

Workshop on Direct Compositionality  
Brown University, June 2003

## The Interpretation of Quantificational Structures: Evidence for the Copy Theory of Traces

Danny Fox, MIT

### Introduction

(1) A girl read every book.

What is the source of Inverse Scope?

Two hypotheses:

- (1') a. **Movement** (Q-Raising, Q-Lowering, Quantifying-in, see Jacobson 2002)  
[A girl read  $t_1$ ] every book<sub>1</sub>
- b. **Function Composition** (Dowty, Jacobson, Steedman *passim*)  
[A-girl'  $\circ$  read'] every-book'

The debate is to some extent about the ways in which one should complicate phonology.

Illustration:

(2) A girl gave every book to John.

When extended to this sentence both hypotheses lead to the conclusion that phonology is deceitful:

- (2') a. [A girl gave  $t_1$  to John] Every book<sub>1</sub> (together with a Special rule of pronunciation: pronounce at the position of the trace)
- b. [A-girl'  $\circ$  gave'(to-John')]every-book' (together with a Special rule of pronunciation: pronounce adjacent to the head: wrap) .

Shared assumptions:

1. *every book* is interpreted as the second order predicate in (3a)
2. its argument is interpreted as the predicate that can be described (in the meta-language) by abstracting over the canonical object position, e.g. (3b) for the Inverse Scope reading of (1).

- (3) a.  $\lambda P_{\langle e, t \rangle}. \llbracket \text{book} \rrbracket \subseteq P$   
b.  $\lambda x_e. \text{A girl read } x$

**Claims:**

A. Shared assumption (1) needs to be modified. The argument of a restrictive quantifier phrase, QP, (at least when there is Inverse Scope) is a partial function defined only for individuals that satisfy the restrictor:

- (3'a) a.  $\lambda P_{\langle e,t \rangle}. \llbracket \text{book} \rrbracket \subseteq P$   
 b.  $\lambda x_e: x \text{ is a book. A girl read } x$

B. Inverse Scope is derived by Movement, and traces are copies.

- (2'') a. [A girl gave [every book]<sub>1</sub> to John] [every book]<sub>1</sub>

A “moved” QP is present at the base position as well as at the landing site.

The restrictor of the QP is accessed as part of the semantic rule for the interpretation of movement (predicate abstraction).

$\llbracket (2'')a \rrbracket = \llbracket \text{every book is such that a girl read that book} \rrbracket$

**Consequences for Direct Compositionality (DC):** See Conclusions

**1. The Interpretation of Chains**

- (4) Mary likes every boy.  $-QR--->$   
 [every boy]<sub>1</sub> [Mary likes [every boy]<sub>1</sub>].

**1.1. Two Syntactic Rules**

Assumption: copies must be replaced with syntactic phrases that contain variables

**Copy Insensitive Rule**

- (5) **Traditional Trace Theory:**  $DP \rightarrow x$   
 (6) every boy A girl talked to every boy.  $\rightarrow$   
 every boy  $\lambda x$  [a girl talked to  $x$ ]?  
 (7) Which boy Mary visited which boy?  $\rightarrow$   
 which boy  $\lambda x$  [Mary visited  $x$ ]?

The inspiration: pronoun binding

- (8) Every boy is such that a girl talked to him.

**Copy Sensitive Rule (Fox 1999, 2003; Sauerland in press; Elbourne 2002):**

**(9) Trace Conversion:**

- a. Variable Insertion (Det) Pred  $\rightarrow$  (Det) [Pred  $\lambda y$  (y=x)]  
 b. Determiner Replacement (Det) Pred  $\rightarrow$  the [Pred  $\lambda y$  (y=x)]

(10) every boy A girl talked to every boy.  $\rightarrow$   
 every boy  $\lambda x$  [a girl talked to **the boy x**]?

(11) Which boy Mary visited which boy?  $\rightarrow$   
 which boy  $\lambda x$  [Mary visited **the boy x**]?

(See Rullmann and Beck (1997))

The inspiration: binding of definite descriptions (e.g., epithets)

(12) Every boy is such that a girl talked to that boy.

**1.2. Two Semantic Rules**

**Chomsky's conceptual argument for the copy theory of movement**

- Even if we have a syntactic rule that replaces copies with variables, we need a semantic rule that would tell us how to interpret variables, and this rule could apply directly to structures with copies.
- So, there is no need for the syntactic rule, hence no need for traces distinct from copies. (So Syntax as apposed to semantics can be totally cyclic: i.e. no need to revise the structure of A when something is "moved out of A". There is no movement only "re-merge".)

**Copy Insensitive Rule**

(13) In a structure formed by XP movement,  $XP_n[\varphi \dots XP_n \dots]$ , the derived sister of XP,  $\varphi$ , is interpreted as a function that maps an individual, x, to the meaning of  $\varphi[x/n]$ .

$\varphi[x/n]$  is the result of replacing every constituent with the index n in  $\varphi$  with the DP,  $him_x$ .

$[[him_x]] = x$ .

**Copy Sensitive Rule (compare with Elbourne 2002)**

(14) In a structure formed by XP movement,  $XP_n[\varphi \dots XP_n \dots]$ , the derived sister of XP,  $\varphi$ , is interpreted as a function that maps an individual, x, to the meaning of  $\varphi[x/n]$ .

$\varphi[x/n]$  is the result of replacing the head of every constituent with the index n in  $\varphi$  with the head  $the_x$ .

$[[the_x]] = \lambda P. [[the]](P \cap \lambda y. y = x)$

I don't know how to distinguish on empirical grounds between the syntactic and the semantic versions of Copy (In)sensitive Rules.

Goal: To present various arguments in favor of Copy Sensitive Rules (which are independent of whether the implementation is syntactic or semantic). These arguments will provide empirical support for the copy theory of traces, hence, for traces.

Chomsky's conceptual argument suggests that the semantic version of the debate is the correct one. Nevertheless, I will stick to the syntactic version because I suspect it will be easier to follow.

The arguments:

- **Condition C**
- **Antecedent Contained Deletion**
- **A constraint on Extraposition**
- **An argument from Focus (Sauerland 2001)**
- **Conservativity**

## 2. The reason why it's so hard to tell:

It is hard to distinguish between Copy Sensitive and Copy Insensitive rules because quantifiers in NL are conservative.

$$\begin{aligned}
 Q(A,B) &= Q(A, A \cap B) && \text{(by conservativity)} \\
 &= Q(A, A \cap \lambda x: A(x).B(x)) && \text{(by defensible assumptions about Pres. Projection)} \\
 &= Q(A, \lambda x: A(x).B(x)) && \text{(by conservativity)} \\
 &= Q(A, \lambda x. B(\text{the}Ax))
 \end{aligned}$$

## 3. Condition C

### *Wh*-movement

- (15) a. \*He<sub>i</sub> visited a friend of John's<sub>j</sub>.  
 b. A friend of John's<sub>j</sub> visited him<sub>i</sub>.

### QR

- (16) A boy talked to every girl. → [every girl] A boy talked to t.  
 (17) ??/\*Someone introduced him<sub>i</sub> to every friend of John's<sub>j</sub>. →  
 [every friend of John's<sub>j</sub>] Someone introduced him to t.  
 (Cf. *Someone introduced John<sub>i</sub> to every friend of his<sub>j</sub>.*)

The copy theory of movement seems to account for the fact that movement doesn't obviate Condition C:

- (15')c [which friend of John's<sub>j</sub>] he<sub>i</sub> visited [which friend of John's<sub>j</sub>].  
 (16') [every friend of John's<sub>j</sub>] Someone introduced him<sub>i</sub> to [every friend of John's<sub>j</sub>].



Sauerland's explanation: The sentences do not obey Parallelism because the trace in the relative clause and the trace of QR are semantically distinct (given the copy theory of movement).

- (22) a. \*[a book about a person<sub>j</sub> you <saw person<sub>j</sub>>]<sub>i</sub>  
           I saw book<sub>j</sub>.  
       b. [the car that belongs to the man<sub>j</sub> you <\*like man<sub>j</sub>>]<sub>i</sub>  
           I like car<sub>j</sub>

#### 4.2.2.Sauerland's extension

- (23) a. \*I visited a city near the lake<sub>j</sub> John did <visited t<sub>j</sub>>.  
       b. (??)I visited a city near the city<sub>j</sub> John did <visited t<sub>j</sub>>.  
       c. I visited a city near the one<sub>j</sub> John did <visited t<sub>j</sub>>.

- (24) a. [a city near the lake<sub>j</sub> John <visited lake<sub>j</sub>>]<sub>i</sub>  
           I visited city<sub>j</sub>.  
       b. [a city near the city<sub>j</sub> John <visited city<sub>j</sub>>]<sub>i</sub>  
           I visited city<sub>j</sub>.

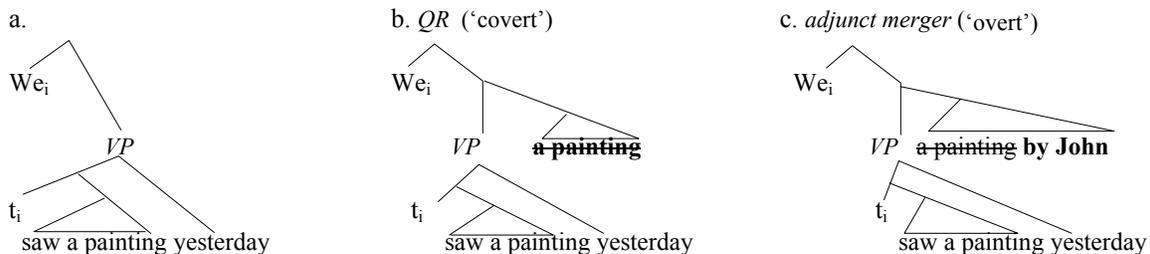
#### 4.3.The Problem

The existence of ACD poses a serious problem for Copy Sensitive Rules. However, Sauerland's contrast suggests that we might want to deal with this problem. Specifically, it suggests that if we had a way to overcome the problems that ACD poses for a Copy Sensitive Rule, we might be able to account for Kennedy's puzzle.<sup>2</sup>

#### 4.4.Extraposition/late merger provides the answer

##### 4.4.1.Background (Fox and Nissenbaum 1999)

- (25) We saw a painting yesterday by John.



<sup>2</sup> Sauerland provides an account that relies on an Economy condition argued for in Fox (1995, 1999). See also Merchant (2000). However, the economy condition is tailored specifically to the puzzle at hand (see Fox (2002) for more details).

Copy Insensitive Rule: [A painting by John]  $\lambda x$  we saw [x]  
 Copy Sensitive Rule: [A painting by John]  $\lambda x$  we saw [the painting x]

A derivation along these lines was proposed for overt *wh*-movement by Lebeaux (1988). The extension to QR is straightforward under the copy theory of movement and the assumption that covert movement involves pronunciation of a lower copy (the phonological theory of QR, Bobaljik, Groat and O'neil, Pesetsky). (For related but different proposals see Guéron and May 1984 and Reinhart 1991.)

- (1) **Williams' generalization:** When an adjunct  $\beta$  is extraposed from a “source NP”  $\alpha$ , the scope of  $\alpha$  is at least as high as the attachment site of  $\beta$  (the extraposition site).<sup>3</sup>

**Scope diagnosed by variable binding in ellipsis:**

- (2) a. I read a/every book before you did.  
 b. I read a/every book that John had recommended before you did.  
 c. I read a/every book before you did that John had recommended.

(2a,b) show an ambiguity that is explained in terms of the relative scope of the object quantifier and the *before*-clause (see Sag 1976 and Williams 1977). (2c) has only the interpretation in which the object quantifier has wide scope.

**Scope diagnosed by NPI licensing:**

*Adjunct-extraposition bleeds Condition C:*

- (3) a. I gave him<sub>j</sub> a book yesterday that John<sub>j</sub> liked  
 b. ??I gave him<sub>j</sub> a book that John<sub>j</sub> liked yesterday

Some speakers don't find (3b) that bad. This is somewhat expected given the availability of string-vacuous extraposition. The fact that (3b) is unacceptable to many speakers probably reflects a processing preference for low attachment. (See Phillips 1996.)

*Conflicting requirements on scope of the source DP — Condition C vs. NPI*

- (4) a. I wanted him<sub>j</sub> not to talk to a (certain) girl yesterday that John<sub>j</sub> has known for years  
 b. \*I wanted him<sub>j</sub> not to talk to any girl yesterday that John<sub>j</sub> has known for years.  
 c. I wanted John<sub>j</sub> not to talk to any girl yesterday that he<sub>j</sub> has known for years.

<sup>3</sup> This is a slight modification (due to Fox and Nissenbaum 1999) of the original statement of the generalization (Williams 1974, chapter 4): originally William made no complement/adjunct distinction, perhaps because he focused on comparative- and result-extraposition, where, as Bhatt and Pancheva argue, complements pattern with NP adjuncts (for reasons having to do with the nature of Trace Conversion).



- (29) [<sub>VP</sub> I visited a city]. -QR-->  
 [[<sub>VP</sub> I visited a city] a-city]. -adjunct insertion-->  
 [[<sub>VP</sub> I visited a city] a-city near the lake that John did <visited lake>].  
 \*[a city near the lake  $\lambda_x$  John did <visited the lake x>]  
 $\lambda_y$  I visited the city y]

### Late Merger is restricted to adjuncts (Lebeuax)

#### Fox and Nissenbaum:

- (30) a. I gave him<sub>i</sub> an argument yesterday that supports John's<sub>i</sub> theory.  
 b. ??/\*I gave him<sub>i</sub> an argument yesterday that this sentence supports John's<sub>i</sub>theory.

Extrapolation does not take an adjunct out of the Noun Phrase (even though it does take the adjunct to a position above elements that c-command the Noun Phrase at SS, compare with Buring and Hartmann 1997)

- (31) a. I told you about John's<sub>i</sub> new argument that supports his<sub>i</sub> theory.  
 b. \*I told you about his<sub>i</sub> new argument that supports John's<sub>i</sub> theory.  
 c. I told John<sub>i</sub> about your new argument that supports his<sub>i</sub> theory.  
 d. \*I t told him<sub>i</sub> about your new argument that supports John's<sub>i</sub> theory.
- (32) a. ?? I told you about John's<sub>i</sub> new argument the other day that supports his<sub>i</sub> theory.  
 b. \* I told you you about his<sub>i</sub> new argument the other day that supports John's<sub>i</sub> theory.  
 c. ?? I told him<sub>i</sub> about your new argument the other day that supports John's<sub>i</sub> theory.

#### Sauerland:

- (33) a. I visited a city **near the one John did <visited>**.  
 b. I made an argument **that was very similar to the one you did**. <made>.  
 c. I saw a picture **that depicts the picture you did <saw>**.
- (34) a. \*I made an argument **that we should adopt the one you did <made>**.  
 b. ??I saw a picture of the picture **you did <saw>**.

**Note:** the restriction to adjuncts follows automatically from the copy-sensitive rule. See Bhatt and Pancheva (this workshop) for an exception that proves the rule.

### 4.5.Evidence that extraposition is crucial for ACD

Extrapolation (analyzed as Late Merger) is necessary to make ACD compatible with a Copy Sensitive Rule. Therefore, evidence that extraposition is obligatory in ACD constructions can be taken as evidence for a Copy Sensitive Rule. (cf. Baltin 1987, Lasnik 1993, Tiedeman 1995, and Wilder 1995).

- (35) a. \*John said that everyone who you did was a genius.  
 b. John said that everyone was a genius who you did.  
 (Tiedeman 1995, Wilder 1995)
- (36) a. \*I know how much every item that you do costs.  
 b. ??I know how much every item costs that you do.  
 (compare Moltmann and Szabolcsi 1994 with Fox 2000)
- (37) a. I discussed nothing that you did <\*discussed in order to upset your teacher> in order to upset my teachers.  
 b. I did none of the dives that my coach told me to <\*do in order to prove that I could> in order to prove that I could.  
 b. Bill read every book that I did <\*read with great care> with great care.
- (38) a. I discussed nothing in order to upset my teachers that you did <discussed in order to upset your teacher>  
 b. I did none of the dives in order to prove that I could that my coach told me to <do in order to prove that I could>.  
 b. Bill read every book with great care that I did. <read with great care>
- (39) a. I read the book after you did <read it>  
 b. \*I read the book Bill did <read t after you read it> after you did <read it>  
 c. I read the book after you did <read it> that Bill did <read t after you did>
- (40) a. You sent him<sub>i</sub> the letter that John<sub>j</sub> expected you would.  
 b. You introduced him<sub>i</sub> to everyone John<sub>j</sub> wanted you to.  
 c. I reported him<sub>i</sub> to every cop John<sub>j</sub> was afraid I would.
- (41) a. \*John will say that Mary put every book that Bill also will <say that Mary put on the shelf> on the shelf.  
 b. John will say that Mary put on the shelf every book that Bill also will <say that Mary put on the shelf>.
- Jacobson (1992): (41a) is bad since here ACD depends on the “wrap” rule for pronunciation, which places the NP after the head of the complex transitive verb, which in turn serves as antecedent for ellipsis. This account doesn’t extend to (41c).
- c. John will say that Mary put every book on the shelf that Bill also will <say that Mary put on the shelf>.

**Note:** There are examples where extraposition doesn’t seem to be necessary (Larson and May 1990). In Fox (2002) I provide various pieces of evidence that this is an illusion. The sentences involve rightward movement of various constituents.

## 5. A constraint on Extraposition

The fact that QPs are conservative makes the semantics of sentences with late-merger (extraposition) identical to the semantics of otherwise identical sentences with early merger (without extraposition).

However, if we can do Late Insertion in a somewhat embedded position, the situation might be different.

### 5.1. Rochemont and Culicover's discovery

- (42) a. John saw an alleged mouse from Mars yesterday  
 b. John saw an alleged Mouse yesterday from Mars.

Rochemont and Culicover: When EC is "extraposed" from DP, NP does not dominate EC.

- (43) 1. An [[alleged mouse] from Mars]  
 2. An [alleged [mouse from Mars]]

- (43') 1.  $\lambda Q.\exists x(Qx \ \& \ \text{from-Mars}'(x) \ \& \ \text{alleged}'(\text{mouse}'(x)))$   
 2.  $\lambda Q.\exists x(Qx \ \& \ \text{alleged}'(\text{mouse}'(x) \ \& \ \text{from-Mars}'(x)))$

- (44)  $\llbracket \text{alleged} \rrbracket (P)(x) = 1$  iff there is an allegation that P is true of x.

When a speaker utters *An alleged mouse is in the room*, there is a feeling that the speaker doesn't believe that the alleged mouse is a real mouse. This is probably an implicature but since the distinction doesn't matter for most of the cases we will look at, we can think of it as part of the meaning of the lexical entry.

- (45)  $\llbracket \text{alleged} \rrbracket (P)(x) = 1$  iff P is not true of x and there is an allegation that P is true of x.

### 5.2. The Complement Adjunct distinction

- (46) a. John read an alleged book yesterday about Irene Heim.  
 b. #John read an alleged book yesterday by Ronald Reagan.

Context (that might help with A, thanks to Tamina Stephenson): Irene's sister wrote a book the protagonist of which is a semanticist who left Germany at a young age to study at a university on the east coast of the united states. Someone makes the allegation that the story is based on the life of Irene Heim.

A similar Context doesn't help with b: there's this book about indefinite and everybody believes the author is Irene Heim. Someone makes the allegation that the author is actually Ronald Reagan.

### 5.3. The constraint is more subtle

(42b) can also be true when the alleged mouse is not necessarily from Mars, i.e., only alleged to be from Mars.

Evidence in favor:

- (47) a. I saw an alleged antique yesterday from the 15<sup>th</sup> century.  
 b. I met an alleged New-Yorker yesterday from the Upper East Side.  
 c. They told me about an alleged alien the other day from Mars.

#### 5.4.An Explanation

The observation is very surprising. There is no obvious way of getting the relevant reading from a parse of *an alleged NP adjunct*.

Take (47a) as an example. The relevant reading, given Rochemont and Cullicover's observation, has to be somewhat similar to the following paraphrase:

- (48)  $\exists x(\text{alleged}(\text{antique}(x)) \wedge \text{alleged}(\text{from-the-15}^{\text{th}}\text{-century}(x)) \wedge \text{saw}(I,x))$

The relevant reading is one under which (47a) is true if the thing I saw is (a) alleged to be an antique and (b) alleged to be from the 15<sup>th</sup> century.

However, we have more options once we consider the independent contribution to meaning that the trace makes under a Copy Sensitive Rule.

- (49) a. John saw an alleged mouse yesterday from Mars.  
 b. Right after John saw an alleged mouse from Mars, he talked about the alleged mouse to various reporters.
- (50) a. John saw an alleged antique yesterday from the 15<sup>th</sup> century.  
 b. Right after John saw an alleged antique from the 15<sup>th</sup> century, he talked about the alleged antique with an art dealer.

- (47a') [An alleged [antique from the 15<sup>th</sup> century]]  
 $\lambda x$  I saw [the alleged antique (identical to) x]

#### 6. An argument from Focus (Sauerland 2003)

- (51) a. I read every book before YOU *read every book*.  
 b. I read every book before YOU *read EVERY book*.

#### 7. Conservativity

It is hard to distinguish between Copy Sensitive and Copy Insensitive rules because quantifiers in NL are conservative.

$$\begin{aligned} Q(A,B) &= Q(A, A \cap B) && \text{(by conservativity)} \\ &= Q(A, A \cap \lambda x: A(x).B(x)) && \text{(by defensible assumptions about Pres. Projection)} \\ &= Q(A, \lambda x: A(x).B(x)) && \text{(by conservativity)} \\ &= Q(A, \lambda x. B(\text{the}Ax)) \end{aligned}$$

But a Copy Sensitive Rule might account for conservativity. (We will return to this at the last section.)

Non-conservative quantifiers (if they existed) would be quite pointless.

(52) #Only girls are such that those girls are intelligent.

For a standard non-conservative quantifier,  $Q$ , the truth value of  $Q(A,B)$  depends on the value of  $B(x)$  for objects  $x$  that are not in  $A$ . However, for such objects  $B$  is undefined (since  $B$  is really  $\lambda x.B(\text{the}Ax)$  ( $= \lambda x: A(x).B(x)$ )).

A very tentative attempt to specify the general principle that is at work here:

$Q(A,B)$  is inappropriate if in all worlds in the context set  $Q(A,B)$  is not true.

- (53) a. #Very few boys in my class have a car, and Every boy drives his car to school.  
 b. Very few boys in my class have a car, but one boy drives his car to school.  
 Beaver 1992

Every  $(A,B) = 1$  iff for every  $x$ , s.t.  $A(x) = 1$ ,  $B(x) = 1$ . If we know that **there is an  $x$  in  $A$** , such that  $B(x)$  is undefined, we know there is no way for the sentence to be true.

Some  $(A,B) = 1$  iff there is some  $x$ , s.t.  $A(x) = 1$  and  $B(x) = 1$ . If we know that **every  $x$  in  $A$**  is such that  $B(x)$  is undefined, we know there is no way for the sentence to be true.

Only  $(A, B) = 1$  iff for every  $x$ , s.t.  $A(x) = 0$ ,  $B(x) = 0$ . If we know that **there is an  $x$  in the complement set of  $A$** , such that  $B(x)$  is undefined, we know there is no way for the sentence to be true.

Under a Copy Sensitive Rule,  $B(x)$  is undefined for every  $x$  in the complement set of  $A$ . Hence, *only* cannot be a natural language determiner.

**Problem with the explanation:** The option of extraosition should allow non-conservative quantifiers to be realized syntactically. In other words, it is predicted that non-conservative quantifiers should exist, but that they should be restricted to extraposition environments.

This, Bhatt and Pancheva argue, is precisely what happens.

**Conclusions:****Claims:**

A. Shared assumption (1) needs to be modified. The argument of a restrictive quantifier (at least when there is Inverse Scope) is a partial function defined only for individuals that satisfy the restrictor:

- (3'a) a.  $\lambda P_{\langle e,t \rangle}. \llbracket \text{book} \rrbracket \subseteq P$   
 b.  $\lambda x_e: x \text{ is a book. A girl read } x$

B. Inverse Scope Is derived by Movement, and traces are copies.

- (2'') a. [A girl gave [every book]<sub>1</sub> to John] [every book]<sub>1</sub>

The quantificational phrase is present at the base position, and its restrictor is accessed as part of the semantic rule for the interpretation of movement (predicate abstraction).

$\llbracket (2'')a \rrbracket = \llbracket \text{every book is such that a girl read that book} \rrbracket$

**Consequences for Direct Compositionality (DC):**

Assuming that claims (A) and (B) are right, does it mean that DC is wrong?

No. It only has consequences for the way we ought to develop a theory within DC. It would tell us that we should develop a theory with (complex) traces, and one can do this (see Jacobson 2002).

However, some of the arguments for claims (A) and (B) are based on particular assumptions, which cannot easily be implemented within DC, in particular assumptions about the nature of Binding Theory, and the analysis of extraposition.

To the extent that the assumptions are independently motivated and to the extent that they converge with argument that are theory neutral (e.g., the *alleged-mouse-from-Mars* argument or the argument based on the Kennedy/Sauerland puzzle) we have a real challenge for DC.