Dahl's Puzzle: Economy or Crossover?

Alex Drummond
UMass Amherst

September 25 2014

Fox (2000) and Reinhart (2006) each present unified analyses of Strong Crossover effects and Dahl's paradigm. Both analyses are compatible with a strict parallelism constraint on VP ellipsis. The connection between Strong Crossover and Dahl's paradigm has recently been called into question (Roelofsen 2008, 2010, 2011). Roelofsen proposes a modified formulation of Fox's Rule H (Free Variable Economy) which, in combination with a relaxed parallelism constraint, successfully accounts for Dahl's paradigm and related phenomena. Since Free Variable Economy permits co-binding in some configurations it cannot account for Strong Crossover. Roelofsen argues that Strong Crossover effects derive from a separate constraint which applies only to movement chains. Contra Roelofsen, I will argue that Fox and Reinhart are correct to link Dahl's paradigm to Strong Crossover. The key to a unified analysis is the generalization of coreference to covaluation proposed in Reinhart (2006), Heim (1998, 2007). Certain features of Reinhart's (2006) analysis, in particular its reliance on a transderivational economy constraint, have been shown to give rise to inconsistencies (Roelofsen 2010). However, Reinhart's basic insight can be captured within a theory containing a fairly simple strong crossover constraint which makes no reference to alternate derivations.

1. Introduction

Dahl (1973, 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.

Fox (2000) and Reinhart (2006) each present influential analyses of Dahl's paradigm. These analyses have two key points in common. First, they are both compatible with a strict parallelism constraint on VP ellipsis.[1] Second, 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's analysis crucially depends on a relaxed parallelism requirement, and he argues that Dahl's paradigm and SCO derive from separate constraints. An important contribution 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 one pronoun is bound across another with the same referential value, it does not block (3b):

(3) a. John thinks that he$_{=\mathrm{John}}$ 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 Strong Crossover (SCO) violation such as (4):

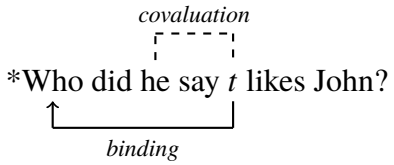
(4) *Who did he say *t* likes John?

Hence, if the constraint responsible for the Dahl effect permits (3b), it follows

[1]Fox (2000) 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 strict/sloppy ambiguities and Dahl's paradigm, see e.g. Dalrymple, Sheiber, and Pereira (1991), Kehler (1993), Fiengo and May (1994), Büring (2005a), Schlenker (2005), Kehler and Büring (2008).

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 analogous more to (3a) than to (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)



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

The main contention of this paper is that Fox and Reinhart are correct to link Dahl's paradigm to Strong Crossover, even though there are significant problems with their specific analyses. The majority of these problems derive from Fox's and Reinhart's use of transderivational economy constraints (Rule H and Rule I respectively). Neither Dahl's paradigm nor SCO phenomena provide any motivation for the use of constraints of this type. Rule H and Rule I were formulated as

[2]Reinhart (2006) 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 (λx (did he say x likes John))], and it is "too late" for *who* to bind *he* (Reinhart 2006, 173-174). In the present theory, the condition in (32) will play a similar role.

transderivational economy constraints for independent binding-theoretic reasons. The key component of my analysis is a formulation of a Strong Crossover constraint which accounts for both standard instances of SCO and Dahl's paradigm. This constraint is not transderivational but is similar in spirit to Rule H. In effect, it is a "local" reformulation of Rule H, analogous to local reformulations of Shortest Move in the mid 90s.

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 represent LFs in the remainder of the paper. Section 5 introduces my formulation of the SCO constraint and presents further evidence to support the hypothesis that SCO is responsible for the unavailability of (2d). Section 7 argues that pronominal binding shows weak crossover effects as well as SCO effects. Section 6 concludes with some remarks on the parallelism constraint on VP ellipsis, addressing arguments against strict parallelism based on certain ellipsis phenomena.

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)

[3]Italics in original.

When both of the pronouns in (1) are interpreted as bound variables, Rule H has the effect of ensuring that only transitive binding — (7a) — is possible in the first conjunct. Co-binding — (7b) — is blocked because it involves a longer dependency:

(7) a. John knows that he loves his mother.

b. *John knows that he loves his mother.

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: 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)–(8c):

(8) a. John knows that he$_{=\text{John}}$ loves his mother.

b. John knows that he loves his$_{=\text{John}}$ mother.

c. John knows that he$_{=\text{John}}$ loves his$_{=\text{John}}$ mother.

All of the LFs in (8) are blocked by Rule I of Grodzinsky and Reinhart (1993), and at times, Fox presents Rule H as a reformulation of Rule I (Fox 2000, 124fn14). On the other hand, Büring (2005a) 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, the VPs in (9a)–(9c) satisfy parallelism but the VP in (9d) does not. Fox's definition of parallelism is reproduced in (10):

(9) a. Bill [$_{\mathrm{VP}}$ knows that he loves his mother].

b. Bill [$_{\mathrm{VP}}$ knows that he$_{=\mathrm{John}}$ loves his$_{=\mathrm{John}}$ mother].

c. Bill [$_{\mathrm{VP}}$ knows that he loves his$_{=\mathrm{John}}$ mother].

d. Bill [$_{\mathrm{VP}}$ knows that he$_{=\mathrm{John}}$ loves his mother].

(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 indentical dependencies (Structural Parallelism).

Since (9a)–(9d) correspond to (2a)–(2d), Fox's analysis correctly predicts the availability of the interpretations glossed in (2a)–(2c), 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)–(8c) are in fact available in addition to (7a), then readings (2a)–(2c) 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 co-binding yields a distinct interpretation from

transitive binding:

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

We will see in section 6 that Roelofsen (2011) has presented strong evidence that certain co-binding configurations are quite generally permitted. 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 six possible LFs for the antecedent, as shown in (12).[4] Of these, four — (12a)–(12d) — are licit, but (12b) and (12c) yield indistinguishable interpretations under ellipsis. Thus, there are three possible interpretations of the elided VP corresponding to (2a)–(2c):

(12) a. John knows that he loves his mother.

b. John knows that he$_{=\text{John}}$ loves his$_{=\text{John}}$ mother.

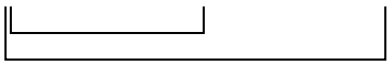
c. John knows that he$_{=\text{John}}$ loves his mother.

d. John knows that he loves his$_{=\text{John}}$ mother.

e. *John knows that he$_{=\text{John}}$ loves his mother.

[4] See Reinhart (2006, 194).

f. *John knows that he loves his mother.



Readings (2a)–(2c) can be obtained simply by copying the antecedent VP in (12a)–(12d) over to the ellipsis site at LF. The unavailable (2d) cannot be derived from any of (12a)–(12d), since Reinhart assumes a strict parallelism requirement.[5] The LFs in (12e)–(12f) 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)–(12f), since as Roelofsen (2010) points out, Reinhart's analysis appears to be internally inconsistent. 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:

(13) *Rule I* (Reinhart 2006)

α and β cannot be covalued 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 covalued interpretation is indistinguishable from what would be obtained if α binds β.

We will discuss the definition of covaluation shortly, but for now all we need is the fact that α and β are covalued if they are coreferential or if α is bound by an expression with which β is coreferential. 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 covaluation relations are added after bound pronouns have been translated as variables via

[5] On 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.

λ-abstraction. If *his* in (12e) is interpreted as a bound pronoun, then (12e) has the following interpretation:

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

Suppose that we now attempt to set *he* covalued 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 & $y =$ 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 (λx (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* = John, so that *he* and *his* are covalued, is indistinguishable from the interpretation which would be obtained if *he* bound *his* (i.e. [John$_1$ (λx (x said that he$_1$(λy (y loves y's mother))))]). 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 SCO:

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

Reinhart assumes that only the trace of *who* can be interpreted as a variable bound by *who* (see footnote 2). 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.

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 (λx (did he say x was intelligent & he $= x$))
b. Who (λy (y (λx (did y say x was intelligent))))

In effect, the additional λ-operator in (19b) adds a counterfeit of x — an expression which always has the same value as x but which is not formally identical to it. To keep track of which variable is a counterfeit of which, it is useful to write the counterfeit of x as x', so that (19b) is written as (20):

(20) Who (λx' (x' (λx (did x' 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 the SCO configuration 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.[6] Considering Dahl's paradigm and standard instances of SCO together, we see that 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 covaluation.

We saw above that Fox's motivations for formulating Rule H as a trans-derivational economy condition had nothing to do with Dahl's paradigm itself. The same holds for Reinhart's formulation of Rule I in this manner. Like the original Rule I of Grodzinsky and Reinhart (1993), Reinhart's modified formulation is is designed to account for the Condition B and C obviation phenomena discussed in Reinhart (1983).[7] Heim (2007) has recently argued that we should opt for a non-Rule-I-based analysis of these data. If we go this route, then it is quite possible to formulate a local constraint which blocks (12e). This will be the aim of section 5.

In contrast to Fox (2000), the account of SCO in Reinhart (2006) does not depend on a ban on co-binding. On Reinhart's assumptions it is simply impossible for an A′-trace and a pronoun to translate as formally identical variables, so the issue of whether co-binding is permitted is moot as far as SCO is concerned.

[6]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.

[7]That is, examples such as "As for John$_1$, Mary likes him$_1$, Bill likes him$_1$ — even John$_1$ likes him$_1$!"

The basic intuition behind Reinhart's analysis of SCO can therefore be maintained in the face of Roelofsen's evidence that certain co-binding configurations are permitted. This evidence will be presented in section 6.

4. Notation

In what follows, it will be 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, 2005b) β-notation can be adapted to this end. In Büring's notation, coreference is indicated via coindexation:

(21) John$_1$ thinks that he$_1$ is intelligent.

Variable binding is represented by the addition of a β bearing its own index. In (22a), *John* binds *he*; in (22b), *John* binds *he* and *John* and *his* are coreferential (hence covalued):

(22) a. John$_1\beta_2$ thinks that he$_2$ is intelligent.

b. John$_1\beta_2$ said that he$_2$ loves his$_1$ mother.

I will show that an expression is covalued with a bound variable via the introduction of an additional β-node. This mirrors the additional λ-operator in examples such as (20):

(23) a. Every boy$\beta_2\beta_1$ said that he$_1$ loves his$_2$ mother.
= R's [E.b. ($\lambda x'(x'(\lambda x$ (x said that x loves x' 's mother)))]

b. Who$\beta_2\beta_1$ t_1 said that he$_2$ likes Mary?
= Reinhart's [Who ($\lambda x'(x'(\lambda x$ (x said that x' likes Mary))))]

Since Reinhart is not explicit regarding the process responsible for introduction of the additional λ-operator, it may be helpful to consider how such additional operators could be introduced via a modified version of Heim's (1998) double indexing system. The relevant modification is to permit expressions to bear multiple outer indices so that an expression may move multiple times to bind multiple distinct variables. The pattern of binding dependencies in (23a) can then be derived by QRing *every boy* twice:

(24) [Every boy]$^{[1,2]}$ said that he$_1$ loves his$_2$ mother.
[Every boy]^2 [λ_1 [t_1 [said that he$_1$ loves his$_2$ mother]]].
[Every boy] [λ_2 [t_2 [λ_1 [t_1 [said that he$_1$ loves his$_2$ mother]]]]].
(Superscripts are outer indices, subscripts are inner indices.)

The pattern of binding dependencies in (23b) can be derived via QR of the *wh*-phrase following *wh*-movement:

(25) Who$^{[1,2]}$ said that he$_2$ likes Mary?
Who^2 [λ_1 [t_1 said that he$_2$ likes Mary]]
Who [λ_2 [t_2 [λ_1 [t_1 said that he$_2$ likes Mary]]]]

The β-notation can therefore be understood as an abbreviation for the corresponding Heim-style LFs.

Links will sometimes be added to LFs to make them more readable. These links always indicate variable binding and never encode any information which the β-notation does not. The LF in (22b), for example, might be rendered as (26):

(26) John$_1$⌊β_2 said that he$_2$ loves his$_1$ mother

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

(27) a. John$_1$⌊β_2 knows that he$_2$⌊β_3 loves his$_3$ mother.

b. John$_1$⌊β_2 knows that he$_2$ loves his$_2$ mother.

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. In particular, Heim's (1998) definition of "codetermination" can be adapted for LFs written using β notation.[8] However, we must take into consideration the fact that SCO effects are triggered by overlapping valuation as well as covaluation:

(28) *Which boy⌊β_2⌊β_1 do they$_{\{2,3\}}$ think t_1 is intelligent.

[8]Heim (1998, 233)'s 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) A is co-indexed with a β-node node associated with B, 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. In fact the anti-A-binding condition turns out to be redundant in the definition of SCO given in (31), and it would therefore be sufficient to take (i) as a definition of covaluation. In more recent work, Heim has offered a somewhat more explicit version of Reinhart's semantic definition of covaluation Heim (2007, 7).

Reinhart's model-theoretic definition of covaluation in (18) is easily adapted:

(29) *Overlapping valuation*

A and B overlap in value iff neither A-binds the other and for all assignments, given the values x of A and y of B, there is a z such that $z \preceq x$ and $z \preceq y$.[9]

Overlapping valuation is not, however, easily defined as a formal relation over LFs. Simple cases can be accommodated by replacing indices with index sets. Insofar as this works, appropriate modifications to the definition in footnote 8 are indicated in the following material. However, we will see several examples overlapping valuation relations which are not easily understood in these terms (e.g. (57)). Heim (2007) notes that there is a price for enriching LF representations so that co-binding, transitive binding, covaluation etc. are all explicitly distinguished: it becomes more difficult to give a purely formal statement of Conditions B and C. If there is evidence that LFs explicitly distinguish all of the aforementioned relations, then the model-theoretic definition of overlapping valuation will be the most straightforward and should perhaps be preferred.

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. This variant depends on formulation of an SCO constraint that is more permissive than Rule H and Rule I in one key respect: it permits co-binding LFs. As we will see in section 6, Roelofsen (2011)

[9]This definition assumes a theory of plurals along the lines of Link (1983)

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 to rule out are exemplified in (30a)–(30c):

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

b. *Who$\beta_2\beta_1$ did he$_2$ say t_1 is intelligent?

c. *Every boy$\beta_2\beta_1$ said that he$_2$ loves his$_1$ mother?

If pronouns interpreted as bound variables were linked to their antecedents via movement then (30a) would immediately fall together with (30b), since the second pronoun in (30a) would then have exactly the status of the A′-trace in (30b).[10] Implementing this approach would take us on something of a detour, however, so I propose here a more conservative definition of a suitable Strong Crossover constraint:

(31) *Strong Crossover*

*[$_\gamma$ β_ι … A_κ … B_ι] (where γ immediately dominates[11] β_ι) if

(i) β_ι c-commands A_κ c-commands B_ι,

(ii) A_κ and B_ι overlap in value, and

(iii) A_κ is not bound within γ.

The condition in (31) immediately rules out both (30a) and (30b). In (30a), *he*$_1$ and *his*$_2$ overlap in value and *he*$_1$ is not bound (hence not bound within the minimal constituent properly containing β_1, the binder of *his*$_2$). In (30b), he$_2$ and t_1 overlap in value and he$_2$ is bound by β_2, which is outside the minimal constituent properly containing the binder of t_1, β_1. The LF in (30c) is illicit for the same reason. In structures such as (30b) and (30c) the condition in (31) effectively

[10] See e.g. Drummond, Kush, and Hornstein (2011), Kayne (2002) for analyses of pronominal binding in terms of movement.

[11] XPβ_ι has the structure [XP [β_ι …]] and XP$\beta_\iota\beta_\kappa$ has the structure [XP [β_ι [β_κ …]]].

functions as a ban on crossing binding dependencies. This way of thinking about the constraint also extends to (30a) if we imagine that he_1 is bound by an operator at the top of the structure.

Note that whereas licit LFs can be derived from (30a) and (30c) by changing the index of *he* from 2 to 1, no licit LF can be derived from (30b) by changing the index of *he* from 2 to 1. That is, if the constraint in (31) permits co-binding and is taken to be responsible for standard SCO effects, then 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 variable bound by it). A key assumption of the present analysis 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 the analogy between standard cases of SCO and the Dahl effect. However, as the aforementioned contrast between (30a) and (30b) indicates, we must take note of the fact that an A′-trace jealously guards its special relationship with its antecedent, refusing to share it with a pronoun:

(32) If α is co-indexed with an A′-trace, α is a trace or a β-node.

It is also crucial that *wh*-movement precede pronominal binding, in the sense that the β-node which binds the trace of a *wh*-phrase must be lower than any other β-nodes associated with it. That is, we must not permit the LF derived by reversing the order of the β-nodes in (30b). If introduction of a second β-node corresponds to QR, as shown in (25) above, then the impossibility of such LFs follows immediately. QR must follow overt *wh*-movement, so the additional λ introduced by QR must be higher than the λ introduced by *wh*-movement.

We now have a formulation of the crossover analysis of Dahl's paradigm which incorporates Reinhart's insights regarding covaluation, but which obviates Roelofsen's (2010) critique of Reinhart's formulation of Rule I. A key dif-

ference between (30a) and (30b) is that only the latter shows an overt distinction between the structure which violates SCO and the structure which does not. This makes it somewhat more difficult to probe the nature of the violation in (30b). Indeed, we have yet to see any direct evidence that (30b) really is an SCO violation. The following subsections present three pieces of evidence for an SCO constraint on pronominal binding. 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 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. 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. The sentence in (33), for example, is ambiguous. It has one reading in which the second *I* is interpreted as a true indexical — (34a) — and another in which the second *I* is a "fake indexical"[12] interpreted as a variable bound by the first *I* — (34b):

(33) Only I think that I'm intelligent.

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

[12]On fake indexicals, see e.g. Kratzer (2009, 1998), Heim (1994, 2005), Rullmann (2004). Examples along the lines of (33) were first discussed in Partee (1989, fn3).

b. I am the only x such that x thinks x is intelligent.[13]

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. The sentence in (35) can have only the readings glossed in (36a)–(36c):

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

(36) 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 would take x's car.

The data in (35)–(36) seem amenable to analysis either in terms of Rule H or crossover. If binding in (35) is evaluated "locally" before *only* makes its contribution to the interpretation, then the LF deriving (36d) will be blocked because it is interpretatively indistinguishable from the LFs deriving (36a)-(36c) and has a longer binding dependency. Under the crossover analysis, the LF deriving (36d) is blocked because the third pronoun is bound across the covalued second pronoun. To distinguish the Rule H account of the pattern in (36) from the crossover account, we can exploit the fact that fake indexicals can be "partially bound" (Rullmann 2004, Heim 2005). For example, (37a) can have the reading glossed in (37b):

(37) a. Out of everyone who's baked with Peter
…only I like our banana muffins.

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

In this light consider (38):

[13](34b) 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.

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

(38) cannot have the interpretation glossed in (39) where *we* and the second *I* each covary with the first *I*:

(39) I am the only x such that [x and Peter] think x is a great cook.

Such a reading would be available only if it were possible for the second and third pronouns in (38) to be co-(partially-)bound by the first.[14] Thus, even before *only* makes its contribution to the interpretation, the co-binding LF for (38) encodes an interpretation which cannot be encoded by any competing LF. Rule H therefore fails to block co-binding in (38), and incorrectly predicts that the interpretation glossed in (39) should be available. Does the crossover analysis fare any better in this instance? Recall the Strong Crossover constraint defined in (31), repeated here in (40):

(40) *Strong Crossover*
*[$_\gamma$ β_ι … A_κ … B_ι] (where γ immediately dominates[15] β_ι) if
(i) β_ι c-commands A_κ c-commands B_ι,
(ii) A_κ and B_ι overlap in value, and
(iii) A_κ is not bound within γ.

The co-binding LF for the relevant portion of (38) is as follows:

[14]Pronouns can be partially bound, but there are no "partial binders," so transitive binding is presumably impossible in (38). 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.

[15]XPβ_ι has the structure [XP [β_ι …]] and XP$\beta_\iota\beta_\kappa$ has the structure [XP [β_ι [β_κ …]]].

(41) *$I_1\beta_2$ said $we_{\{2,3\}}$ think I_2'm a great cook.

If we understand contraindexation as non-identity of index sets, then *we* and the second *I* are overlap in value and *I* is not bound within [β_2 ...], so that the Strong Crossover constraint is violated. Thus, the crossover analysis correctly rules out interpretation (39).

We do not expect to find a crossover violation if the order of the partially and fully bound fake indexicals in (41) is reversed. The Strong Crossover constraint permits the second *I* to partially bind *we* in (42):

(42) $I_1\beta_2$ said $I_2\beta_3$ think $we_{\{3,4\}}$'re great cooks.

As expected, (43) can receive the reading corresponding to (42):

(43) Out of everyone who's baked with Peter
...only I said I think we're great cooks.

We therefore see that the crossover analysis correctly predicts the range of available readings for (38) and (43).

Before moving on, it may be worth noting that the present formulation of Strong Crossover correctly permits binding in acceptable examples such as (44):

(44) Every boyβ_1 persuaded every girlβ_2 that $they_{\{1,2\}}$ should meet his_1 parents.

Here, *every boy* binds *his* over another pronoun, *they*, which overlaps in value. However, since *they* is (with respect to both of its indices) bound within [β_1 ...], there is no violation.

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 (45), there is no way of telling which pronoun is the true resumptive, and which is simply a pronoun receiving a covarying interpretation:[16]

(45) $Wh_1 \ldots \text{ProA}_1 \ldots \text{ProB}_1$

This is because (45) can always be parsed with ProA as the resumptive, so that there is no way of telling whether an SCO violation would arise if ProB 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 (45) 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 should be incurred if the epithet c-commands the pronoun. This prediction is indeed borne out — (46b) is unacceptable in Irish:

(46) a. $Wh_1 \ldots \text{Pro}_1 \ldots \text{Ep}_1$.

b. $*Wh_1 \ldots \text{Ep}_1 \ldots \text{Pro}_1$.

In principle, it should be possible to use a similar trick with (47) to exclude the licit parses in (48a)–(48b):

[16] 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.

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

(48) a. Every boyβ_1 said that he$_1$ loves his$_1$ mother.

b. Every boy$\beta_2\beta_1$ said that he$_1$ loves his$_2$ mother.

c. *Every boy$\beta_2\beta_1$ said that he$_2$ loves his$_1$ 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 Strong Crossover violation which masks the violation we wish to isolate. In (49), 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:[17]

(49) *Every boyβ_1 said that [the x_1 little brat]$_2$ loves his$_1$ mother.

This problem is easily addressed. Since Strong Crossover violations are (in contrast to variable binding relations) conditioned on strict c-command, we need only embed the antecedent slightly:[18]

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

[17]For concreteness, I assume in (49) that covariation is effected via binding of a variable x within the epithet. The key point is that the epithet is not itself a variable bound by *every boy*.

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

The resulting sentence nonetheless fails to instantiate a crossover violation. This is because (50) 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*:

(51) Every boyβ_1's mother thinks that [the x_1 little brat]β_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 quantifier such as *every bride*, but *the happy couple* is not a possible antecedent for the pronoun *she*:

(52) *The happy couple said that she was going to be late.

If we replace the epithet in (50) 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 (53a), it is difficult to obtain a reading in which *the happy couple* covaries with the quantifier. In contrast, antecedence is fully acceptable in (53b), since the pronoun does not have to "reach over" the epithet:

(53) 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 (53a) leads to an SCO effect when the bound pronoun is replaced by a *wh*-trace:

(54) 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 (53a) 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 (53a) and (55):

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

There is a parallel contrast between (54a) and (56):

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

This reinforces the connection between the effect in (53a) and standard instances of SCO. Rule H cannot account for the deviance of (53a), since there is no alternative interpretatively-identical LF in which the pronoun is bound by a closer antecedent.

5.3. Crossover effects triggered by DPs embedded in QPs

A shown in (57), it is sometimes possible for a DP embedded inside a *wh*-phrase to trigger an SCO effect:

(57) 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 (58):

(58) 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 (57b) and with *he* in (58b).[19] The pronoun *they* cannot be bound as a variable in (57b) or (58b) since *the boys* is too deeply embedded to bind it. In (57b) the trace must be bound as a variable by the *wh*-phrase. In (58b) *he* must be bound as a variable by *each of the boys* if it is to receive a covarying interpretation. Taking *they* as A and *t*/*he* as B, Strong Crossover is therefore violated in both cases:

(59) a. *[Which of [the boys]$_1$]β_2 did they$_2$ say t_2 would win?

b. *[Each of [the boys]$_1$]β_2 said that they$_1$ think he$_2$ is intelligent.

Rule H cannot account for the deviance of (57b) or (58b). Although *they* is a closer potential antecedent for *t* in (57b) and for *he* in (58b), binding the *t*/*he* by *they* would not derive an LF with the same interpretation as (57b)/(58b).

6. 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

[19] This is clear from the semantic definition of overlapping valuation in (29). It is less clear that overlapping valuation in this particular instance could be captured by the Heim-style definition proposed in footnote 8. Examples of this sort may constitute an argument that a semantic definition of overlapping valuation is to be preferred.

the approach of Fox and Roelofsen.[20] 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 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.[21] 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.

6.1. Reverse Dahl effects

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

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

(61) a. ...Bob claimed Bob called Max's mother.

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

To obtain the reading (61b) for the second conjunct of (60) without violating parallelism, the pronoun in the second conjunct would have to be bound by the

[20] While Fox (2000) argues against a strict parallelism constraint on VP ellipsis, his analysis of Dahl's paradigm is compatible with a strict parallelism constraint (see also footnote 1).

[21] 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.

first instance of *Bob*. This violates Rule H:[22]

(62) LFs for second conjunct of (60)

a. Bob_1 claimed $\text{Bob}_1\beta_2$ called his_2 m. (Violates parallelism)

b. $\text{Bob}_1\beta_2$ claimed Bob_1 called his_2 m. (Violates Rule H)

The formulation of the crossover analysis presented here also blocks reading (61b). We can obtain this reading only by assigning the LF in (63a) to the first conjunct, which yields the LF in (63b) for the second:

(63) a. *First conjunct*

$\text{Max}_1\beta_2$ claimed that Bob_3 called his_2 mother.

b. *Second conjunct*

$\text{Bob}_3\beta_4$ claimed that Bob_3 called his_4 mother.

(63b) is, of course, a Strong Crossover violation under the definition in (31). So far, there is nothing to choose between Fox's analysis and the present analysis. However, Roelofsen (2011) points out that examples such as (64) are problematic for Fox:

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

[22] Both (62a) and (62b) violate Condition C. However, this violation can be overcome via vehicle change (Fiengo and May 1994) (i.e., by replacing the second instance of *Bob* with a coreferential pronoun). Vehicle change is independently motivated by acceptable examples such as (i), which becomes (ii) following vehicle change:

(i) John likes Bill_1's mother, and Bill_1 does like Bill_1's mother too.

(ii) John likes Bill_1's mother, and Bill_1 does like his_1 mother too.

As Roelofsen notes, (64) has a reading on which every boy added that he loved his own dish. Given Fox's theory, candidate LFs for deriving this reading are as follows:

(65) a. Every boyβ_1 claimed that the jury loved his$_1$ dish

and added that he$_1\beta_2$ did l. his$_2$ dish too.

b. Every boyβ_1 claimed that the jury loved his$_1$ dish

and added that he$_1\beta_2$ did l. his$_2$ dish too.

Neither LF is in fact available, since (65a) violates parallelism and (65b) violates Rule H (there is co-binding in the second conjunct). On the present account, however, co-binding is not blocked and so (65b) is correctly predicted to be available. Roelofsen also has an explanation for the availability of (65b). He develops an ingenious alternative to Rule H, "Free Variable Economy", which also permits co-binding in (65b). FVE is defined in (66)–(68):

(66) *Free Variables*[23]

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.

[23]The definition in (66) has been modified 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."

(67) *Economy measure*

Let Σ and Π be alternatives.[24] Then we say that Π is more economical than Σ iff some subconstituent Π' of Π contains fewer free variables than the corresponding subconstituent $\Sigma\prime$ of Σ.

(68) *Free Variable Economy (FVE)*

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

Owing to the manner in which (66) defines free variables in terms of indices, FVE permits both co-binding and transitive binding in simple cases, as shown in (69a)–(69b). However, it does not permit LFs such as (70a), 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 (70b), for example, because the subconstituent [called his_2 mother on his_1 birthday] contains more free variables according to (68) than the corresponding subconstituents of the alternative derivations in which *her* is bound by *his* or *he*. Examples of LFs permitted by FVE are shown in (69) and examples of LFs blocked by FVE in (70):

(69) *Permitted by FVE*

a. $\text{Everyone}\beta_1$ said he_1 called his_1 mother.

b. $\text{Everyone}\beta_1$ said $\text{he}_1\beta_2$ called his_2 mother.

[24] Two LF constituents are alternatives iff they are (i) semantically equivalent, and (ii) formally identical modulo binding indices on pronouns.

(70) *Blocked by FVE*

a. John$_1$|β_2 said he$_1$ called his$_2$ mother.

b. Everyoneβ_1 said he$_1$|β_2 called his$_2$ m. on his$_1$ b.day.

That FVE blocks (70a) 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 (64)–(65), we can now see that FVE permits co-binding in (65b), as is required to derive the relevant interpretation of (64).

6.2. Can we keep parallelism?

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 straightforwardly derived *if* one assumes that co-binding is quite generally available.[25] 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

[25] 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.

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 (71):

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

As he observes, "[(71)] has a 'mixed' reading, on which every worker x said that the boss knows when he, the boss, can take home x's tools." The two logical forms which could on Roelofsen's assumptions in principle derive this reading are illustrated in (72):

(72) a. E.w.β_1 said he$_1\beta_2$ knows when he$_2\beta_3$ can tk. hm. his$_3$ tools,

and TBβ_4 does k. when he$_4$ can tk. hm. his$_1$ tools too.

b. E.w.β_1 said he$_1\beta_2$ knows when he$_2$ can tk. hm. his$_1$ tools,

and TBβ_3 does k. when he$_3$ can tk. hm. his$_1$ tools too.

The problem for Fox and Roelofsen is that (72a) violates parallelism, whereas (72b) violates Rule H and FVE. Once again, the crossover analysis derives the relevant reading of (71) unproblematically, since (72b) does not violate the formulation of Strong Crossover in (31).

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 begin-

ning of the sentence, interpret the first *he* referentially, and set *his* coreferential with the first *he*. In contrast, Roelofsen proposes to replace parallelism with a less stringent requirement (Focus Match) to accommodate the relevant reading of (71). This requirement is satisfied in LFs like (73):

(73) Every studentβ_1 thinks that he$_1\beta_2$ loves his$_2$ mother

and every prof.β_3 does t. that he$_3$ loves his$_3$ mother too

The details of Focus Match 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 without any relaxation of the parallelism requirement.

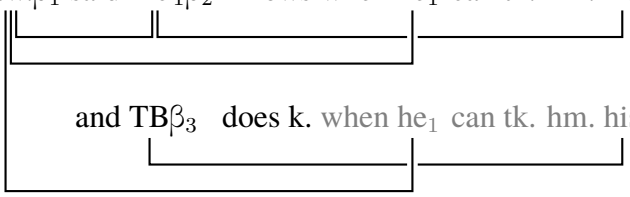
Roelofsen discusses one last example which motivates a further complication. This time, the problem is one of overgeneration. As it stands, Roelofsen's theory predicts that (71), repeated here as (74a), should permit the reading for the second conjunct glossed in (74b):

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

b. Every worker x said that the boss knows when x can take home the boss's tools.

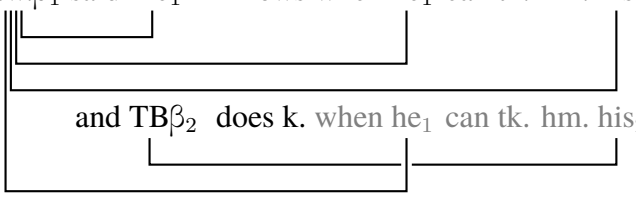
Since (74a) does not in fact have such a reading, Roelofsen proposes to add an additional constraint on VP ellipsis, Preserve Coindexing. Again, I do not wish to go into the details of Roelofsen's analysis here, but only to note that the absence of the relevant reading follows immediately on the crossover analysis. This reading could in principle derive from either of the following LFs:

(75) a. E.w.β_1 said he$_1\beta_2$ knows when he$_1$ can tk. hm. his$_2$ tools,



and TBβ_3 does k. when he$_1$ can tk. hm. his$_3$ tools too.

b. E.w.β_1 said he$_1$ knows when he$_1$ can tk. hm. his$_1$ tools,



and TBβ_2 does k. when he$_1$ can tk. hm. his$_2$ tools too.

The LF in (75b) violates parallelism, while the LF in (75a) violates the formulation of Strong Crossover in (31).[26]

7. 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.

[26]The last two pronouns in the first conjunct of (75a) overlap in value according to the definition in (29). Since β_2 c-commands *he* c-commands *his*, *he* overlaps in value with *his* and *he* is not bound with [β_2 ...], the Strong Crossover constraint is violated. Note that *he* and *his* are also covalued according to the definition in footnote 8. With regard to the definition in footnote 8, 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 (75a) is codetermined with *his*. We can (redundantly) verify that neither pronoun A-binds the other, so the pronouns are covalued.

7.1. WCO effects triggered by DPs embedded in QPs

The examples in this subsection are constructed simply by embedding the relevant pronoun in the examples from section 5.3. Like SCO effects, weak crossover effects can also be triggered by DPs embedded inside a *wh*-phrase — (76). There is a parallel constraint on pronominal binding — (77):

(76) a. [Which of [the boys]$_1$]β_2 t_2 said their$_1$ mother would win?

b. ?[Which of [the boys]$_1$]β_2 did their$_1$ mother say t_2 would win?

(77) a. [Each of [the boys]$_1$]β_2 knows that his$_2$ mother loves them$_1$.

b. ?[Each of [the boys]$_1$]β_2 knows that their$_1$ mother loves him$_2$.

7.2. WCO effects triggered by embedding the first pronoun

If variable binding is subject to WCO, we expect the SCO violation in the LF for (2d) to become a weak crossover violation if the first pronoun in (1) is further embedded, as in (78). This is by analogy with (79b):

(78) John thinks all of his friends love his mother and Bill does too.

(79) 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 (78) 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 and May (1994) and Kehler (1993)) points to examples such as (80), which he reports as permitting the interpretation of the elided VP glossed in (81):

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

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

Fiengo and May (1994, 156) provide the following example, which they claim makes the (81)-type reading reasonably accessible:

(82) I thought my 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 (81) far less accessible than the other three readings glossed in (83a)–(83c):

(83) 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 (82) extremely awkward under the relevant reading. Thus, even if (81) is more accessible than (2d), we still need an account of why it is less accessible than (83a)–(83c).[27] Under the crossover analysis of Dahl's paradigm, these facts are readily explicable. The reading in (81) requires a WCO violation, and is therefore less accessible than (83a)-(83c). The reading in (2d) requires an SCO violation, so it less accessible than both (81) and (83a)–(83c).

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 (84) from Sag (1976):

(84) Edith said that finding her husband nude had upset her, and Martha did

[27]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.

too.

(85) 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.

Kehler attempts to explain the absence of (85d) by appealing to special properties of experiencer verbs such as *upset*. Following Fiengo and May (1994, 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 (87) as in (85), but there is no experiencer verb or PRO in (86):

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

(87) 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 entirely. Rather, it ameliorates it by replacing an SCO violation with a WCO violation. To show that Rule H cannot account for the preceding data 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 second as a variable.[28] To decide between Fox's analysis and the crossover analysis, we must determine whether or not the following generalization holds:

(88) *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 (88) is not entirely straightforward. As Shan and 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:

(89) *[The teacher who liked every student$_1\beta_2$] praised him$_2$.

The "weak" form of the Dahl effect is nonetheless manifested in (90)–(91):

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

(91) a. …Bill said that the t. who liked B. gave B. extra homework.

[28] With regard to (86), many speakers at least marginally allow a sloppy reading for (i) and binding in (ii):

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

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

b. …Bill said that the t. who liked J. gave J. extra homework.

c. …Bill said that the t. who liked B. gave J. extra homework.

d. ??…Bill said that the t. who liked B. gave J. extra homework.

It seems, then, that the generalization in (88) does not hold. This supports the crossover analysis of Dahl's paradigm, which predicts that reading (91d) should be difficult to access due to WCO.

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 LF copying of its antecedent. There are, however, a number of arguments in the literature against strict parallelism. 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, (92a) can have the reading glossed in (92b):

(92) 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 (93):

(93) *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 indentical dependencies (Structural Parallelism).

It would not be possible to adopt (93) in conjunction with the crossover analysis, since (93) permits the co-binding LF in (94a) to license the elided VP in (94b):

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

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

Thus, examples such as (92) 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 and 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 (92a) glossed in (92b) has the LF in (95):

(95) 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. I will depart here from Fiengo & May's technical implementation.

Let us begin by noting that bound variable interpretations are subject to some structural constraint. That is, in order for the pronoun in (96) to be interpreted as a variable bound by the DP, some condition on the structural configuration represented by the ellipsis must be met:

(96) $\text{DP}_1\beta_2 \ldots \text{pro}_2$

Call the relevant structural relation S, and assume (i) that 'a c-commands b' implies '$\mathrm{S}(a, b)$' and (ii) that S holds between *Homer* and *his* in the third line of (95). It is an interesting feature of the β-notation that we have two ways of constraining binding relations via S. That is, using (96) as an example, we can either require that $\mathrm{S}(\beta_2, \text{pro}_2)$ or that $\mathrm{S}(\text{DP}_1, \text{pro}_2)$. 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):

(97) $[_{\text{XP}}\ \text{DP}_1\beta_2 \ldots] \ldots [_{\text{YP}}\ \text{DP}_1 \ldots \text{pro}_2]$

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 LFs in (98):

(98) $\text{Smithers}_1\beta_2$ thinks that his_2 job sucks.
$\text{Homer}_3\beta_4$ 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, this pronoun does stand in S to a DP co-indexed with its antecedent. (The antecedent of the pronoun is the first instance of *Homer*, and the pronoun 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 (98) is licit. The following definitions accomplish this:

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

(100) *Constraint on bound variable interpretation*
B can be 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 $\mathrm{S}(C, B)$

The preceding definitions implement a variant of Fiengo & May's analysis of (92). 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 (92a) glossed in (92b). This suggests that a somewhat looser relation than covaluation may be able to license variable binding interpretations via (100). This is consistent with the data in (101)–(102). (102a) has a reading under which each captain's first mate shares each captain's fear regarding the fate of the relevant ship. (102b) — 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). By contrast, (102c) has no such reading:

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

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

b. ?The first mate does too.

c. *John does too.

It is not entirely clear how the constraint in (100) should be modified to account for (102b). However, (102b) 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 because 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 (102b). Rather, this example suggests an explanation in terms of a modified form of (100). That is, although the pronoun in the elided VP of (102b) 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 (92) and (102b) 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. The constraint responsible for strong crossover effects need not be formulated as a transderivational economy condition. With regard to Dahl's paradigm, the crossover analysis has a number of advantages over analyses stated in terms of Rule H, Rule I or Free Variable Economy. It is also directly supported by data showing that pronominal binding triggers both weak and strong crossover effects.

References

Büring, Daniel. 2005a. Bound to bind. *Linguistic Inquiry* 36:259–274.

Büring, Daniel. 2005b. *The syntax and semantics of binding theory*. Cambridge University Press.

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

Dahl, Östen. 1974. How to open a sentence: Abstraction in natural language. In *Logical grammar reports*, volume 12. University of Götenburg.

Dalrymple, Mary, Stuart M. Sheiber, and Fernando C. N. Pereira. 1991. Ellipsis and higher-order unification. *Linguistics and Philosophy* 14:399–452.

Drummond, Alex, Dave Kush, and Norbert Hornstein. 2011. Minimalist construal: Two approaches to A and B. In *The Oxford handbook of linguistic minimalism*, ed. Cedric Boeckx, 396–426. Oxford: Oxford University Press.

Fiengo, Robert, and Robert May. 1994. *Indices and identity*. Cambridge, Massachusetts: MIT Press.

Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, Massachusetts: MIT Press.

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

Heim, Irene. 1994. Puzzling reflexive pronouns in *de se* reports. Unpublished handout, March 1994.

Heim, Irene. 1998. Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach. In *The interpretive tract (25)*, ed. Orin Percus and Uli Sauerland, 205–246. Cambridge, Massachusetts: MIT, Department of Linguistics.

Heim, Irene. 2005. Features on bound pronouns. Unpublished ms., MIT.

Heim, Irene. 2007. Forks in the road to Rule I. In *Proceedings of NELS 38*, ed. M Abdurrahman, A Schardl, and M Walkow.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Malden, MA: Blackwell.

Hornstein, Norbert. 1995. *Logical Form: from GB to Minimalism*. Cambridge, Massachusetts: Basil Blackwell.

Kayne, Richard S. 2002. Pronouns and their antecedents. In *Derivation and explanation in the minimalist program*, 133–166. Blackwell.

Kehler, Anderew, and Daniel Büring. 2008. Be bound or be disjoint! In *NELS 38*, ed. Anisha Schardl, Martin Walkow, and Muhammad Abdurrahman, 487–501. CreateSpace.

Kehler, Andrew. 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. Saarland University.

Kratzer, Angelika. 1998. More structural analogies between pronouns and tense. In *Proceedings of SALT VIII*, ed. Devon Strolovitch and Aaron Lawson, 92–110. Ithaca, NY: CLC Publications.

Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40:187–237.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, use and interpretation of language*, ed. R. Baeuerle, C. Schwarze, and Arnim von Stechow. DeGruyter.

McCloskey, James D. 2011. Resumptive pronouns, A′-binding, and levels of representation in irish. In *Resumptive pronouns and the interfaces*, ed. Alain Rouveret, 65–120. John Benjanims.

Partee, Barbara. 1989. Binding implicit variables in quantified contexts. In *Papers from the 25th Regional Meeting of the Chicago Linguistics Society, Part One*, ed. Caroline R. Wiltshire, Randolph Graczyk, and Bradley Music, 342–365. Chicago Linguistics Society, Chicago Illinois: University of Chicago.

Reinhart, Tanya. 1983. Coreference and bound anaphora: A restatement of the

anaphora questions. *Linguistics and Philosophy* 6:47–88.

Reinhart, Tanya. 2006. *Interface strategies*. MIT Press.

Roelofsen, Floris. 2008. *Anaphora resolved*. University of Amsterdam, Amsterdam: Institute for Logic, Language and Computation.

Roelofsen, Floris. 2010. Condition B effects in two simple steps. *Natural Language Semantics* 18:115–140.

Roelofsen, Floris. 2011. Free Variable Economy. *Linguistic Inquiry* 42:692–697.

Rullmann, Hotze. 2004. First and second person pronouns as bound variables. *Linguistic Inquiry* 35:159–168.

Sag, Ivan. 1976. Deletion and logical form. Doctoral Dissertation, MIT.

Schlenker, Philippe. 2005. Non-redundancy: Towards a semantic reinterpretation of Binding Theory. *Natural Language Semantics* 13:1–92.

Shan, Chung-Chieh, and Chris Barker. 2006. Explaining crossover and superiority as left-to-right evaluation. *Linguistics and Philosophy* 29:91–134.

Shan, Chung-Chieh, and Chris Barker. 2008. Donkey anaphora is in-scope binding. *Linguistics and Philosophy* 29:91–134.