Dahl's paradigm: In defense of the crossover analysis

Alex Drummond, McGill University

October 18 2012

alex.drummond@mcgill.ca

1. Introduction

Dahl (1974) observes that the interpretation of the elided VP in (1) is restricted in a surprising way. When both pronouns in the first conjunct are anteceded by *John*, the pronouns in the elided VP may receive either strict or sloppy readings. However, as shown in (2), the second pronoun may receive a sloppy reading only if the first does also:

(1) John knows that he loves his mother and Bill does too.

(2) John knows that John loves John's mother and

a. …Bill knows that Bill loves Bill's mother.

b. …Bill knows that John loves John's mother.

c. …Bill knows that Bill loves John's mother.

d. * …Bill knows that John loves Bill's mother.

The influential analyses of Dahl's paradigm presented in Fox 2000 and Reinhart 2006 have two key points in common. (i) They are both compatible with a strict parallelism

constraint on VP ellipsis.[1] (ii) They both account for strong crossover effects using the same constraint that is responsible for blocking reading (2d). Both of these points are challenged in Roelofsen's (2008, 2011) careful studies of Dahl's paradigm. Like Fox and Reinhart, Roelofsen analyzes Dahl's paradigm in terms of an economy constraint. However, Roelofsen argues that this constraint must be separate from the constraint responsible for strong crossover effects, and his analysis crucially depends on a relaxed parallelism requirement. One of the most important contributions of Roelofsen's work is to show, contra Fox and Reinhart, that the constraint responsible for the Dahl effect does not block co-binding. That is, while this constraint blocks the LF in (3a), in which a pronoun is bound across an intervening pronoun with the same referential value, it does not block (3b):

(3) a. * John thinks that he$_{=John}$ loves his mother.
|_____|

b. John thinks that he loves his mother.
|_____| |
|_____|

On Fox's and Roelofsen's assumptions, (3b) is also the pattern we find in an SCO violation such as (4):

[1] As we will see, Fox does not in fact assume a strict parallelism constraint on VP ellipsis. However, his analysis of Dahl's paradigm in terms of Rule H is nonetheless compatible with such a constraint. For other treatments of Dahl's paradigm, see e.g. Kehler (1993), Fiengo & May (1994), Büring (2005b), Schlenker (2005), Kehler & Büring (2008).

(4) * Who did he say *t* likes John?

Hence, if the constraint responsible for the Dahl effect permits (3b), it follows that (4) must be ruled out by a separate constraint. There is, however, another way of looking at SCO configurations according to which they are more analogous to (3a) than (3b). According to Reinhart (2006:173), co-binding is not available in crossover configurations due to a constraint on logical syntax. The only candidate LF is one in which the pronoun is not bound as a variable by the *wh*-trace, but merely "covalued" with it:

(5) *covaluation*

* Who did he say *t* likes John?

binding

This is essentially the configuration which we have in (3a), where a pronoun is bound across an intervening pronoun with the same value. Building on Reinhart's ideas, I will show that it is possible to formulate a constraint which blocks (5) and (3a) while permitting (3b).[2]

[2] As we will see, this constraint is a modified formulation of Rule I of Grodzinsky & Reinhart (1993). Reinhart formulates Rule I in such a way that it blocks (3b) as well as (5). However, Reinhart notes that nothing in her analysis hangs on (3b)'s being illicit, so her analysis could easily be revised in light of Roelofsen's evidence that (3b) is in fact available. Co-binding in crossover configurations is, for Reinhart, blocked not only by Rule I, but also by the (implicit) assumption that the λ-operators introduced for the purpose of interpreting *wh*-traces are introduced prior to the λ-operators introduced for the purpose of interpreting pronouns as bound variables. Thus, by the time we come to interpret *he* in (5), we already have [Who (λ*x* (did he say *x* likes John))], and it is "too late" for *who* to bind *he* (Reinhart 2006:173-174). In the

The main contention of this paper is that the crossover analysis of Dahl's paradigm is correct, but that Fox's and Reinhart's implementations of the crossover analysis are on the wrong track. Both Fox and Reinhart make use of transderivational economy constraints (Rule H and Rule I respectively) to block the unavailable reading (2d). It turns out that Dahl's paradigm itself does not really motivate the use of constraints of this type — Rule H and Rule I were formulated in this manner for independent binding-theoretic reasons. The key component of my proposal is a version of Condition C designed to block (5) together with the LFs responsible for the unavailable reading (2d). This version of Condition C can be viewed as a "local" reformulation of Rule H, by analogy with the shift in the 1990s from global to local formulations of Shortest Move.

The paper is organized as follows. I will begin in sections 2-3 by outlining Fox's and Reinhart's analyses of Dahl's paradigm. Section 4 introduces the notation which will be used to distinguish binding from other forms of antecedence relation in the remainder of the paper. Section 5 introduces my formulation of Condition C, and presents further evidence to support the hypothesis that SCO is responsible for the unavailability of (2d). Section 6 argues that pronominal binding shows weak crossover effects as well as strong crossover (SCO) effects. Section 7 concludes with some remarks on the parallelism constraint on VP ellipsis, addressing arguments against strict parallelism based on certain ellipsis phenomena.

present theory, the condition in (30) will play a similar role.

2. Fox (2000)

Fox's analysis of Dahl's paradigm has two components: an economy constraint on variable binding, and a particular formulation of the parallelism constraint on VP ellipsis. The economy constraint, Rule H, is stated as follows:

(6) *Rule H*

A pronoun, α, can be bound by an antecedent, β, only if there is no closer antecedent, γ, such that it is possible to bind α by γ and *get the same semantic interpretation*.[3]

(Fox 2000:115)

When both of the pronouns in (1) are interpreted as bound variables, Rule H has the effect of ensuring that only transitive binding is possible in the first conjunct. Co-binding is blocked because it involves a longer dependency:

(7) a. John$_1$ knows that he$_1$ loves his$_1$ mother. *(Transitive binding)*

b. * John$_1$ knows that he$_1$ loves his$_1$ mother. *(Co-binding)*

What if one of the pronouns in the first conjunct of (1) is taken to be coreferential with *John* rather than bound by it? It is clear that Fox intends Rule H to rule out one such LF:

[3] Italics in original.

the LF illustrated in (8a), where the second pronoun is bound and the first is coreferential. It is somewhat less clear whether Fox would permit first conjunct of (1) to have either of the LFs illustrated in (8b-c):

(8) a. John knows that he$_{=John}$ loves his mother. (*J* binds *his*; *J*/*he* coref.)
|_____|

b. John knows that he loves his$_{=John}$ mother. (*J* binds *he*; *J*/*his* coref.)
|_____|

c. John knows that he$_{=John}$ loves his$_{=John}$ mother. (*J*, *he* and *his* coref.)

All of the LFs in (8) are blocked by Rule I of Grodzinsky & Reinhart (1993), and at times, Fox presents Rule H as a reformulation of Rule I (Fox 2000, 124fn14). On the other hand, Büring (2005b) notes that there is some reason to doubt that Rule H can subsume or replace Rule I. Leaving these exegetical issues aside, it is in any case clear from Fox's discussion on pp. 116-117 that Fox does not assume that strict readings require the use of coreference in the antecedent VP. Thus, even if (7a) is not the only LF available for the first conjunct of (1), it must presumably be possible, on Fox's assumptions, to derive all available readings of the elided VP with (7a) as the LF of the antecedent. Fox (2000, 117) states the parallelism requirement on VP ellipsis in such a way that if the first conjunct has a transitive binding LF, only the elided VPs in (9a-c) satisfy parallelism. Fox's definition of parallelism is reproduced in (10):

(9) a. $Bill_1$ [$_{VP}$ knows that he_1 loves his_1 mother].
|_____||_____|

b. $Bill_1$ [$_{VP}$ knows that $he_{=John}$ loves $his_{=John}$ mother].

c. $Bill_1$ [$_{VP}$ knows that he_1 loves $his_{=John}$ mother].
|_____|

d. * $Bill_1$ [$_{VP}$ knows that $he_{=John}$ loves his_1 mother]. *(Not parallel to (7a))*
|_____|

(10) *NP Parallelism* (Fox 2000, 117)

NPs in the antecedent and elided VPs must either

a. have the same referential value (Referential Parallelism) or

b. be linked by identical dependencies (Structural Parallelism)

Since (9a-d) correspond to (2a-d), Fox's analysis correctly predicts the availability of the interpretations glossed in (2a-c), and the absence of the interpretation glossed in (2d).

There are two points regarding Fox's analysis which I would like to emphasize here. The first is that if the LFs in (8b-c) are in fact available in addition to (7a), then readings (2a-c) can be derived even if VP ellipsis is subject to a strict parallelism constraint. Thus, although Fox rejects strict parallelism, his analysis of Dahl's paradigm is not in itself incompatible with a strict parallelism constraint. The second point is that Fox's motivation for formulating Rule H as a transderivational economy condition derives not from properties of Dahl's paradigm itself, but from his analysis of Heim's (1998) "exceptional co-binding" examples. In short, Fox wishes Rule H to exceptionally permit co-binding in examples such as (11), where it yields an interpretation distinct from transitive binding:

(11) Every politician$_1$ is worried that only HE$_1$ voted for him$_1$.

We will see in section 7 that Roelofsen (2011) has provided strong evidence that co-binding LFs are quite generally available. If this is the case, then Fox's attempt to bring (11) together with Dahl's paradigm is probably on the wrong track, and there is no prima facie reason to suppose that the condition which blocks (3a) (and hence accounts for Dahl's paradigm) is a transderivational economy condition.

3. Reinhart (2006)

Reinhart assumes that there is a distinct LF for the antecedent VP corresponding to each of the four readings in (2). In principle, there are five possible LFs for the antecedent, as shown in (12).[4] Of these, four — (12a-d) — are licit, but (12b) and (12c) yield indistinguishable interpretations under ellipsis. Thus, there are three possible interpretations of the elided VP corresponding to (2a-c):

[4] See Reinhart (2006:194).

(12) a. John$_1$ said that he$_1$ loves his$_1$ mother. → (2a)

b. John$_1$ said that he$_{=John}$ loves his$_{=John}$ mother. → (2b)

c. John$_1$ said that he$_{=John}$ loves his$_1$ mother. → (2b)

d. John$_1$ said that he$_1$ loves his$_{=John}$ mother. → (2c)

e. * John$_1$ said that he$_{=John}$ loves his$_1$ mother.

f. * John$_1$ said that he$_1$ loves his$_1$ mother.

Readings (2a-c) can be obtained simply by copying the antecedent VP in (12a-d) over to the ellipsis site at LF. The unavailable (2d) cannot be derived from any of (12a-d), since Reinhart assumes a strict parallelism requirement.[5] The LFs in (12e-f) are ruled out by Reinhart's (2006) reformulation of Rule I. It is difficult to give a precise explanation of how Rule I blocks (12e-f), since as Roelofsen (2010) points out, Reinhart's analysis contains a technical flaw. However, the basic idea behind Reinhart's proposal is reasonably clear, and I will attempt to give a faithful exposition here. Rule I is defined as follows:[6]

[5] In Fox's theory, (12f) would, if it were not blocked by Rule H, be a possible source for (2d), but this is not the case given Reinhart's assumptions.

[6] As we will see shortly, Reinhart (2006) actually states Rule I in terms of *covaluation*, not coreference.

(13) *Rule I* (Reinhart 2006)

α and β cannot be coreferential in a derivation D if

(i) α is in a configuration to A-bind β,

(ii) α cannot A-bind β in D, and

(iii) The coreferential interpretation is indistinguishable from what would be obtained if α binds β.

It is easy to see why Rule I blocks the co-binding LF (12f) if we assume, following Reinhart, that A-binding is defined in such a way that *he* A-binds *his* in (12f). With regard to (12e), Reinhart's reasoning is somewhat more involved. Reinhart implicitly assumes that coreference relations are added after bound pronouns have been translated as variables via λ-abstraction. If *his* in (12e) is interpreted as a bound pronoun, then (12e) has the following semantic representation:

(14) John (λx (*x* said that he loves *x*'s mother)).

Suppose that we now attempt to set *he* coreferential with *his*. This would amount to translating *he* as another variable, *y*, and then adding the condition *y* = John:

(15) John (λx (*x* said that *y* loves *x*'s mother & *x* = John))

The interpretation of (15) is for all intents and purposes identical to that of (16), which is derived if both *he* and *his* are bound by *John*:

(16) John (λ*x* (*x* said that *x* loves *x*'s mother)).

We can now show that (12e) violates Rule I. In (14), *he* is in a position to A-bind *his*, and the interpretation derived by setting *he* and *his* coreferential is indistinguishable from the interpretation which would be obtained if *he* did in fact bind *his*. Conditions (i) and (iii) of Rule I are therefore met. As for condition (ii), observe that *he* in (12e) cannot in fact bind *his* because *his* has already been translated as a bound variable in (14) — to bind it again would be a violation of logical syntax. All three conditions of Rule I are therefore met, and (12e) is consequently illicit.

As mentioned above, Rule I also accounts for instances of SCO such as (17):

(17) * Who$_1$ did he$_1$ say t_1 was intelligent?

The relation between *he* and the trace in (17) is abstractly the same as that between *he* and *his* in (12e). In both cases, the second element is a bound as a variable by the antecedent, whereas the first is construed with the antecedent in some other fashion. Since it would be semantically illiterate to say that *he* in (17) is "coreferential" with the trace of *who*, Reinhart introduces the broader notion of *covaluation*:

(18) *Covaluation:* (Reinhart 2006, 172)

α and β are covalued iff neither A-binds the other and they are assigned the same value.

Two variables which are co-bound by the same quantifier are covalued on this definition, as are two variables in a transitive binding configuration. However, Reinhart notes that it is possible for two elements to be covalued in such a way that that neither is bound by the other or by the same quantifier. In the case of (17), it is consequently possible for the pronoun to be linked to the trace without being a variable bound by the *wh*-phrase (just as coreference makes it possible for *he* to be linked to *John* in (15) without being bound by *John*). If the pronoun in (17) is identified with the trace, as in (19a), then we can cache out this relation via the introduction of an additional λ operator, as in (19b):

(19) a. Who (λ*x* (*x* did he say *x* was intelligent & he = *x*))

b. Who (λ*x* (*x* (λ*y* (did *y* say *x* was intelligent))))

If Rule I is now stated in terms of covaluation rather than coreference, we can see that all three of its conditions are met in (17): (i) *he* is in a position to bind the trace, (ii) *he* cannot bind the trace (because it is already interpreted as a variable bound by the *wh*-phrase), and (iii) an identical interpretation would obtain if *he* were to bind the trace.[7] The basic intuition behind Reinhart's analysis is that both (12e) and (17) involve the use of covaluation to "sneak in" an interpretation which is blocked by a grammatical condition (the prohibition on rebinding). Rule I is in effect a prohibition on sneaky uses of

[7] With regard to (iii), Reinhart is presumably working on the assumption that if the pronoun were to bind the trace, this would derive a semantic representation in which both the pronoun and the trace would be interpreted as variables bound by the *wh*-phrase. In effect, then, (17) is blocked by the availability of "Who$_1$ t_1 said that he$_1$ is intelligent?" However, it is important to bear in mind that Rule I does not directly compare alternative syntactic derivations.

covaluation. The core thesis of this paper is that Reinhart is correct to draw a parallel between (12e) and (17), even if Rule I itself is problematic.

Just as with Fox's Rule H, we find that Reinhart's motivations for formulating Rule I as a transderivational economy condition have nothing to do with Dahl's paradigm itself. Rather, Reinhart formulates Rule I in this way because she wishes it to do all the work of the original Rule I of Grodzinsky & Reinhart (1993) in accounting for the classic Condition B and C obviation phenomena discussed in Reinhart (1983).[8] If we opt for a non-Rule-I-based analysis of these data (as Heim (2007) has recently suggested that we should), then it is quite possible to formulate a local constraint which blocks (12e). This will be the aim of section 5.

4. Notation

In what follows, it will important to have a clear notational system for LFs. Some care will be required in interpreting this notation, since we will be comparing theories which make substantially different assumptions about the semantic relations available in addition to variable binding. For Fox and Roelofsen, the only other semantic relation relevant to the discussion of Dahl's paradigm is coreference. For Reinhart, and for the analysis presented in this paper, the only other relevant semantic relation is covaluation. What is therefore required is a notation which clearly distinguishes variable binding from every other kind of semantic relation, whatever exactly may fall within this class. Büring's (2005a,b) β notation can be adapted to this end. In Büring's notation, coreference is indicated via coindexation:

[8] That is, examples such as "As for $John_1$, Mary likes him, Bill likes him_1 — even $John_1$ likes him_1!"

(20) John$_1$ thinks that he$_1$ is intelligent. *('John' and 'he' are coreferential)*

Variable binding is represented by the addition of a β bearing its own index:

(21) a. John$_1$β$_2$ thinks that he$_2$ is intelligent. *('John' binds 'he')*

b. John$_1$β$_2$ said that he$_2$ loves his$_1$ mother. *('John' binds 'he'; 'John' and 'his' are coreferential and covalued)*

I will say that *John* "binds *he* via a β node" in (21a-b), and that β$_2$ is "*John*'s β node". In this paper, coindexation with a β node will always indicate a bound variable reading. Depending on the theoretical context, different interpretations will be given to the relation holding between *John* and *he* in (20) and between *John* and *his* in (21b). When looking at an LF from the point of view of Fox or Roelofsen, this relation is interpreted as coreference. When looking at an LF from Reinhart's or my own point of view, it is interpreted as covaluation (which may be effected via coreference, but need not be). For example, (22) indicates that *he* is covalued with the *wh*-trace, and (22b) indicates that *he* and *his* are covalued:

(22) a. Who$_1$β$_2$ t_2 said that he$_1$ likes Mary

(= Reinhart's "Who (λ*x* (*x* said that *y* likes Mary & *x* = *y*))")

b. Every boy$_1$β$_2$ said that he$_2$ loves his$_1$ mother.

(= Reinhart's "Every boy (λ*x* (*x* said that *x* loves *y*'s mother & *x* = *y*))")

In cases where a quantifier does not bind a variable at all, co-indexation with the quantifier simply indicates that an element receives a covarying interpretation under that quantifier. For example, (23) indicates that *the little brat*, though it is not bound as a variable, nonetheless receives a covarying interpretation under *every boy*:

(23) Every boy_1's mother wants the little $brat_1$ to eat his greens.

The notation in (23) can be understood in a manner consistent with the notation in (22a-b). In (22b), for example, *he* is covalued with the *wh*-trace precisely because both *he* and the *wh*-trace receive a covarying interpretation under the *wh*-phrase.

Links will sometimes be added to LFs to make them more readable. These links always indicate binding dependencies, and never encode any information which the β notation does not. (21b), for example, might be rendered as (24):

(24) $John_1\beta_2$ said that he_2 loves his_1 mother.
|_____|

Transitive binding and co-binding configurations are rendered as follows:

(25) a. $John_1\beta_2$ said that $he_2\beta_3$ loves his_3 mother. *(Transitive binding)*
|_____||_____|

b. $John_1\beta_2$ said that he_2 loves his_2 mother. *(Co-binding)*
|_____| |
|_____|

Two final notes on covaluation. First, although covaluation is most simply defined in semantic terms — as in Reinhart's definition (18) — it is possible to replace (18) with a definition which makes reference only formal properties of LFs. This is essentially Heim's (1998) definition of "codetermination," which is easily adapted for LFs written using β notation.[9] Second, we must take into consideration the well-known fact that Condition C/SCO effects are triggered by overlapping valuation as well as covaluation:

(26) a. * They$_{\{1,2\}}$ think that John$_1$ is intelligent.

b. * Which boy$_1\beta_2$ do they$_{\{1,3\}}$ think t_2 is intelligent.

[9] Heim's (1998, 233) definition of the *codetermination* relation can be adapted for LFs written using β notation as follows:

(i) Two DPs (or DP traces) A and B are *codetermined* iff

(a) A = B, or

(b) either one of A or B binds the other via a β node, or

(c) A and B are coindexed, or

(d) for some C, A and C are codetermined and so are B and C.

We can now say that A and B are *covalued* if neither of A and B A-binds the other and A and B are co-determined. It will turn out in section 5, however, that the anti-A-binding condition is unnecessary, and it would therefore be sufficient to take (i) as a definition of covaluation. To handle instances of overlapping valuation such as (26a-b), (c) of (i) can be replaced by "The index sets of A and B intersect." In more recent work, Heim has offered a somewhat more explicit version of Reinhart's semantic definition of covaluation (Heim 2007:7).

I will use the set notation exemplified in (26) to indicate overlapping valuation. Reinhart's definition of covaluation in (18) is easily adapted to cover cases of overlapping valuation:

(27) *Overlapping valuation:*

α and β overlap in value iff neither A-binds the other and their values intersect.

Heim's definition of co-determination can also be adapted to handle covaluation (see footnote 9).

5. Dahl's paradigm as a strong crossover effect

Recall that Fox (2000) and Reinhart (2006) agree that the Dahl effect is a kind of SCO effect. That is, they agree that the constraint which is responsible for blocking the reading (2d) is also the constraint responsible for standard SCO effects. The aim of this section is to present a more more theoretically conservative variant of the Fox/Reinhart analysis which treats SCO violations as Condition C violations. My formulation of Condition C is more permissive than Rule H and Rule I in one key respect: it permits co-binding LFs. As we will see in section 7, Roelofsen (2011) has shown that the availability of certain types of co-binding LF is necessary to account for a number of Dahl-type phenomena.

The kinds of LF that we wish Condition C to rule out are exemplified in (28a-c):

(28) a. * Everyone$_1\beta_2$ said that he$_1$ loves his$_2$ mother.

b. * Who$_1\beta_2$ did he$_1$ say t_2 is intelligent?

c. * He$_1$/John$_1$ said that John$_1$ is intelligent.

The standard Condition C of GB theory already blocks (28b-c) (if appropriately modified to handle LFs in β notation[10]). Thus, the key question is how to define Condition C so that it blocks (28a) too. One way of achieving this would be to adopt the hypothesis that pronouns which are interpreted as bound variables are linked to their antecedents via A′-movement.[11] The second pronoun in (28a) would then have exactly the status of the A′-trace in (28b). However, since adopting this approach would take us on something of a detour, I would like to consider more conservative means of formulating Condition C to block (28a). The formulation of Condition C I have in mind is as follows:

(29) *Condition C:*

* [A … B] if

(i) A c-commands B,

(ii) B is co-indexed with a β node which c-commands A and B, and

(iii) A and B are contraindexed but overlap in value.[12]

[10] Since I will be proposing my own formulation of Condition C — (29) — I will not go into the question of how the standard Condition C can be updated to handle LFs in β notation. Leaving aside unimportant differences in notation, this issue has already been addressed in Heim (1998, 2007).

[11] See e.g. Drummond, Hornstein & Kush (2011), Kayne (2002).

[12] Overlapping valuation should be understood here in accord with the definition in (27), or alternatively according to the Heim-style definition in footnote 9. As mentioned in footnote 9, the requirement in (18)

The condition in (29) immediately rules out both (28a) and (28b). As with any other formulation of Condition C, it is necessary to make a stipulation to bring (28c) together with (28b). Traditionally, this is accomplished by taking (28c) as the base case and stipulating that A′-trace is an r-expression. I will go the opposite route, adapting a suggestion of Irene Heim's mentioned in Fox (2000,124fn14). Heim's suggestion is that proper names denote variables with the presupposition that the variable is identical to a given individual. If these variables are eventually bound by a quantifier of some sort, then (28c) will be blocked by (29).[13]

Note that whereas a licit LF can be derived from (27a) by changing the index of he from 1 to 2, no licit LF can be derived from (27b) by changing the index of he from 1 to 2. That is, if we permit co-binding, we must ensure that co-binding is not licit in crossover configurations, so that only the trace of a wh-phrase can be interpreted as a true variable bound by the wh-phrase. A key assumption of the analysis presented here is that the special relationship traces enjoy with their antecedents is also shared with a certain class of pronoun — pronouns which are true bound variables. This is what gives rise to

---

that neither of A and B A-bind the other is redundant for the purposes of stating (29). A can only A-bind B (or vice versa) if A binds B via a β node or if A and B are co-indexed. But if either of these scenarios obtains, then (ii) and (iii) of (29) will respectively fail to be satisfied quite independently of whether or not A and B overlap in value.

[13] With Q the relevant quantifier, the LF for (28c) would be as shown in (ii). Here, (29) applies with A as either *he* or the first *John* and B as the bound variable *x*:

(ii) * $Q_1\beta_2$ … $John_1/he_1$ said that [x_2 John] is intelligent.

See also Heim (2007:13-14) on giving a presuppositional semantics for names.

the analogy between standard cases of SCOand the Dahl effect. However, as the aforementioned contrast between (28a) and (28b) indicates, we must take note of the fact that traces jealously guard the special relationship that they have with their antecedents, refusing to share it with pronouns.[14] There are many possible ways of enforcing this constraint. Perhaps the simplest is via the following principle:

(30) If A is co-indexed with a trace, A is a trace or a β node.

With (29) and (30) in place, we now have a formulation of the crossover analysis of Dahl's paradigm which incorporates Reinhart's insights regarding covaluation while obviating Roelofsen's (2010) critique of Reinhart's formulation of Rule I. We will see in section 7 that the availability of covaluation makes it possible to maintain a strict parallelism constraint on VP ellipsis in the face of certain problematic ellipsis phenomena identified in Roelofsen 2011. A key difference between (28a) and (28b) is that only the latter shows an overt distinction between the structure which violates SCO and the structure which does not. This obviously makes it more difficult to probe the nature of the violation in (28b). Indeed, we have yet to see any direct evidence that (28b) really is an SCO. The following subsections present three pieces of evidence for an SCO constraint on pronominal binding which are independent of the original Dahl paradigm. Section 5.1 discusses variations on Dahl's paradigm involving fake indexicals. Section 5.2 makes use of epithets, building on an idea of McCloskey (2011). Section 5.3 presents an example of

[14] Resumptive pronouns may be an exception to the generalization that only traces may enter into the relevant kind of "special relationship."

a crossover constraint on variable binding in which the crossover effect is triggered by a DP embedded in the relevant QP. The data discussed in these subsections turn out not to have any obvious account in terms of Rule H. (Recall that Fox (2000) takes Rule H to be responsible for SCO effects.) In each case to be considered, the problem for Rule H is one of overgeneration due to the absence of a suitable interpretatively-equivalent competitor derivation. We have already seen that there is no direct empirical motivation for treating Dahl's paradigm or SCO in terms of a transderivational economy constraint. We will now see the empirical case against such an analysis.

5.1. Fake indexicals

There are certain focus constructions in which indexical pronouns appear to be interpreted as bound pronouns. (31), for example, is ambiguous. It has one reading in which the second *I* is interpreted as a true indexical — (32a) — and another in which the second *I* is a "fake indexical"[15] interpreted as a variable bound by the first *I* — (32b):

(31) Only **I** think that I'm intelligent. *(Boldface indicates focus)*

(32) a. I am the only *x* such that *x* thinks I am intelligent.

b. I am the only *x* such that *x* thinks *x* is intelligent.[16]

[15] On fake indexicals, see e.g. Kratzer (2009, 1998), Heim (1994, 2005), Rullman (2004). Examples along the lines of (31) were first discussed in Partee (1989, fn3).

[16] (32b) is not a fully adequate gloss of the relevant reading, since a *de se* interpretation is (for me at least) obligatory here. But this is not relevant to the present discussion.

It seems reasonable to assume that fake indexical readings are only available for pronouns bound as variables. This gives us another variation on Dahl's paradigm. (33) can have only readings (34a-c):

(33) Only **I** said that I'd take my car.

(34) a. I am the only *x* such that *x* said that *x* would take *x*'s car.

b. I am the only *x* such that *x* said that I would take my car.

c. I am the only *x* such that *x* said that *x* would take my car.

d. * I am the only *x* such that *x* said that I'd take *x*'s car.

The data in (33)-(34) seem amenable to analysis in terms of either Rule H or crossover. If binding in (33) is evaluated "locally" before *only* makes its contribution to the interpretation, then the LF deriving (34d) will be blocked because it is interpretatively indistinguishable from the LFs deriving (34a-c) and has a longer binding dependency. Under the crossover analysis, the LF deriving (34d) is blocked because the third pronoun is bound across the covalued second pronoun. To distinguish the Rule H account of (34) from the crossover account, we can exploit the fact that fake indexicals can be "partially bound" (Rullman 2004, Heim 2005). For example, (35a) can have the reading glossed in (35b):

(35) a. Out of everyone who's baked with Peter

…only **I** liked our banana muffins.

b. I am the only *x* such that *x* liked [*x* and Peter]'s banana muffins.

Now consider (36):

(36) Out of everyone who's baked with Peter
…only **I** said we think I'm a brilliant cook.

(36) cannot have an interpretation in which *we* and the second *I* each covary with the first *I*:

(37) I am the only *x* such that [*x* and Peter] think *x* is a brilliant cook.
(Not a possible interpretation of (36))

Such a reading would be available only if it were possible for the second and third pronouns in (36) to be co-(partially-)bound by the first.[17] Thus, even before *only* makes its contribution to the interpretation, the co-binding LF for (36) encodes an interpretation which cannot be encoded by any competing LF. Rule H therefore fails to block co-binding in (36), and incorrectly predicts that the interpretation glossed in (37) should be available. Does the crossover analysis fare any better in this instance? Recall Condition C as defined in (29), repeated here in (38):

[17] Pronouns can be partially bound, but there are no "partial binders", so transitive binding is presumably impossible in (36). Partial antecedence in general is possible (e.g. "They$_{\{1,2\}}$ think that he$_1$ is intelligent", "Every couple's counselor thinks that the husband is to blame.") However, I am working on the assumption that fake indexicals are always true bound variables.

(38) *Condition C:*

* [A … B] if

(i) A c-commands B,

(ii) B is co-indexed with a β node which c-commands A and B, and

(iii) A and B are contraindexed but overlap in value.

The co-binding LF for the relevant portion of (36) is as follows:

(39) $I_1\beta_2$ said $we_{\{2,3\}}$ think I_2'm a brilliant cook.
|_____| |
|_____ _____|

If we understand contraindexation as non-identity of index sets, then *we* and the second *I* are contraindexed but overlap in value, so that (iii) of Condition C is violated. Thus, the crossover analysis correctly rules out interpretation (37).

We do not expect to find a crossover violation if the order of the partially and fully bound fake indexicals in (39) is reversed. In an LF such as (40), the second *I* can partially bind *we*, so that *we* is does not overlap in value with the second *I*:

(40) $I_1\beta_2$ said $I_2\beta_3$ think $we_{\{3,4\}}$ are brilliant cooks.
|_____||_____|

As expected, (41) can receive the reading corresponding to (40):

(41) Out of everyone who's baked with Peter

…only **I** said I think we're brilliant cooks.

(Can mean: "I am the only x *such that* x *said that* x *thinks* x *and Peter are brilliant cooks.")*

We therefore see that the crossover analysis correctly predicts the range of available readings for (36) and (41).

Before moving on, we must be careful to define covaluation in such a way as to permit binding in acceptable examples such as (42):

(42) Every $boy_1\beta_2$ persuaded every $girl_3\beta_4$ that $they_{\{2,4\}}$ should meet his_2 parents.

If *they* and *his* are taken to overlap in value in (42), then binding of *his* by *every* is incorrectly predicted to induce a Condition C violation. However, there is a clear asymmetry in the relation between the two pronouns in (42): *they* is assigned a different value for each value of *his*, but *his* is not assigned a different value for each value of *they*. In contrast, *we* in (36) covaries symmetrically with the second *I*. It seems that only symmetrically covarying overlapping pronouns trigger Condition C.[18]

[18] Further to footnote 9, note that it is once again possible to adapt Heim's (1998) notion of codetermination to give an appropriate definition of overlapping valuation. This time, the challenge is to ensure that *they* and *his* are not taken to overlap in value in (iii), even though *we* and *I* are taken to overlap in (iv):

(iii) Every $boy_1\beta_2$ persuaded every $girl_3\beta_4$ that $they_{\{2,4\}}$ should meet his_2 parents.

(iv) $I_1\beta_2$ said $we_{\{2,3\}}$ think I_2'm the best.

5.2. McCloskey (2011) on epithets

McCloskey (2011) argues that resumptive pronouns in Irish have the status of variables linked syntactically via an A′-chain to an operator position. To support this analysis, it is crucial to show that resumptive pronouns pattern with *wh*-traces in triggering SCO violations. The problem is that in an abstract structure such as (43), there is no way of telling which pronoun is the true resumptive, and which is simply a pronoun receiving a covarying interpretation:[19]

(43) Wh_1 … $PrnA_1$ … $PrnB_1$

This is because (43) can always be parsed with PrnA as the resumptive, so that there is no way of telling whether an SCO violation would arise if PrnB were the resumptive. McCloskey's ingenious solution to this problem takes advantage of epithets. Epithets in Irish cannot be resumptives. Thus, if one of the pronouns in (43) is replaced by a covarying epithet, the remaining pronoun is disambiguated as a true resumptive. McCloskey points out that if true resumptives behave as variables, an SCO violation

---

The key difference between (iii) and (iv) is that in (iii), *they* and *his* are not bound by the same set of β nodes (*they* is bound by β_2 and β_4, whereas *his* is bound only by β_2). Thus, we can add the additional condition that in order for A and B to overlap in value, each must be bound by the same set of β nodes.

[19] On McCloskey's assumptions, the non-resumptive pronoun can receive a covarying interpretation by being bound as a variable by the *wh*-phrase. We will see that under the theory presented in this paper, the non-resumptive pronoun can receive a covarying interpretation via covaluation with the *wh*-trace, but it cannot be bound by the *wh*-phrase as a variable.

should be incurred if the epithet c-commands the pronoun. This prediction is indeed borne out — (44b) is unacceptable in Irish:

(44) a. Wh_1 … Prn_1 … Ep_1

b. * Wh_1 … Ep_1 … Prn_1

In principle, it should be possible to use a similar trick with (45) to exclude the licit parses in (46a-b):

(45) Every boy said that he loves his mother.

(Where both pronouns are related to 'every boy')

(46) a. Every boy_1 β_2 said that he_2 loves his_2 mother.

b. Every boy_1 β_2 said that he_2 loves his_1 mother.

c. * Every boy_1 β_2 said that he_1 loves his_2 mother.

The logic would go as follows. If we replace the first pronoun with an epithet, then the epithet cannot be bound as a variable by the antecedent. Thus, the second pronoun must be so bound, and a crossover violation should result. There are, however, some complications. First of all, we cannot simply replace the first pronoun with an epithet without introducing an additional Condition C violation which masks the crossover violation we wish to isolate. In (47), for example, binding of *his* by *every boy* may well

induce an SCO violation if *the little brat* receives a covarying interpretation, but since the relation between *every boy* and *the little brat* also violates Condition C, it is hard to tell:

(47) * Every $boy_1\beta_2$ said that the little $brat_1$ loves his_2 mother.

This problem is easily addressed. Since Condition C violations are (in contrast to variable binding relations) conditioned on strict c-command[20], we need only embed the antecedent slightly:

(48) Every boy's mother thinks that the little brat should do his homework.

The resulting sentence nonetheless fails to instantiate a crossover violation. This is because (48) has a parse in which *his* is bound as a variable by the epithet, which in turn receives a covarying interpretation in the scope of *every*:

(49) Every boy_1's mother thinks that the little $brat_1\beta_2$ should do his_2 homework.

To work around this issue, we must ensure that it is only the quantifier, not the epithet, which is a suitable antecedent for the pronoun. This can be achieved in English by introducing a φ-feature mismatch between the epithet and the pronoun. For example, an epithet such as *the happy couple* can receive a covarying interpretation in the scope of a

[20] Variable binding seems to be constrained by a structural configuration along the lines of "almost c-command" (Hornstein 1995) (though c.f. Shan & Barker 2006, 2008).

quantifier such as *every bride*, but *the happy couple* is not a possible antecedent for the pronoun *she*:

(50) * The happy couple said that she was going to be late.
('she' cannot refer to one half of 'the happy couple')

If we replace the epithet in (48) with one which cannot antecede the pronoun, we therefore expect an SCO effect to be triggered (since the pronoun must then "reach over" the covalued epithet to find its antecedent). Antecedence does indeed seem to be degraded in this configuration. For example, in (51a), it is difficult to obtain a reading in which *the happy couple* covaries with the quantifier. In contrast, antecedence is fully acceptable in (51b), since the pronoun does not have to "reach over" the epithet:[21]

(51) a. ?? Every bride's father said that the happy couple would be taking her new car to the honeymoon.
b. Every bride's father said that she would be taking the happy couple's new car to the honeymoon.

We can verify that the configuration in (51a) leads to an SCO effect when the bound pronoun is replaced by a *wh*-trace:

[21] Note that swapping the pronoun and the epithet in (51a) will not remove the violation, since epithets cannot be c-commanded by pronouns which overlap in value. In the same way, it will not help to swap the pronoun and the epithet in McCloskey's examples (shown abstractly in (44)).

(52) a. * Which bride did the happy couple say *t* would be honeymooning in Hawaii?

b. Which bride *t* said that the happy couple would be honeymooning in Hawaii?

The effect in (51a) is observed only if the epithet overlaps in denotation with the pronoun. That is, not just any φ-incompatible epithet which receives a covarying interpretation will do the trick. This is shown for example by the contrast between (51a) and (53):

(53) Every bride's father said that the groom would drive her to the hotel.

There is a parallel contrast between (52a) and (54):

(54) Which bride did the groom refuse to marry *t*?

This reinforces the connection between the effect in (51a) and standard instances of SCO. Rule H cannot account for the deviance of (51a), since there is no alternative interpretatively-identical LF in which the pronoun is bound by a closer antecedent.

5.3. Crosover effects triggered by a DP embedded within the quanitfier

A shown in (55), It is sometimes possible for a DP embedded inside a *wh*-phrase to trigger an SCO effect:

(55) *(Where 'they' refers to the boys)*

a. [Which of [the boys]] *t* said they would win?

b. ?? [Which of [the boys]] did they say *t* would win?

A similar effect is found in (56):

(56) *(Where 'they' refers to the boys and 'he' is bound by 'each of the boys')*

a. [Each of [the boys]] said that he thinks they are intelligent.

b. ?? [Each of [the boys]] said that they think he is intelligent.

These SCO effects appear to arise because *they* overlaps in value with the trace in (55b) and with *he* in (56b). This is clear from the semantic definition of overlapping valuation in (27).[22] To my knowledge, no analysis has previously been given of the contrast between (56a) and (56b). The hypothesis that pronouns bound as variables are subject to SCO offers a straightforward explanation of this fact. Rule H cannot account for the deviance of (55b) or (56b) because although *they* is a closer potential antecedent for the trace in (55b) and for *he* in (56b), binding the pronoun/trace by *they* would not derive an LF with the same interpretation as (55b)/(56b).

[22] It is less clear that overlapping valuation in this particular instance could be captured by the Heim-style definition proposed in footnote 9. Examples of this sort may constitute an argument that a semantic definition of covaluation / overlapping valuation is to be preferred.

6. Dahl's paradigm and Weak Crossover

If Reinhart (2006) is correct that the absence of reading (2d) is essentially the result of SCO, it is natural to ask whether there are any analogs of Dahl's paradigm demonstrating WCO effects. The following subsections examine two sets of phenomena which appear to match this description.

6.1. WCO effects triggered by a DP embedded within the quantifier

The examples in this subsection are constructed simply by embedding the relevant pronoun in the examples in section 5.3. We see that weak crossover effects, like SCO effects, can also be triggered by DPs embedded inside a *wh*-phrase — (57) — and that there is a parallel constraint on pronominal binding — (58):

(57) a. [Which of [the boys]$_1$]$_2$ t_2 said their$_1$ mother would win?

b. ? [Which of [the boys]$_1$]$_2$ did their$_1$ mother say t_2 would win?

(58) a. [Each of [the boys]$_1$]$_2$ knows that his$_2$ mother loves them$_1$.

b. ? [Each of [the boys]$_1$]$_2$ knows that their$_1$ mother loves him$_1$.

6.2. WCO effects triggered by embedding the first pronoun in a Dahl sentence

If variable binding is subject to WCO, we should also expect to find that the SCO violation in the LF for (2d) should become a weak crossover violation if the first pronoun in (1) is further embedded, as in (59). This is by analogy with (60b):

(59) John knows that all of his friends love his mother and Bill does too.

(60) a. * Who does **he** think *t* is intelligent? *(Strong crossover)*

b. ? Who does **his mother** think *t* is intelligent? *(Weak crossover)*

The presence of a Dahl effect in (59) would clearly be unexpected under Fox's analysis, since if transitive binding is impossible, Rule H should not block co-binding. Indeed, Fox (citing Fiengo & May (1994) and Kehler (1993)) points to examples such as (61), which he reports as permitting the interpretation of the elided VP glossed in (62):

(61) John said that all of his friends love his mother and Bill did too.

(62) Bill said that all of John's friends love Bill's mother. *(Compare (2d))*

Fiengo & May (1994, 156) provide the following example, which they claim makes the (62)-type reading reasonably accessible:

(63) I thought my tax accountant should do my taxes, and you did, too.

On the basis of such examples, Fox follows Fiengo & May in concluding that Dahl's paradigm is restricted to configurations in which the first pronoun is in a position to bind the second. However, all speakers who I have consulted find (62) far less accessible than the other three readings glossed in (64a-c):

(64) a. …Bill said that all of Bill's friends love Bill's mother.

b. …Bill said that all of John's friends love John's mother.

c. …Bill said that all of Bill's friends loves John's mother.

These speakers also find (63) extremely awkward under the relevant reading. Thus, even if (62) is comparably more accessible than (2d), we still need an account of why it is harder to access than (64a-c).[23] Under the crossover analysis of Dahl's paradigm, these facts are readily explicable. The reading in (62) requires a WCO violation, and is therefore more difficult to access than (64a-c). The reading in (2d) requires an SCO violation, making it even more difficult to access.

Kehler 1993 discusses an example in which Dahl's paradigm obtains in the absence of a c-command relation between the pronouns. The example in question is (65) from Sag (1976):

(65) Edith said that finding her husband nude had upset her, and Martha did too.

(66) a. …Martha said that finding Martha's husband nude had upset Martha.

b. …Martha said that finding Edith's husband nude had upset Edith.

c. …Martha said that finding Martha's husband nude had upset Edith.

d. * …Martha said that finding Edith's husband nude had upset Martha.

[23] Fiengo & May make an interesting and precise hypothesis regarding the relative "accessibility" of various readings which may explain this fact. I have nothing to say against this aspect of their theory, except to note that it is rather stipulative, whereas the present analysis accommodates the same data in terms of the independently required distinction between weak and strong crossover violations.

Kehler attempts to explain the absence of (66d) by appealing to special properties of experiencer verbs such as *upset*. Following Fiengo & May (p. 157), one might also consider the role of the PRO subject of *finding*, which may be the true antecedent of both pronouns. In any case, these grammatical details do not seem to be essential for obtaining a weaker version of the Dahl effect in the absence of strict c-command. For example, the same pattern of judgments is found in (68) as in (66), but there is no experiencer verb or PRO in (67):

(67) Edith said that the death of her husband had impoverished her, and Martha did too.

(68) a. Martha said that the death of Martha's husband had impoverished Martha.
b. Martha said that the death of Edith's husband had impoverished Edith.
c. Martha said that the death of Martha's husband had impoverished Edith.
d. * Martha said that the death of Edith's husband had impoverished Martha.

It seems, then, that embedding of the first pronoun does not get rid of the Dahl effect. Rather, it ameliorates it by replacing an SCO violation with a WCO violation. This is expected under the crossover analysis.[24] To show that Rule H cannot account for this fact will require a little more work. This is because in all of the examples we have seen so far, it is not clear that the first pronoun is sufficiently embedded to prevent it from binding the

[24] I leave open how WCO effects are to be analyzed. The key point is that once again, pronouns interpreted as true bound variables have the same status as *wh*-trace with regard to strong/weak crossover.

second as a variable.[25] To decide between Fox's analysis and the crossover analysis, we must determine whether or not the following generalization holds:

(69) *Fiengo & May's Generalization:*

The Dahl effect obtains only if the first pronoun is in a configuration to bind the second as a variable.

Testing (69) is slightly trickier than one might think. As Shan & Barker (2006, 2008) have recently emphasized, pronouns can sometimes receive co-varying interpretations even with very deeply embedded antecedents. Nonetheless, there are ways of embedding an antecedent which at least typically have the effect of making such interpretations difficult or impossible. For example, it is very difficult for a strong quantifier to bind out of a relative clause contained in a definite DP:

(70) * [The teacher who liked every student$_1\beta_2$] gave him$_2$ extra homework.

The Dahl effect (at least in its "weak" form) is nonetheless manifested in (71)-(72):

(71) John said that the teacher who liked him gave him extra homework,
and Bill did too.

[25] With regard to (67), many speakers at least marginally allow a sloppy reading for (v) and binding in (vi):

(v) The death of Mary's husband impoverished her, and the death of Jane's husband did too.

(vi) The death of every factory$_1$ worker was due to his$_1$ negligence.

(72) a. …Bill said that the teacher who liked Bill gave Bill extra homework.

b. …Bill said that the teacher who liked John gave John extra homework.

c. …Bill said that the teacher who liked Bill gave John extra homework.

d. ?? …Bill said that the teacher who liked Bill gave John extra homework.

It seems, then, that the generalization in (69) does not hold. This supports the crossover analysis of Dahl's paradigm, which predicts that reading (72d) should be difficult to access due to WCO.

7. Maintaining a simple parallelism requirement

Theoretical approaches to Dahl's paradigm and related phenomena can be roughly divided into two types. Theories of the first type hold constant some relatively simple formulation of the parallelism requirement and assume that Dahl's paradigm informs us which of the logically possible patterns of binding relations in the first conjunct are in fact licit. This is the approach taken by Reinhart and by Fiengo & May. Theories of the second type hold constant the assumption that binding in the first conjunct must be maximally local, and assume that Dahl's paradigm tells us something about the extent to which two LFs can be non-identical and nonetheless meet the parallelism constraint on VP ellipsis. This is the approach of Fox and Roelofsen.[26] The theory presented here is of the first type. As far as I am aware, no attempt has yet been made to work out the

[26] The case of Fox is somewhat complex, since while Fox argues against a strict parallelism constraint on VP ellipsis, his analysis of Dahl's paradigm is compatible with a strict parallelism constraint (see also foonote 1).

consequences for type 1 theories of the various ellipsis configurations discussed in Roelofsen (2011). As Roelofsen shows, these require a fairly substantial modification of Fox's original parallelism constraint.[27] We will see, however, that they are straightforwardly consistent with the hypothesis that Dahl's paradigm is a crossover effect. I take this to be a further point in favor of the crossover analysis.

7.1. Reverse Dahl effects

Fox (2000) shows that Rule H, in addition to accounting for Dahl's original puzzle, also accounts for what Kehler & Büring (2008) call "reverse" Dahl effects:

(73) Max claimed that Bob called his mother and Bob did too.

(74) a. …Bob claimed that Bob called Max's mother.

b. * …Bob claimed that Bob called Bob's mother.

To obtain the reading (74b) for the second conjunct of (73) without violating parallelism, the pronoun in the second conjunct would have to be bound by the first instance of *Bob*. This violates Rule H:[28]

[27] Roelofsen argues that they also require a modification of Rule H. This is not my concern here, since I am addressing the question of whether the data to be discussed in this section can be accommodated under a crossover analysis without significantly complicating the parallelism constraint on VP ellipsis.

[28] Both (75a) and (75b) violate Condition C (in both its standard formulation and the one in (29)).. However, this violation can be overcome via vehicle change (Fiengo & May 1994) (i.e., by replacing the second instance of *Bob* with a coreferential pronoun). Vehicle change is independently motivated by

(75) *LFs for second conjunct of (73):*

a. Bob_1 claimed that $Bob_1\beta_2$ called his_2 mother. *(Violates parallelism)*

b. $Bob_1\beta_2$ claimed that Bob_1 called his_2 mother. *(Violates Rule H)*

The formulation of the crossover analysis presented here also blocks reading (74b). We can obtain this reading only by assigning the LF in (76a) to the first conjunct, which yields the LF in (76b) for the second:

(76) a. *First conjunct:*

$Max_1\beta_2$ claimed that Bob_3 called his_2 mother.

b. *Second conjunct:*

$Bob_3\beta_4$ claimed that Bob_3 called his_4 mother.

(76b) is, of course, a Condition C violation by the definition in (29). So far, then, there is nothing to choose between Fox's analysis and the present analysis. However, Roelofsen (2011) points out that examples such as (77) are problematic for Fox:

(77) Every boy claimed that the jury loves his dish, and added that he did too.

---

acceptable examples such as (vii), which becomes (viii) following vehicle change:

(vii) John likes Bill's mother, and Bill does ~~like Bill's mother~~ too.

(viii) John likes $Bill_1$'s mother, and $Bill_1$ ~~does like his_1 mother~~ too.

As Roelofsen notes, (77) has a reading on which every boy added that he loved his own dish. Given Fox's theory, candidate logical forms for deriving this reading are as follows:

(78) a. Every $boy_1\beta_2$ claimed that the jury loved his_2 dish
|_____|
|
| and added that $he_2\beta_3$ loved his_3 dish too.
|_____||_____|

b. Every $boy_1\beta_2$ claimed that the jury loved his_2 dish
|_____|
|
| and added that he_2 loved his_2 dish too.
|_____| |
|_____ __|

Neither LF is in fact available, since (78a) violates parallelism and (78b) violates Rule H (there is co-binding in the second conjunct).[29] On the present account, however, co-binding is not blocked and so (78b) is correctly predicted to be available. Roelofsen also has an explanation for the availability of (78b). He develops an ingenious alternative to Rule H, "Free Variable Economy", which also permits co-binding in (78b). FVE is defined in (79)-(81):

[29] I note in passing that the problem which Roelofsen raises for Fox's analysis also arises for Heim 2007, which develops a theory in which only transitive binding can be encoded at LF.

(79) *Free variables* (Roelofsen 2011, 688)[30]

Let Σ be a logical form constituent, and let P be a bound pronoun in Σ which is c-commanded by a co-indexed β node, but which is not c-commanded by a co-indexed β node within Σ. Then the index of P is called a *free variable* in Σ.

(80) *Economy measure*

Let Σ and PI be alternatives. Then we say that Π is more economical than Σ iff some subconstituent Π′ of Π contains fewer free variables than the corresponding subconstituent Σ′ of Σ.

(81) *Free Variable Economy* (FVE)

A logical form constituent is illicit if it has a more economical alternative.

Owing to the manner in which (79) defines free variables in terms of indices in (79), FVE permits both co-binding and transitive binding in simple cases, as shown in (82a-b). However, it does not permit LFs such as (82c), in which one pronoun is bound across another pronoun coreferential with the first pronoun's antecedent. Nor does it permit co-binding in all configurations. FVE blocks (82d), for example, because the subconstituent [called his_3 mother on her_2 birthday] contains more free variables according to (81) than the corresponding subconstituents of the alternative derivations in which *her* is bound by *his* or *he*:

[30] The definition in (79) has been modified slightly from Roelofsen's original definition so as to apply to LFs written in β notation (which Roelofsen does not use). The original definition reads "....and let P be a pronoun in Σ that has a binding index, but no binder within Σ...then the binding index of P is called a *free variable* in Σ."

(82) a. Everyone$_1\beta_2$ said that he$_2$ called his$_2$ mother. *(✔ FVE)*

b. Everyone$_1\beta_2$ said that he$_2\beta_3$ called his$_3$ mother. *(✔ FVE)*

c. Max$_1\beta_2$ said that he$_1$ called his$_2$ mother. *(✖ FVE)*

d. Everyone$_1\beta_2$ said that he$_2\beta_3$ called his$_3$ mother on her$_2$ birthday. *(✖ FVE)*

That FVE blocks (82c) is an important result, since on Roelofsen's assumptions this is the only kind of LF which could license the illicit reading of the second conjunct in Dahl's paradigm. Returning to the issue raised by (77-78), we can now see that FVE permits co-binding in (78b), as is required to derive the relevant interpretation of (77).

7.2. Need we further complicate the parallelism constraint?

This subsection will consider a number of rather complex examples presented in Roelofsen (2011). Before addressing these, it may be useful to indicate my overall line of argument. The examples to be considered are problematic both for Rule H and for Free Variable Economy. Roelofsen proposes to solve the problems posed by these examples by relaxing the parallelism constraint still further (and in fact weakening it to the extent that it is no longer really a parallelism constraint). These problems can all be traced back to the hypothesis that Dahl's paradigm is a result of a ban on (some instances of) co-binding. That is, all of the problematic readings can be perfectly straightforwardly

derived *if* one assumes that co-binding is quite generally available.[31] This rather suggests that the aforementioned hypothesis is wrong. Restricting co-binding buys us an appealing account of Dahl's paradigm, but when we consider a fuller range of facts regarding the interpretation of elided VPs, restricting co-binding creates more problems than it solves. I conclude that it is better to retain a simple parallelism requirement, permit co-binding, and analyze Dahl's paradigm in terms of crossover.

Roelofsen begins by examining embedded instances of Dahl's paradigm such as (83):

(83) Every worker says that he knows when he can take home his tools, and the boss does too.

As he observes, "[(83)] has a 'mixed' reading, on which every worker *x* said that the boss knows when he, the boss, can take home *x*'s tools." On Roelofsen's assumptions, the two logical forms which could in principle derive this reading are illustrated in (84):

[31] As we have seen in the preceding subsection, although FVE permits instances of co-binding which Rule H does not, FVE does nonetheless block co-binding in certain configurations.

(84) a. Every worker$_1\beta_2$ said that he$_3\beta_4$ knows when he$_4\beta_5$ can take home his$_5$ tools,

and that TB$_6\beta_7$ ~~knows when he$_7$ can take home his$_2$ tools~~

b. Every worker$_1\beta_2$ said that he$_3\beta_4$ knows when he$_4$ can take home his$_2$ tools,

and that TB$_5\beta_6$ ~~knows when he$_6$ can take home his$_2$ tools~~

The problem for Fox and Roelofsen is that (84a) violates parallelism, whereas (84b) violates Rule H and FVE. Once again, the crossover analysis derives the relevant reading of (83) unproblematically. This is because, due to the availability of covaluation, there is an additional LF available for the first conjunct in which *his* is not bound as a variable:

(85) Every worker$_1\beta_2$ said that he$_3\beta_4$ knows when he$_4$ can take home his$_1$ tools,

and that TB$_5\beta_6$ ~~knows when he$_6$ can take home his$_1$ tools~~

This LF derives the correct interpretation, and does not violate strict parallelism or Condition C. The key point here is that given Reinhart's notion of covaluation, we can treat the embedded Dahl paradigm in exactly the same terms as the original Dahl paradigm. It is as if we were to chop off *every worker said that* from the beginning of the sentence, interpret the first *he* referentially, and set *his* coreferential with the first *he*. I.e., the LF of the first conjunct in (85) is exactly parallel to (12d), but with covaluation

instead of coreference. In contrast, Roelofsen is forced to revise parallelism to accommodate the relevant reading of (83). He does this essentially by relaxing parallelism in such a way that parallelism is satisfied in LFs like (86):

(86) "Every student thinks that his mother loves him and every professor does too."

Every student$_1\beta_2$ thinks that his$_2\beta_3$ mother loves him$_3$
|_____||_____|

and every professor$_4\beta_5$ ~~thinks that his$_5$ mother loves him$_5$~~ too.
|_____| |
|_____|

Roelofsen achieves this by replacing parallelism with a Focus Match constraint. The details of this condition need not concern us here. The point I wish to make is simply that accounting for the examples above is entirely straightforward under the crossover analysis, requiring no weakening of the parallelism requirement.

Roelofsen discusses one last example which motivates a further complication of his replacement for the parallelism constraint. This time, the problem is one of overgeneration. As it stands, Roelofsen's theory predicts that (83), repeated here as (87), should have a reading "every worker *x* said that the boss knows when *x* can take home the boss's tools":

(87) Every worker says that he knows when he can take home his tools, and the boss does too.

Since (87) does not in fact have such a reading, Roelofsen proposes to add an additional constraint on VP ellipsis, Preserve Coindexing. Again, I do not wish to go into the details of Roelofsen's analysis here, but only to note that the absence of the relevant reading follows immediately on the crossover analysis. On Roelofsen's assumptions, this reading could in principle derive from either of the following LFs:

(88) a. Every worker$_1\beta_2$ said that he$_3\beta_4$ knows when he$_2$ can take home his$_4$ tools

and that TB$_5\beta_6$ ~~knows when he$_2$ can take home his$_5$ tools~~

b. Every worker$_1\beta_2$ said that he$_2$ knows when he$_2$ can take home his$_2$ tools

and that TB$_3\beta_4$ ~~knows when he$_2$ can take home his$_4$ tools~~

(88b) violates parallelism. (88a) violates the formulation of Condition C in (29), since the last two pronouns in the first conjunct of (88a) are contraindexed but covalued (and hence

overlap in value).[32] We must also verify that there is no licit variation on (88a) in which covaluation is used to avoid binding the second pronoun :

(89) a. Every worker$_1\beta_2$ said that he$_3\beta_4$ knows when he$_1$ can take home his$_4$ tools
|_____||_____|

and that TB$_5\beta_6$ ~~knows when he$_1$ can take home his$_5$ tools~~
|_____ ___|_____|

Indeed, (89) is also blocked by Condition C.[33]

[32] The last two pronouns in the first conjunct of (88a) clearly overlap in value according to the definitions in (27) and footnote 9 (and also covary symmetrically, as required in section 5.1). With regard to the definition in footnote 9, note that by (c) of (i), the two instances of *he* in the first conjunct are codetermined. By (b) of (i), the first *he* is codetermined with *his*. Now consider (d) of (i), taking C as the first *he*, A as the second *he*, and B as *his* in the first conjunct. We have just seen that A and C are codetermined and that B and C are codetermined. Hence, A and B are codetermined — i.e., the second *he* in the first conjunct of (88a) is codetermined with *his*. We can redundantly (see footnote 12) verify that neither pronoun A-binds the other, so the pronouns overlap in value. Since *he* c-commands *his*, *he* is co-indexed with a β node which c-commands *he* and *his*, and the pronouns are contraindexed but overlap in value, Condition C is violated.

[33] The second *he* in the first conjunct c-commands *his* in the first conjunct, *his* is coindexed with a β node which c-commands *he* and *his*, and the pronouns are contraindexed but overlap in value..

8. Is a strict parallelism requirement viable?

The crossover analysis accounts for the Dahl paradigm (and the related phenomena considered above) without requiring any relaxation of the strict parallelism requirement. Indeed, everything said so far is compatible with the assumption that the LF of an elided VP is reconstructed via copying of its antecedent. There are, however, a number of arguments in the literature against such a strict interpretation of the parallelism requirement. If the more relaxed constraints assumed by Fox and Roelofsen were shown to be required for independent reasons, then the case made for the crossover analysis in this paper would be weakened.

Fox (2000), pointing to an observation of Dahl (1973), notes that a single antecedent VP can license both strict and sloppy ellipsis. For example, (90a) can have the reading glossed in (90b):

(90) a. Smithers thinks that his job sucks. Homer does too. However, Homer's wife doesn't.

b. Smithers thinks that his job sucks. Homer does ~~think that Homer's job sucks~~ too. However, Homer's wife doesn't ~~think that Homer's job sucks~~.

The strict reading for *Homer's wife doesn't* is unexpected under a strict parallelism requirement, since if the antecedent VP uses coreference, then *Homer does, too* should also receive a strict interpretation. Fox offers data of this sort in support of his formulation of the parallelism constraint, according to which the elided VP of *Homer's*

wife doesn't is parallel with an antecedent VP in which *his* is bound by *Smithers*. Fox's definition of the parallelism constraint, given in (10), is repeated in (91):

(91) *NP Parallelism* (Fox 2000, 117)

NPs in the antecedent and elided VPs must either

a. have the same referential value (Referential Parallelism)

b. be linked by identical dependencies (Structural Parallelism)

It would not be possible to adopt (91) in conjunction with the crossover analysis, since (91) permits the co-binding LF in (92a) to license the elided VP in (92b):

(92) a. John$_1\beta_2$ said that he$_2$ loves his$_2$ mother.

b. …and Bill$_3\beta_4$ did too ~~say that he$_1$ loves his$_4$ mother~~.

Thus, examples such as (90) poses a prima facie problem for the crossover analysis, which necessarily depends on a fairly strict parallelism requirement. These problematic examples are, however, dealt with quite extensively by Fiengo & May (1994, 165), and it is straightforward to reformulate their analysis within the present theory. Translated into the present framework, F&M's proposal is essentially that the reading of (90a) glossed in (90b) has the LF in (93):

(93) Smithers$_1\beta_2$ thinks that his$_2$ job sucks.

Homer$_3\beta_4$ does ~~think that his$_4$ job sucks~~ too.

However, Homer$_3\beta_5$'s wife doesn't ~~think that his$_5$ job sucks~~.

Here, the pronoun in the last conjunct is, like the other two pronouns, bound. The strict reading is derived due to "an anaphoric connection otherwise established between clauses," i.e., the anaphoric relation between the two instances of *Homer*. According to Fiengo & May, the existence of this anaphoric relation removes the need for the index of the pronoun in the third conjunct to be licensed by parallelism, so that the absence of structural parallelism between *Homer* and *his* in the second and third conjuncts is not an issue. Some technical artifice is required to make this idea precise, and I will depart here from Fiengo & May's technical implementation.

Let us begin by taking note of the fact that there is some structural constraint on bound variable interpretations. That is, in order for the pronoun to be interpreted as a variable bound by the DP in (94), some condition on the structural configuration represented by the ellipsis must be met:

(94) DP$_1\beta_2$ … pro$_2$.

Call the relevant structural relation S, and assume (i) that c-command(*a*, *b*) implies S(*a*, *b*) and (ii) that S holds between *Homer* and *his* in the third line of (93). It is an interesting artifact of the β-notation that we have two ways of constraining binding relations via S. That is, using (94) as an example, we can either require that S(β_2, pro$_2$) or that S(DP$_1$,

pro$_2$). Ordinarily, since a DP c-commands its β node and nothing of interest intervenes, both requirements are equivalent. However, consider the following abstract configuration (where S does not hold between the first DP and the pronoun):

(95) $[_{XP}\ DP_1\ \beta_2\ \ldots\ ]\ \ldots\ [_{YP}\ DP_1\ \ldots\ pronoun_2]$

This structure is licit if we require S to hold between a DP bearing the index of the antecedent and the pronoun, but illicit if we require S to hold between the β node and the pronoun. Now consider the LF in (96):

(96) Smithers$_1\beta_2$ thinks that his$_2$ job sucks.
Homer$_3\beta_4$ does ~~think that his$_4$ job sucks~~ too.
However, Homer$_3$'s wife doesn't ~~think that his$_4$ job sucks~~.

Here, although the last instance of *his* is not c-commanded by the co-indexed β node, it does stand in S to a DP co-indexed with its antecedent. (Its antecedent is the first instance of *Homer*, and it stands in S to the second instance of *Homer*, which bears the same index as the first.) Thus, we can impose structural constraints on bound variable readings in such as way that (96) is licit. The following definitions accomplish this:

(97) A is the *binding antecedent* of B if B is co-indexed with A's β-node.

(98) *Rule of bound variable interpretation:*

B is interpreted as a variable bound by A iff A is the binding antecedent of B and there exists a C covalued with A such that S(C, B).

(97)-(98) implement a variant of Fiengo & May's analysis of (90).

Fox (2000, 116fn8) briefly objects to Fiengo & May's analysis on the basis of examples where *Homer's wife* is replaced by *Marge*. He points out that people who know that Marge is Homer's wife are still able to get the reading of (90a) glossed in (90b). This suggests that a somewhat looser relation than covaluation may be able to license variable binding interpretations via (98). This is consistent with the data in (99)-(100). (100a) has a reading under which each captain's first mate shares each captain's fear regarding the fate of the relevant ship. (100b) — in which *the captain* is no longer present to establish an overt anaphoric link to the quantifier — has a similar reading under which each captain's first mate agrees with each captain. (100c), by contrast, clearly has no such reading:

(99) Every captain in the fleet fears that his ship will founder.

(100) a. The captain's first mate does, too.

b. ? The first mate does, too.

c. * John does, too.

It is not entirely clear how (98) should be modified to account for (100b). However, (100b) does remove some of the force from Fox's objection. Fox's argument is that the *Marge* examples are evidence against a strict parallelism requirement, since they cannot be accounted for on Fiengo & May's analysis. However, relaxing the parallelism requirement in the manner that Fox proposes is of no help in accounting for (100b). Rather, this example suggests an explanation in terms of a modified form of (98). That is, although the pronoun in the elided VP of (100b) does not stand in S to the quantifier which binds it (*every captain*), binding is nonetheless licit because this pronoun *does* stand in S to another DP (*the captain*), which is in some sense anaphorically linked to the quantifier. Thus, while (90) and (100b) remain problematic, they do not provide strong motivation for abandoning a strict parallelism requirement.

9. Conclusion

There is good evidence for the hypothesis that variable binding is constrained by both strong and weak crossover. Condition C, the constraint responsible for strong crossover effects, need not be formulated as a transderivational economy condition. With regard to Dahl's paradigm, the Condition C analysis has a number of advantages over analyses stated in terms of Rule H, Rule I or Free Variable Economy. It also directly supported by data showing that pronominal binding triggers both weak and strong crossover effects.

References

Büring, D. 2005a. *Binding Theory*. Cambridge University Press.

Büring, D. 2005b. "Bound to Bind." Linguistic Inquiry 36-259-274.

Chomsky, N. 1993. "A Minimalist Program for Linguistic Theory." In K. Hale & S.J. Keyser (eds.), The View from Building 20. MIT Press.

Dahl, Ö. 1973. "On So-Called Sloppy Identity." Synthese 26:81-112.

Dahl, Ö. 1974. "How to Open a Sentence: Abstraction in Natural Language." In *Logical Grammar Reports*, volume 12, University of Götenberg.

Dalrymple, M. S. Shieber & F. Pereira. 1991. "Ellipsis and Higher-Order Unification." *Linguistics and Philosophy* 14:399-452.

Fiengo, R. & R. May. 1994. *Indices and Identity*. MIT Press.

Fox, D. 2000. *Economy and Semantic Interpretation*. Linguistic Inquiry Monograph Thirty-Five, MIT Press.

Grodzinsky, Y. & T. Reinhart. 1993. "The Innateness of Binding and Coreference." *Linguistic Inquiry* 24:69-102.

Heim, I. 1994. "Puzzling Reflexive Pronouns in *de se* Reports." Paper presented at Bielfeld Conference on Indexicals. University of Bielfield. Available at http://semanticsarchive.net/Archive/mVkNjZlY/Bielefeld%20handout%2094.pdf.

Heim, I. 1998. "Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach." *MIT Working Papers in Linguistics* 25:205-246.

Heim, I. 2005. "Features on Bound Pronouns." Unpublished Ms., MIT.

Heim, I. 2007. "Forks in the Road to Rule I." In M. Abdurrahman, A. Scharl & M. Walkow (eds.) *Proceedings of NELS 38*.

Heim, I. & I. Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.

Higginbotham, J. 1983. "Logical Form, Binding and Nominals." *Linguistic Inquiry* 14:395-420.

Hornstein, 1995. *Logical Form: From GB to Minimalism*. Oxford: Blackwell.

Kayne, R. 2002. “Pronouns and their antecedents.” In S.D. Epstein & T. Seely (eds), *Derivation and Explanation in the Minimalist Program*, Blackwell, 133-166.

Kehler, A. 1993. “A Discourse Copying Algorithm for Ellipsis and Anaphora Resolution.” In *Proceedings of the Sixth Eurpoean Chapter of the Association for Computational Linguistics* (EACL-93), 203-212.

Kehler, A. & D. Büring. 2008. “Be bound or be disjoint!” In A. Schardl, M. Walkow & M. Abdurrahman (eds.) *NELS 38*, 487-501.

Koopman, H. & D. Sportiche. 1982. “Variables and the Bijection Principle.” *Linguistic Review* 2:139-160.

Kratzer, A. 1998. “More Structural Analogies Between Pronouns and Tenses.” In D. Strolovitch & A. Lawson (eds.), *Proceedings of SALT VIII*, Ithaca, NY: CLC Publications, 92-109.

Kratzer, A. 2009. “Making a Pronoun: Fake Indexicals as Windows into the Properties of Pronouns.” *Linguistic Inquiry* 40:187-237.

Kurafuji, T. 1998. “Dynamic Binding and the E-Type Strategy: Evidence from Japanese.” In D. Strolovitch & A. Lawson (eds.), *Proceedings of SALT VIII*, Ithaca, NY: CLC Publications, 129-144.

McCloskey, J. 2011. “Resumptive Pronouns, A′-binding, and levels of representation in Irish.” In A. Rouveret (ed.), *Resumptive Pronouns at the Interfaces*, John Benjamins, 65-120.

Nishigauchi, T. 1990. *Quantification in the Theory of Grammar*, volume 37. Dordrecht: Kluwer.

Partee, B. 1989. “Binding implicit variables in quantified contexts.” In C. Wiltshire, R. Graczyk & B. Music (eds.), *Papers from the 25th Regional Meeting of the Chicago Linguistic Society*, Part One, *The General Session,* Chicago: University of Chicago, Chicago Linguistic Society, 342–365.

Reinhart, T. 1983. “Coreference and Bound Anaphora: A Restatement of the Anaphora Questions.” *Linguistics and Philosophy* 6:47-88.

Reinhart, T. 2006. *Interface Strategies*. MIT Press.

Roelofsen, F. 2008. *Anaphora Resolved*. Doctoral dissertation, University of Amsterdam.

Roelofsen, F. 2010. “Condition B Effects in Two Simple Steps.” *Natural Language Semantics* 18:115-140.

Roelofsen, F. 2011. “Free Variable Economy.” *Linguistic Inquiry* 42:692-697.

Rullman, H. 2004. “First and Second Person Pronouns as Bound Variables.” *Linguistic Inquiry* 35:159-168.

Sag, I.A. 1976. *Deletion and Logical From*. Doctoral dissertation, MIT.

Schlenker, P. 2005. “Non-redundancy: Towards a Semantic Reinterpretation of Binding Theory”. *Natural Language Semantics* 13:1-92.

Shan, C. & C. Barker. 2006. “Explaining Crossover and Superiority as Left-to-Right Evaluation.” *Linguistics and Philosophy* 29:91-134.

Shan, C. & C. Barker. 2008. “Donkey Anaphora is In-Scope Binding.” *Semantics & Pragmatics* 1:1-46.

Tamioka, S. 1999. “A Sloppy Identity Puzzle.” *Natural Language Semantics* 7:217-241.

Watanabe, S. 1993. “A Note on So-Called ‘donkey’ Sentences in Japanese: A Preliminary Study.” In H. Hoji (ed.), *Japanese/Korean Linguistics Volume 2*, Stanford: CSLI, 299-315.