob's mother, and Max called Max's mother.
      But Tom didn't call Max's mother.

It must immediately be noted that there are also certain instances of cascaded ellipsis where mixed readings *are* available, an observation going back at least to Schiebe (1973) and Dahl (1973). Fox (2000, p. 117) presents the following example (explicitly in support of DP Parallelism):

(31) Smithers thinks that his job sucks, and Homer does, too. But Marge doesn't.

Notice that (31) is structurally analogous to (29). However, (31) does have a mixed reading, at least for people who recognize that it is about a popular American sitcom, in which Marge is Homer's wife and does not have a job of her own. The availability of such a mixed reading is predicted by DP Parallelism, and not by most other theories of VP ellipsis, at least not without additional stipulations. However, I think it goes too far to say that this really is an argument in favor of DP Parallelism, and against most other theories of VP ellipsis. DP Parallelism has trouble with cases like (29), while other theories struggle with cases like (31). To the best of my knowledge, none of the available theories provides a completely satisfactory account of cascaded ellipsis. My hunch is that mixed readings arise in cases like (31) because such constructions somehow license *inference* in the interpretation of the elided VP. It is well known that the interpretation of elided VPs must involve inference in certain cases (cf. Webber 1978; Hardt 1993; Sag 2006). Unfortunately, not much is known about the exact circumstances under which inference is licensed (although see Hardt 2005 for a proposal). I will leave this issue open here, but must conclude that cascaded ellipsis does not in general provide evidence for DP Parallelism.

Thus, HLB requires a very specific account of VP ellipsis, which—at least in its present formulation—seems problematic. Notice that HLB essentially inherits this problem from Reinhart's Coreference Rule. I think that the persistence of this problem was in fact the main reason for Reinhart herself to eventually depart from her 1983 theory and develop a new account in the mid-1990s.

5 Reinhart's Interface Rule

Reinhart's original approach was based on two assumptions. First, binding is encoded in syntactic structure, while coreference is not; and second, speakers generally try to avoid risks of being misinterpreted. Reinhart argued that it follows from these two assumptions that speakers generally prefer to use bound pronouns, which explicitly encode the intended anaphoric relations, rather than referential pronouns, which may well not be resolved as intended.

In her later work, Reinhart concludes that these assumptions eventually yield the wrong predictions and should be reconsidered. More specifically, she proposes to leave the first assumption intact (binding is encoded in syntactic structure, coreference is not) but to replace the second assumption, which is about *speakers*, with an alternative assumption about *hearers*. The general idea she pursues in Reinhart (2006) is that hearers *minimize interpretive options*. In the specific case of anaphora, this means that if a certain interpretation is ruled out by grammatical restrictions on binding, then a hearer will recognize that this interpretation was not intended, even if it could in principle be derived via other anaphoric mechanisms. In other words, interpretations which are ruled out by restrictions on binding cannot be sneaked in via other anaphoric mechanisms. Reinhart points out that the existence of such an interpretive restriction would be extremely useful. For communication to proceed efficiently, it is crucial for a hearer to keep interpretive options to a minimum at all times.

Reinhart, (2006) formalizes this idea in terms of a notion she calls *covaluation*.

Def. 11: Covaluation

Let C be a context, let LF be a logical form, and let A and B be two DPs in LF. Then A and B are covalued in LF/C iff

(i)   A does not bind B and B does not bind A in LF, and
(ii)   A and B are cobound in LF or A and B corefer in C.

Notice that covaluation is essentially a generic term for coreference and cobinding. As such, it is more general than coreference alone, but less general than Heim's notion of codetermination, which covered other kinds of anaphora as well. Reinhart proposes the following constraint on covaluation:

Def. 12: Interface Rule

A logical form LF is ruled out if one of its binding alternatives LF′ is such that

a.   LF and LF′ are semantically indistinguishable, and
b.   The transition from LF to LF′ is illicit, because

(i)   LF′ is ruled out by restrictions on binding (Condition B), or
(ii)   The existing binding relations in LF are not preserved in LF′, or
(iii)   LF′ is ruled out by another application of the Interface Rule.

The notion of a *binding alternative* has to be revised slightly. The only difference between Def. 13 below and the earlier Def. 3 is that in Def. 13 A and B are supposed to be *covalued*, while in Def. 3 A and B were supposed to *corefer*.

Def. 13: Binding Alternatives

Let C be a context, let LF be a logical form, and let A and B be two determiner phrases such that A c-commands B in LF and such that A and B are covalued in LF/ C. Then the structure obtained from LF by

(i)   quantifier raising A in case it has not been raised yet
(ii) replacing B with a pronoun or trace bound by A

is called a *binding alternative* of LF.

Let us see whether the Interface Rule accounts for the relevant data. Let us first consider (12), repeated here as (32), which exhibits a standard Condition B effect on coreference. The Interface Rule correctly rules out (32a) because it is semantically indistinguishable from its binding alternative in (32b), and (32b) violates Condition B.

(32)   Max voted for him.
    a. $[\text{Max}]^1$ $[t_1$ voted for him] $\qquad\qquad\qquad\qquad$ him = Max
    b. $[\text{Max}]^1$ $[t_1$ voted for $\text{him}_1]$

Now, let us see whether the Interface Rule can deal with Condition B environments in which coreference is exceptionally permitted. Consider (13), repeated here as (33):

(33)   Only Max himself voted for him.
    a.   [only] $[[\text{Max himself}]^1$ $[t_1$ voted for him]] $\qquad$ him = Max himself
    b.   [only] $[[\text{Max himself}]^1$ $[t_1$ voted for $\text{him}_1]]$

The Interface Rule does not rule out coreference here, because (33a) is not semantically indistinguishable from its binding alternative in (33b). Intuitively speaking, coreference is not sneaking in an interpretation that is ruled out by restrictions on binding, but rather gives rise to an interpretation that is different from what would be obtained from binding. Evans' examples are dealt with in a similar way. So the Interface Rule accounts for standard Condition B effects on coreference, and also for the cases in which coreference is exceptionally allowed.

The issues that the Coreference Rule was facing and which were addressed by Heim's Exceptional Codetermination Rule are also satisfactorily dealt with by the Interface Rule. In particular, the Interface Rule does not only account for cases of illicit coreference, but also for cases of illicit cobinding, and it allows strict identity readings in VP ellipsis. First consider a case of illicit cobinding:

(34)   Every man said that he voted for him.
    a. $[\text{Every man}]^1$ $[t_1$ said that $[\text{he}_1]^2$ $[t_2$ voted for $\text{him}_1]]$
    b. $[\text{Every man}]^1$ $[t_1$ said that $[\text{he}_1]^2$ $[t_2$ voted for $\text{him}_2]]$

The logical form in (34a), in which [he₁] and [him₁] are cobound, should be ruled out. The Interface Rule accounts for this: (34a) is ruled out because it is semantically indistinguishable from its binding alternative (34b), and (34b) violates Condition B.

Next, consider the simple case of VP ellipsis in (16), repeated here as (35):

(35)   Max called his mother and Bob did too.

The Coreference Rule prohibited coreference in the source clause and thus ruled out the strict reading. The Interface Rule, on the other hand, does *not* rule out coreference in the source clause (coreference is only ruled out if binding is, too, and binding is certainly possible here) and thus correctly admits the strict reading of (35).

Finally, Reinhart claims that the Interface Rule also accounts for Dahl's puzzle. To see whether this is indeed the case, let me briefly resume the puzzle:

(20)   Max said that he called his mother. Bob did too.

(21)   a.   ... Bob too said that Bob called Bob's mother.
       b.   ... Bob too said that Max called Max's mother.
       c.   ... Bob too said that Bob called Max's mother.
       d.   #...Bob too said that Max called Bob's mother.

(22)   a. $[\text{Max}]^1$ $[t_1$ said $[\text{he}_1]^2$ $[t_2$ called his$_2$ mother$]]$ (21a)
       b. $[\text{Max}]^1$ $[t_1$ said $[\text{he}_1]^2$ $[t_2$ called his$_1$ mother$]]$ (21a)
       c. $[\text{Max}]^1$ $[t_1$ said $[\text{he}]^2$ $[t_2$ called his mother$]]$ he = his = Max   (21b)
       d. $[\text{Max}]^1$ $[t_1$ said $[\text{he}]^2$ $[t_2$ called his$_2$ mother$]]$ he = Max   (21b)
       e. $[\text{Max}]^1$ $[t_1$ said $[\text{he}_1]^2$ $[t_2$ called his mother$]]$ his = Max   (21c)
       f. $[\text{Max}]^1$ $[t_1$ said $[\text{he}]^2$ $[t_2$ called his$_1$ mother$]]$ he = Max   (21d)

The Interface Rule is supposed to do two things: first, it is supposed to rule out (22f) as a logical form of the source clause, and thus (21d) as a possible reading of the target clause. Second, it is supposed to allow enough logical forms of the source clause to derive each of the legitimate readings of the target clause. In particular, it should not rule out both (22a) and (22b), or both (22c) and (22d), or (22e).

Let us see whether this is indeed established. First consider (22f). This LF is indeed ruled out. To see this, we have to consider the binding alternative of (22f), which is (22b). First observe that (22f) and (22b) are semantically indistinguishable. Next, observe that (22b) is ruled out *by another application of the Interface Rule*: (22b) is semantically indistinguishable from its binding alternative, (22a), and (22a) does not leave the existing binding relations in (22b) intact: [his] is no longer bound by [Max] in (22a). So (22b) is ruled out by the Interface Rule, and this means that (22f) itself is prohibited as well.

So far so good. The problem is that the Interface Rule also rules out (22e), by exactly the same line of reasoning: (22e)'s binding alternative is (22b), just like that of (22f). (22e) and (22b) are semantically indistinguishable, and we have already

seen that (22b) is ruled out by the Interface Rule. Thus, (22e) is ruled out as well. As a consequence, (21c) is wrongly excluded as a possible reading of the target clause.

Thus, the Interface Rule does not account for Dahl's puzzle. Note that it is not just too *weak* in this respect. If this were the case, we could simply adopt an independent account of Dahl's puzzle, such as Fox's Rule H, to complement the Interface Rule. The problem is more serious, though. The Interface Rule is too *strong*. It rules out one of the readings of Dahl's sentence that is actually attested.

Apart from this empirical problem, there are two other aspects of Reinhart's theory that call for further attention. First, there is a striking discrepancy between (the simplicity of) the intuition behind the Interface Rule and (the complexity of) its actual formulation. Recall the basic intuition: interpretations which are ruled out by restrictions on binding cannot be sneaked in via other anaphoric mechanisms. We should expect, then, that the formal statement of the rule should say something like: a logical form LF is ruled out if it is semantically indistinguishable from one of its binding alternatives LF′, and LF′ is ruled out by constraints on binding. The actual formulation of the Interface Rule is much more complicated. In particular, it additionally requires that the existing binding relations in LF are preserved in LF′ and that LF′ is not ruled out by recursive applications of the Interface Rule.

The second issue that needs further attention is that the Interface Rule is not only undesirably complex in its formulation, but also in its workings. The analysis of a sentence like (36) illustrates this:

(36)   Max said that he voted for him.

First note that a simpler example, like *Max voted for him*, is straightforwardly accounted for: Condition B prohibits binding and (as a consequence) the Interface Rule prohibits covaluation. We would like the Interface Rule to deal with the more complex example in (36) in a similar way. But this turns out not to be the case. To see this, consider the logical form in (37).

(37) $[Max]^1$ $[t_1$ said that he voted for $him_1]$ he $=$ Max

The reading represented by (37) is not an available reading for (36), so the LF should be ruled out. To see if it is, we should consider its binding alternative:

(38) $[Max]^1$ $[t_1$ said that he_1 voted for $him_1]$

Is the transition from (37) to (38) illicit? Only if (38) is ruled out by another application of the Interface Rule. To see if it is, we must consider the binding alternative of (38):

(39) $[Max]^1$ $[t_1$ said that $[he_1]^2$ $[t_2$ voted for $him_2]]$

The fact that [he] binds [him] in (39) is in conflict with Condition B. Now we can start to calculate backwards to the original LF: (39) is ruled out by Condition B; therefore, the Interface Rule rules out (38); and therefore, another application of the

Interface Rule rules out (37). So the Interface Rule does account for the illegitimacy of (37), but in a roundabout way. And more complex examples can easily be constructed of course.

In conclusion, Reinhart's Interface Rule successfully accounts for Condition B effects. Moreover, it allows for strict readings in VP ellipsis. But it does *not* account for Dahl's puzzle, its actual formulation is more complex than its underlying intuition, and its workings are (sometimes) undesirably complicated.

6 Rule S

Now we come to the positive contribution of the present paper. I will propose an interface rule that is simpler than the one proposed by Reinhart, but still reflects the underlying intuition. The rule will be referred to as *Rule S*; it literally prohibits *sneaking in* interpretations that are ruled out by restrictions on binding. We will see that Rule S compares favorably to all proposals considered so far.

Recall, again, the basic intuition underlying Reinhart's Interface Rule:

Interpretations which are ruled out by restrictions on binding cannot be sneaked in via other anaphoric mechanisms.

Let me first consider, and slightly revise, Reinhart's formal rendering of *other anaphoric mechanisms*. Reinhart assumes that these mechanisms are all instances of covaluation (cobinding and coreference). But the notion of covaluation does not cover all the relevant instances of anaphoric relatedness. In particular, it does not cover the *indirect* instances of anaphoric relatedness *via third parties*, so to speak. We have just seen in Sect. 5 that it took various recursive applications of the Interface Rule to rule out logical forms like (40) below, whereas (41) was ruled out in one simple step.

(40) $[\text{Max}]^1$ $[t_1$ said that he voted for $\text{him}_1]$ he = Max

(41) $[\text{Max voted for him}]$ him = Max

Essentially, this is because the indirect way in which [he] and [him] are anaphorically related in (40) does not count as covaluation. The complication can easily be avoided by adopting Heim's notion of *codetermination* instead of Reinhart's notion of covaluation. The definition is repeated here for convenience.

Def. 14: Codetermination

Let C be a context, let LF be a logical form, and let A and B be two DPs in LF. We say that A and B are codetermined in LF/C iff

(i)   A binds B in LF, or
(ii)   A and B corefer in C, or
(iii)   There is a third DP which is codetermined with A and B in LF/C.

Next, I propose the following simplified interface rule:

Def. 15: Rule S (preliminary version)

A logical form is illicit if it is semantically indistinguishable from one of its binding alternatives LF′, and LF′ is ruled out by constraints on binding (Condition B).

The notion of a *binding alternative* has to be slightly revised again. The only difference between Def. 16 and Def. 13 (Sect. 5) is that in the former A and B are supposed to be *codetermined*, while in the latter A and B were supposed to *covalued*.

Def. 16: Binding Alternatives

Let C be a context, let LF be a logical form, and let A and B be two determiner phrases such that A c-commands B in LF, A and B are codetermined in LF/C, but A does not bind B in LF. Then the structure obtained from LF by

(i)   quantifier raising A in case it has not been raised yet, and
(ii)   replacing B with a pronoun or trace bound by A.

is called a *binding alternative* of LF.

Notice that Rule S is in correspondence with Reinhart's original intuition. Also, the workings of Rule S are as straightforward as we would like them to be. In particular, (40) is now dealt with just like (41): codetermination is ruled out in one simple step.

Rule S is weaker than the Interface Rule. In particular, it allows too *many* (instead of too *few*) readings of the target clause in Dahl's puzzle. In fact, Rule S does not rule out *any* of the possible logical forms of the source clause of Dahl's sentence, because none of their respective binding alternatives violates Condition B.

Thus, Rule S accounts just for Condition B effects, and Dahl's puzzle must be accounted for in some other way, for instance by Fox's Rule H.[11,12] The crucial point, however, is that the excessive strength of Reinhart's Interface Rule has been tempered.

---

[11] Alternative accounts of Dahl's puzzle have been proposed by Kehler (1993), Fiengo and May (1994), Sem (1994), Schlenker (2005), Jäger (2005), Hardt (2008), and Kehler and Büring (2008), possibly among others. In Roelofsen (2008b), I discuss a slight variant of Dahl's example that is problematic for most, if not all, of these analyses, and propose a modification of Rule H. In Roelofsen (2008a), I discuss yet another possible approach. In any case, most of these accounts are directly compatible with Rule S (in some cases, this is not immediately clear, as the underlying assumptions about the nature of pronominal anaphora and the syntax-semantics interface are rather different from the ones adopted here).

[12] Reuland (2008, p. 16) explicitly argues that Condition B effects and Dahl's puzzle should be accounted for by two distinct mechanisms, because Condition B effects are "categorical, and hard or impossible to obviate," while "the local binding requirement expressed by Rule H and its alternatives is quite sensitive to plausibility considerations." To the extent that this empirical difference is real, it supports the present proposal.

6.1 Further simplification

There is one aspect of Rule S (and of the other constraints discussed so far) that has perhaps not been paid sufficient attention to. Namely, strictly speaking, Rule S does not apply to plain logical forms. Rather, it applies to what we may call 'extended logical forms', which include a specification of how the referential pronouns occurring in the logical form are supposed to be resolved. Having drawn attention to this subtlety, we may content ourselves with the present formulation of Rule S. However, there is a simpler way to state the constraint, which avoids the issue altogether.

Def. 17: Rule S (final version)

Any interpretation of a given clause X that could be obtained via a logical form of X that violates Condition B (or other syntactic constraints on binding) is illicit.

In this formulation, Rule S directly rules out interpretations rather than 'extended logical forms'. Instead of checking the binding alternatives of each extended logical form that involves some form of codetermination, it directly filters out any interpretation that could be obtained via a logical form that violates constraints on binding. Thus, there is no comparison of extended logical forms. Rather—and this brings us perhaps even closer to Reinhart's original intuition—once an interpretation is filtered out by syntactic constraints on binding, it is also prevented from being established via other interpretive strategies.

6.2 Rule S and Heim's ECR

It turns out that Rule S, in combination with the standard version of Condition B, is empirically equivalent to Heim's ECR together with her own, stronger version of Condition B. The standard version of Condition B is only concerned with *binding*: it rules out logical forms in which a pronoun is bound by one of its coarguments. Rule S, then, extends this restriction to codetermination: it rules out logical forms in which a pronoun is codetermined with one of its coarguments, *except* when such logical forms are semantically distinguishable from their binding alternatives. In Heim's system, the division of labor between Condition B and the ECR is just a little bit different. Heim's version of Condition B is directly concerned with codetermination: it rules out any logical form in which a pronoun is codetermined with one of its coarguments. The ECR, then, takes care of the exceptions to this rule: it says that a logical form in which a pronoun is codetermined with (but not bound by) one of its coarguments is exceptionally allowed if it is semantically distinguishable from its binding alternative. Thus, both systems make exactly the same predictions. The fact that they were arrived at via different routes could be considered a positive sign in itself.

The main advantage of Rule S over the ECR is that it allows us to stick to the light version of Condition B, which is only concerned with binding. This means that we are not committed to the assumption that all forms of codetermination are syntactically encoded.

6.3 Rule S and Büring's HLB

HLB is much stronger than Rule S. First of all, HLB incorporates the effects of Fox's Rule H, thus accounting for Dahl's puzzle, about which Rule S has nothing to say. Moreover, HLB also incorporates the effects of Reinhart's Coreference Rule, which in itself is already stronger than Rule S. And we have seen that this is problematic. On most accounts of VP ellipsis, the Coreference Rule—and therefore also HLB—prohibit strict readings in basic cases of VP ellipsis, and also rule out readings of Dahl's sentence that are in fact available. To make the right predictions, HLB must be combined with a very specific theory of VP ellipsis, which, at least in its present formulation, seems to be rather problematic. Rule S avoids this problem. It is compatible with virtually any account of VP ellipsis (including DP Parallelism). Thus, the weakness of Rule S is a virtue rather than a vice. It seems to do exactly what is needed to capture Condition B effects, and nothing more than that.

6.4 The bigger picture

Rule S is a direct implementation of the idea that interpretations that are ruled out by syntactic constraints on binding cannot be sneaked in via other anaphoric mechanisms. As mentioned in Sect. 5, Reinhart (2006) presents this idea as a particular instance of the more general hypothesis that 'hearers minimize interpretive options'. She points out that there are a priori reasons to believe that hearers adopt such a strategy. Namely, doing so would significantly restrict their processing load. However, real support for the hypothesis can of course only come from empirical considerations. In order to provide such support, Reinhart (2006) attempts to show that the effects of minimizing interpretive options are not only exhibited by the interpretation of anaphora but also by the assignment of quantifier scope. Rule S refines Reinhart's account of anaphora. Thus, to the extent that its predictions are empirically correct, it provides evidence for the underlying general hypothesis.

Reuland (2008) points out that the idea underlying Rule S does not *necessarily* have to be considered a particular instance of the hypothesis that hearers minimize interpretive options. It can also be seen as an instance of a closely related, but different idea, namely that rejection of a certain interpretation by one component of the language system cannot be overruled by another component.

Reuland argues on empirical grounds that this idea may actually be more suitable than Reinhart's principle of minimizing interpretive options. I will not present Reuland's argument in detail here, and I will remain agnostic for now on which of these two general hypotheses will ultimately turn out to be more appropriate. The pertinent observation to make is just that Rule S, if correct, provides support for both.

7 Some remarks on alternative approaches

Reinhart's approach to capturing Condition B effects has been particularly prominent in the literature, but several alternatives have been proposed as well. I cannot

do justice to all of these alternative accounts here, but I will briefly discuss two of them: the *pragmatic* account of Dowty (1980), Sadock (1983), Levinson (2000), and Huang (2000), among others, and the *direct compositional* account of Jacobson (2007).[13]

Proponents of the pragmatic account argue that Condition B effects are the result of a division of labor between regular pronouns and reflexive pronouns. They assume that reflexive pronouns are more informative than regular pronouns (since regular pronouns are ambiguous, while reflexive pronouns enforce a particular 'reflexive' interpretation). Therefore, they argue, the use of a regular pronoun *implicates* that a reflexive interpretation is *not* intended (just as *Some students passed the test* implicates that not all students passed the test).

The main problem with this proposal, as has also been noted by Jacobson (2007), is that Principle B effects are not *cancelable* in the way implicatures generally are. To see this, consider the contrast between (42) and (43):

(42)   a.   Some students passed the test.
       b.   In fact, it is possible that all of them passed.

(43)   a.   John thinks that Bill voted for him.
       b.   In fact, it is possible that John thinks that Bill voted for himself.

Example (42a) implicates that not all students passed the test. This implicature is canceled in (42b). It is a characteristic feature of implicatures that they are cancelable in this way. Thus, if Principle B effects are implicatures, we should expect that they are cancelable too. Example (43) shows that this is not the case. So, at face value, pragmatic inference patterns do not seem to constitute a suitable account of Condition B effects.

However, it *does* seem plausible, at least to me, that pragmatic inference patterns have played an important role in the *diachronic* realization of Condition B effects (as described, for instance, by Levinson 2000). The strict grammatical constraints on pronominal anaphora that we observe today should then be viewed as resulting from such pragmatic inference patterns through a process of 'conventionalization'. I believe that this approach is well worth further exploring (see Roelofsen 2008a for some more discussion).

The *direct compositional* account of Jacobson (2007) assumes that pronouns belong to a different syntactic category than ordinary determiner phrases. It assumes furthermore that a transitive verb normally takes an ordinary determiner phrase as its argument. To take a pronoun as its argument, the verb must be transformed, and this transformation has an impact on the meaning of the verb: it becomes undefined for the reflexive pairs in the original denotation. Take, for instance, the verb *to praise*. This verb denotes a function from pairs of individuals to truth values. The function maps a pair $\langle x, y \rangle$ to 1 if x praises y, and to 0 if x does not praise y. Now, to take a pronoun as its argument, the verb must be transformed, and as a result of this

---

[13] Some other recent proposals worth mentioning are those of Kiparsky (2002), Schlenker (2005), and Heim (2007) (see also the second part of Roelofsen 2008a).

transformation the denotation of the verb becomes a *partial* function from pairs of individuals to truth values. It maps $\langle x, y \rangle$ to 1 if $x \neq y$ and x praises y; it maps $\langle x, y \rangle$ to 0 if $x \neq y$ and x does not praise y; but it is *undefined* in case $x = y$. Thus, Jacobson attributes Condition B effects to the very core machinery of syntactic and semantic composition. Given the above assumptions, it follows directly from the way in which lexical items are composed into larger linguistic entities that transitive verbs with pronominal arguments can only apply to pairs consisting of two distinct individuals.

This mechanism provides a general account of Condition B effects, not only on binding, but also on cobinding, coreference, and other kinds of codetermination. And it does so in one fell swoop, without assuming a distinction between bound and referential pronouns, and even without assuming any syntactic representation of anaphoric relations. In this respect, the account is certainly attractive, and one of the reviewers of the present article even suggested that it may be possible to think of Reinhart's 'two-step' approach as 'simulating' the effects of the direct compositional account.

I don't think that this is exactly the right picture. Reinhart's two-step approach was motivated by the empirical observation that coreference sometimes seems to escape the constraints that invariably apply to binding. Jacobson's approach is simpler, which is attractive from a methodological point of view, but—as she herself notes—it does not yield any predictions regarding the supposed difference between binding and coreference. To be sure, more empirical work is needed to establish the exact differences between binding and coreference (see the remarks in Sect. 2.2). But in any case, I don't think that the 'two-step' approach should be thought of as simulating the effects of a direct compositional account such as Jacobson's. It is really designed to capture certain distinctions that the latter account does not deal with at all.

Now, apart from this issue, my main worry with the direct compositional account is that, as it stands, it actually places certain restrictions on interpretation that have nothing to do with Condition B effects per se. Jacobson (2007, p. 233) in fact provides an example to illustrate this:

(44)   Bill thinks that no one voted for him.   him = Bill

On the direct compositional account, this sentence does not entail that Bill thinks that he didn't vote for himself. It only entails that Bill thinks that no one *else* voted for him. Jacobson argues that this prediction is in fact not all that bad: the 'no one else' reading, which does indeed arise in this case, is usually explained on pragmatic grounds, but in principle there is nothing against deriving it semantically. However, I think that there are other cases where 'no one' is really read as 'no one including Bill', while Jacobson's account would still derive the 'no one else' reading. For instance:

(45)   *Context*:
        All club members, including Bill, elected a new president. The candidates
        were Bill and Sue. The votes were submitted anonymously. Mary told us
        that she had voted for Bill, but when we counted the votes it turned out that
        she had lied.

        Target sentence:
        In fact, no one had voted for him.

(46)   *Context*:
        All club members, including Bill, elected a new president. The candidates
        were Bill and Sue. The votes were submitted anonymously. Mary counted
        them and announced the results.

        Target sentence:
        Mary told Bill that no one had voted for him.
        (But Bill knew she was lying, because he had actually voted for himself.)

Thus, Jacobson's account deals with Condition B effects in one fell swoop, but the
proposed mechanism is so strong that it also yields restrictions on interpretation
outside the realm of Condition B effects, which seems undesirable.

8 Conclusion

Building on the work of Reinhart, Heim, Fox, and Büring, we have arrived at a
theory that captures Condition B effects in two simple steps:

– Condition B: a syntactic constraint on binding (which may be derived from
  primitive syntactic mechanisms, cf. Reuland 2001);
– Rule S: an interface strategy that constrains anaphoric relations other than
  binding.

We have seen that the proposed theory only requires syntactic encoding of binding
relations (and not of other kinds of codetermination); it is compatible with a wide
range of theories of VP ellipsis; and, if correct, it provides support for the general
conceptual outlook of Reinhart (2006) and Reuland (2008).

References

Asher, N., D. Hardt, and J. Busquets. 2001. Discourse parallelism, ellipsis, and ambiguity. *Journal of Semantics* 18 (1): 1–25.

Büring, D. 2005a. *Binding theory*. Cambridge: Cambridge University Press.

Büring, D. 2005b. Bound to bind. *Linguistic Inquiry* 36 (2): 259–274.

Dahl, O. 1973. On so-called sloppy identity. *Synthese* 26: 81–112.

Dahl, O. 1974. How to open a sentence: Abstraction in natural language. Logical grammar reports 12, University of Götenborg.

Dalrymple, M., S. Shieber, and F. Pereira. 1991. Ellipsis and higher-order unification. *Linguistics and Philosophy* 14: 399–452.

Dowty, D. 1980. Comments on the paper by Bach and Partee. In *Papers from the parasession on pronouns and anaphora*, ed. J. Kreiman and A. Ojeda, 29–40. Chicago: Chicago Linguistic Society.

Elbourne, P. 2008. Ellipsis sites as definite descriptions. *Linguistic Inquiry* 39 (2): 191–220.

Evans, G. 1980. Pronouns. *Linguistic Inquiry* 11: 337–362.

Fiengo, R., and R. May. 1994. *Indices and identity*. Cambridge, MA: MIT Press.

Fox, D. 1999. Focus, parallelism and accommodation. In *Proceedings of SALT 9*, ed. T. Matthews and D. Strolovitch, 70–90. Ithaca: CLC Publications.

Fox, D. 2000. *Economy and semantic interpretation*. Cambridge, MA: MIT Press.

Geurts, B. 1999. *Presuppostions and pronouns*. Oxford: Elsevier.

Grodzinsky, Y. 2007. *Coreference and self-ascription*. Manuscript, McGill University.

Grodzinsky, Y., and T. Reinhart. 1993. The innateness of binding and coreference. *Linguistic Inquiry* 24: 69–102.

Hardt, D. 1993. *Verb phrase ellipsis: Form, meaning, and processing*. PhD dissertation, University of Pennsylvania.

Hardt, D. 1999. Dynamic interpretation of verb phrase ellipsis. *Linguistics and Philosophy* 22: 187–221.

Hardt, D. 2005. Inference, ellipsis and deaccenting. In *Proceedings of the fifteenth Amsterdam colloquium*, ed. P. Dekker and M. Franke, 197–221.

Hardt, D. 2008. VP-ellipsis and constraints on interpretation. In *Topics in Ellipsis*, ed. K. Johnson, Cambridge: Cambridge University Press.

Hardt, D., and M. Romero. 2004. Ellipsis and the structure of discourse. *Journal of Semantics* 21: 375–414.

Heim, I. 1997. Predicates or formulas? Evidence from ellipsis. In *Proceedings of the SALT 7*, ed. A. Lawson and E. Cho, 197–221. Ithaca: CLC Publications.

Heim, I. 1998. Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach. In *The interpretive tract. MIT working papers in linguistics*, ed. U. Sauerland and O. Percus, 205–246. Originally written and distributed as a technical report at the University of Tübingen in 1992.

Heim, I. 2007. *Forks in the road to Rule I*. Invited talk at NELS-38. http://Semanticsarchive.net/Archive/TAOZGI30/Nels%2038%20talk.pdf

Heim, I., and A. Kratzer. 1998. *Semantics in generative grammar*. Blackwell Publishers.

Horn, L. 2008. "I love me some him": The landscape of non-argument datives. In *Empirical Issues in Syntax and Semantics 7*, ed. O. Bonami and C. Hofheu, 169–192. http://www.cssp.cnrs.fr/eis7.

Huang, Y. 2000. *Anaphora: A cross-linguistic study*. Oxford: Oxford University Press.

Jacobson, P. 2007. Direct compositionality and variable-free semantics: The case of 'Principle B' effects. In *Direct compositionality*, ed. C. Barker and P. Jacobson, 191–236. Oxford: Oxford University Press.

Jäger, G. 2005. *Anaphora and type logical grammar*. Heidelberg: Springer.

Johnson, K., ed. 2008. *Topics in ellipsis*. Cambridge, MA: Cambridge University Press.

Kehler, A. 1993. A discourse copying algorithm for ellipsis and anaphora resolution. In *Proceedings of the sixth European chapter of the association for computational linguistics*, 203–212. Morristown: ACL.

Kehler, A. 2002. *Coherence, reference, and the theory of grammar*. Stanford: CSLI Publications.

Kehler, A., and D. Büring. 2008. Be bound or be disjoint! In *Proceedings of the NELS 38*. http://Semanticsarchive.net/Archive/2gxMGVIO/kehler.buring.nels07.pdf.

Kiparsky, P. 2002. Disjoint reference and the typology of pronouns. In *More than words*, ed. I. Kaufmann and B. Stiebels. Berlin: Akademie Verlag.

Levinson, S. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.

Merchant, J. 2001. *The syntax of silence*. Oxford: Oxford University Press.

Pollard, C., and Sag, I. 1992. Anaphors in English and the scope of binding theory. *Linguistic Inquiry* 23: 261–303.

Prüst, H., R. Scha, and M. van den Berg. 1994. Discourse grammar and verb phrase anaphora. *Linguistics and Philosophy* 17: 261–327.

Reinhart, T. 1983. *Anaphora and semantic interpretation*. London: Croom Helm.

Reinhart, T. 2006. *Interface strategies*. Cambridge, MA: MIT Press.

Reinhart, T., and E. Reuland. 1993. Reflexivity. *Linguistic enquiry* 24: 657–720.

Reuland, E. 2001. Primitives of binding. *Linguistic Inquiry* 32: 439–492.

Reuland, E. 2008. Minimal versus not so minimal pronouns: Feature transmission, feature deletion, and the role of economy in the language system. Manuscript, Utrecht University.

Roelofsen, F. 2008a. Anaphora resolved. PhD dissertation, University of Amsterdam.

Roelofsen, F. 2008b. Free variable economy. In *Proceedings of Sinn und Bedeutung 13*, ed. A. Riester and T. Solstad, 415–424.

Rooth, M. 1992. Ellipsis redundancy and reduction redundancy. In *Proceedings of the Stuttgart workshop on ellipsis*, ed. S. Berman and A. Hestvik, 1–26.

Sadock, J. 1983. The necessary overlapping of grammatical components. In *Papers from the parasession on the interplay of phonology, morphology, and syntax*, ed. J. Richardson, M. Marks, and A. Chukerman, 198–221 Chicago: Chicago Linguistic Society.

Sag, I. 1976. Deletion and Logical Form. PhD dissertation, MIT.

Sag, I. 2006. What's LF got to do with it? In *Proceedings of the annual meeting of the Linguistic Society of America, workshop on ellipsis*, Albuquerque, NM.

Schiebe, T. 1973. Zum Problem der grammatisch relevanten Identität. In *Generative grammar in Europe*, ed. F. Kiefer and N. Ruwet, 482–527. Dordrecht: Reidel.

Schlenker, P. 2005. Non-redundancy: Towards a semantic reinterpretation of binding theory. *Natural Language Semantics* 13: 1–92.

Schwarzschild, R. 1999. Givenness, avoid F, and other constraints on the placement of accent. *Natural Language Semantics* 7: 141–177.

Sem, H. 1994. VP-ellipsis and DRT. Technical report, Dynamic interpretation of natural language, ESPRIT Basic Research Project Deliverable R2.2.B.

Tancredi, C. 1992. Deletion, de-accenting, and presupposition. PhD dissertation. MIT.

Tomioka, S. 1997. Focussing effects and NP interpretation in VP-ellipsis. PhD dissertation, University of Massachusetts, Amherst.

van der Sandt, R. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9: 333–377.

Webber, B. 1978. A formal approach to discourse anaphora. PhD dissertation, Harvard University.

Williams, E. 1977. Discourse and logical form. *Linguistic Inquiry* 8: 101–139.