Direct Compositionality: Is there any reason why not?

Pauline Jacobson
Dept. of (cognitive and) Linguistic Science(s)
Brown University
pauline_jacobson@brown.edu

University of Michigan Workshop on Linguistics and Philosophy. November, 2004
draft version, Sept. 9; bibliography to be added

0. Goals

The hypothesis of Direct Compositionality - familiar (in one form) from, the work of Montague (e.g., Montague, 1974) and more recently from considerable work within Categorial Grammar - is that the syntax and semantics work "in tandem". The syntax is a system of rules (call them "principles" if one prefers) which prove the well-formedness of linguistic expressions while the semantics works simultaneously to provide a model-theoretic interpretation for each expression as it is proved well-formed in the syntax. (I will use the usual metaphor that the syntax "builds" expressions.) Under this conception, there is no intermediate representation such as LF mediating between the surface syntax of linguistic expressions and their model-theoretic interpretation and consequently (and more importantly) no need for an additional set of rules providing a mapping between expressions and their LFs.

This paper is a plea for this view of the organization of the grammar. I will argue that direct compositionality is (a) the simplest conception of the organization of the grammar amongst competing alternatives, that (b) it should therefore be abandoned only in the face of strong evidence to the contrary, and that (c) there has never been convincing evidence to the contrary. The paper divides into two parts. The first part overlaps largely with my paper "The (dis)organization of the grammar: 25 years" which appeared in *Linguistics and Philosophy* ***25****.5-6* (the anniversary issue). There I will sort out two (or three) versions of the hypothesis of direct compositionality - essentially, the differences between these versions are differences in theories of *syntax:* the weaker versions allow a richer set of operations in the syntax than do stronger versions. But I will argue that even what I call Weak Direct Compositionality is much simpler and more explanatory than a view of the grammar in which the syntax first "runs" and then sends syntactic representations to the semantics for compositional interpretation. No arguments for this kind of view over a Weak Direct Compositional theory have, to my knowledge, ever been put forth. To the extent that one finds articulated arguments in the literature against direct compositionality, these are *at best* arguments only against what I call Strong Direct Compositionality - not arguments against direct compositionality per se. (I say "at best" because I think that most of these have been answered especially in much of the Categorial Grammar (including Type Logical Grammar) and related literature. But there do remain some potential problems for Strong Direct

Compsoitionality, and so the question remains open as to just how rich a set of operations are available in the syntax.) It seems, then, that the discussion should not be about "LF based" approaches vs. direct compositionality: rather rational discussion should be about just how strong a version of direct compositionality is possible.

The second part of this paper turns to a group of phenomena which, I argue, are not only compatible with direct compositionality, but with strong direct compositionality. I use this as a case study because these phenomena - and their interactions - have been used to argue for a variety of devices incompatible with strong direct compositionality. The phenomena involve the interaction of extraction, pronoun binding, and "Condition C" effects. The apparent challenges that these pose are (a) they are often accounted for in such a way as to require reference to an abstract level like LF which is the input to the model-theoretic interpretation, and (b) they seem to require reference to representational properties (things like c-command). But under a strong direct compositional view, things like trees are just convenient representations for the linguist of how the syntax worked to prove something well formed and how the semantics worked to put the meanings together - they are not representations that the grammar gets to "see" and so the grammar could not possibly contain principles stated on these representations. From much of the work in Generalized Phrase Structure Grammar in the 80's and continued work within Categorial Grammar, we have learned that when we seem to find a generalization about representations, this is often an artefact - a consequence of how the rules worked, not a consequence of an actual principle in grammar that refers to representations. I should stress one point at the outset: trying to derive facts from the way the combinatory rules (the syntax, and the semantics) works as opposed to trying to derive facts from principles on representations is *not* just a trade-off to be determined by what is fashionable. Any theory needs combinatory principles in the syntax and in the semantics - to prove well-formed strings and assign them a meaning. Thus if we can get all the work to come from there (without, of course, unduly complicating these principles) then we have less machinery than a theory which also needs the rules (principles) to keep track of representations, so as to allow these to be the input to further statements.

1. Direct Compositionality

1.1. Four Conceptions of the syntax-semantics "interface"

Thus let me outline here four theories of the syntax-semantics "interface". The first three embrace direct compositionality in some form - and I want to make these all explicit to show that any of these three is preferable to the "modern" conception of things whereby surface structures are mapped into a separate level which is then compositionally interpreted. (I realize that things are a bit different from this picture under "Minimalism", but the key point remains the same - the syntax works separately to ultimately build LFs which the semantics then interprets.) Since the possibility of having quantified NPs in object position, and the further possibility of scope ambiguities as in (1) has always played a fundamental role in the discussion of the syntax/semantics

interaction (see, for example, any textbook discussion of "LF"), I will illustrate each of these theories with a brief consideration of its treatment of scope:

(1) Some man read every book.

A. *Strong Direct Compositionality*. Under this view, there is a set of syntactic rules which "build" (that is, prove the well-formedness of) the set of sentences (or other expressions) in the language - generally "building" (i.e., specifying the well-formedness of) larger expressions in terms of the well-formedness of smaller subexpressions. Assume here that each such rule is a context-free phrase structure rule (or, highly generalized rule schema). Note that if this is the only form of syntactic rule in the grammar then the grammar need keep no track of structure: the rules "building" complex expressions merely concatenate strings. Thus under this conception, a tree is just a convenient representation of how the grammar worked to prove a string well-formed; it is not something that the grammar can nor ever would need to "see". This will become crucial in much of the discussion in Sec. 2: I wish to argue that there is no need for constraints in the grammar which are stated on global chunks of representation and that thinking of various phenomena in that way often just gets in the way of providing an adequate account. (By a "global" chunk of representation I mean a chunk which is larger than a local tree - any statement concerning a local tree can obviously be recast as part of a rule.) Coupled with each syntactic rule is a semantic rule specifying how the meaning of the larger expression is derived from the meaning of the smaller expressions.

The above terminology may be misleading in one respect: strong direct compositionality is not committed to the view that the grammar consists of a list of phrase structure rules each of which is idiosyncratically associated with a semantic rule. Such a view has often come under fire for containing "construction specific" semantic rules - a state which complicates the grammar because it adds to the amount of idiosyncratic information which must simply be listed. But the issue of construction-specific rules is completely independent of the other issues that I am concerned with here, and nothing in the strong direct compositional view requires construction-specific rules. One can maintain that the actual semantic operations are predictable from each syntactic rule. In fact, the earliest discussions of "type driven" interpretation (e.g., Klein and Sag, 1983) were framed within a strong direct compositional view: their use of the term meant that the semantic operation associated with each syntactic (phrase structure) rule schema was predictable from the syntax of the *rule* combined with the semantic types of the expressions referred to in the rule. Moreover, a strong direct compositional theory can also assume that the syntax itself consists of just a few very general rule schemata (each of which is associated with a general semantic rule schema). Indeed, this is exactly the view taken in various versions of Categorial Grammar; thus the rejection of Strong Direct Compositionality surely cannot be motivated by a rejection of "construction specific" semantic rules.

There is a further issue of relevance here. Much research within the tradition of Strong Direct compositionality - or within one of its slightly weaker versions to be discussed below - has also advocated the existence of unary rules. These are rules

which map single linguistic expressions into new ones, and in so doing change the meaning and/or category of an expression (generally without changing the phonology). Such rules have generally gone under the rubric of "type-shift" (and/or category-changing) operations. There is disagreement in the literature as to just how many such operations there are and how generally they should be stated. This in turn ties in with the question of how much "construction-specific" information the grammar contains – or – to put this in less sloganistic terms: just how complex the grammar is. (Note too that any theory allowing for silent operators is not very different from one allowing type-shift rules: a type-shift operation can always be recast as a silent lexical item which applies to the expression with which it combines.) Thus, while one can argue about just how many and how predictable are the individual rules, this too is a separate issue.

There are some very appealing properties of Strong Direct Compositionality. One is the fact that in building strings, the syntax need keep no track of structure, since all combinatory operations simply concatenate strings, and all unary rules have no effect on the internal structure. We can think of each linguistic expression as a triple of <[phonology]; syntactic category; meaning>, where the rules take one or more such triples as input and give back a triple as output.

How are quantifier scope ambiguities handled under strong direct compositionality? As with most things, there is more than one proposal in the literature. The most influential proposal during the late 70's and 80's was probably that of Cooper (1975). But there are other proposals for quantifier scopes within strong direct compositionality; in general these involve some type-shift rule or rules. One well-known proposal is developed in Hendriks (1993) and is a generalization of ideas in Partee and Rooth (1983). Here a transitive verb like *read* is listed in the lexicon with a meaning of type <e,<e,t>>, but there is a generalized type-shift rule allowing any e-argument position to lift to an <<e,t>,t> argument position. If the subject position lifts first and then the object position lifts the result is the wide-scope reading on the object. Further generalizations of this can be used to allow for wide scope of embedded quantifiers; other proposals for wide scope embedded material makes use of a combination of type-shift rules and function composition. (For another kind of proposal, see Barker, 2001.)

B. *Weak(er) Direct Compositionality*. The above picture has often been weakened by the adoption of two related revisions: (a) the combinatory syntactic rules are not all equivalent to context-free phrase structure rules but may perform some other operations, and (b) the syntactic rules do not build only completely unstructured strings but may build objects with more "structural" information. The mildest weakening of A is to be found, perhaps, in those proposals that add only Wrap operations in addition to concatenation operations (for the original Wrap proposal, see Bach, 1979, 1980). Here the combinatory syntactic operations allow two strings not only to not concatenate, but also for one to be infixed into another. As such, the input to the combinatory operations has to be not just unstructured strings, but these strings need to contain at least enough additional information so as to define the infixation point. This has been formalized in a variety of ways; I will not pursue this here, although it is worth noting that Montague's Quantifying In rule can be recast as an infixation operation, and so Weak Direct

Compositional. systems with infixation operations are one way to account for quantifier scopes. I actually do think that Wrap operations are needed - we know from much discussion in the 80's that natural languages cannot be described purely with context-free phrase structure rules - but the addition of only Wrap operations seems quite reasonable. Although a theory with Wrap is not really what I am calling "Strong(est) Direct Compositionality", I will continue to use the term Strong Direct Compositionality to encompass a theory in which the only weakening is the addition of infixation operations.

But one can imagine other kinds of weakenings -- including one in which certain transformation-like operations can be performed in the "building" of syntactic structures. (Much of the work in classical Montague grammar allowed substitution and deletion operations, though not much else.) Of course just what can and can't occur also interacts with the question of how much structure is kept track of in the building of syntactic operations. At the extreme end, one can imagine that the output of each syntactic operation is a full-blown tree rather than a string (see, e.g., Partee, 1976); hence linguistic expressions are now richer, and can be thought of as triples of the form <[phonological representations]; syntactic structure – i.e., a full tree; meaning>.

Montague's "Quantifying-In" rule - and conceivable variants of it - is compatible with various versions of Weak Direct Compositionality (as noted above, it can be formulated just in terms of infixation, for example; and in Montague's work it was stated as a substitution operation.) Since my concern is not with Montague's particular proposal but with this general picture of the architecture of the grammar, I will recast the Quantifying-In rule(s) so as to provide a maximal basis for comparison with other theories. Thus assume (contra Montague, 1974) that a transitive verb like *read* has a lexical meaning of type <e,<e,t>>. Assume further the usual theory of variables, and assume (along the lines of Montague's treatment) that we build syntactic representations with indexed pronouns each of which corresponds to the samely-indexed variable in the semantics. We can have expressions like he_1 *reads* he_2 whose meaning – relative to some assignment function g - will be $[[\text{reads}]]^g([[x_2]]^g)([[x_1]]^g)$. In addition, we will let the syntax keep track of the indices on the unbound pronouns (this is not really necessary, but will facilitate the exposition). More concretely, assume that every node label is enriched with an IND feature, whose value is a set of indices, and that – unless a rule specifies otherwise - the IND value on the category which is the output of a combinatory rule is the union of the IND values of expression on the input. Thus the category of the expression given above is S [IND: {i,j}]. We can thus accomplish Quantifying-In by the following two rules, the first of which is a type-shift rule:

(2) Let α be an expression of the form $<[\alpha]$; S [IND: X where i ε X]; $[[\alpha]]>$. Then there is an expression β of the form $<[\alpha]$; L [IND: X-i]; $[[\beta]]^g$ is that function which assigns to each individual a in D, $[[\alpha]]^{g[a/x(i)]}>$ (this of course is just the semantics of λ-abstraction).

(3) Let α be an expression of the form $<[x, he_i\ y]$; $\Lambda_{,}$ $[[\alpha]]>$ and β be an expression of the form $<[\beta]$; DP; $[[\beta]]>$. Then there is an expression γ of the form: $<[x$ DP $y]$; S; $[[\gamma]]^g = [[\beta]]^g([[\alpha]]^g)>$.

(Notice that, unlike in Montague's treatment, I have broken the Quantifying In rule down into two steps; Montague collapsed into a single but more complex rule.) One can also build Weak Crossover into this picture: Montague's rule itself required that if there is more than one pronoun with the same index, the substitution could apply only to the leftmost one. Should one want to take the more usual modern view that the appropriate restriction is stated in terms of c-command rather than linear order (cf., Reinhart, 1983) the rule can be restricted so that he_i in the above SD must c-command all other occurrences of the same indexed pronoun (such a restriction, of course, commits to the view that the input to these rules are as rich as trees).

Incidentally, here too I have stated the rules in a construction-specific way, but here too this is not a necessary requirement of the program. To be sure, most work within "classical Montague Grammar" of the 70's made heavy use of construction-specific information, but there is no reason why the program seeking to minimize or eliminate this is incompatible with the general architecture of Weak Direct Compositionality (if there is such a reason, it has never to my knowledge been demonstrated).

C. *Generative Semantics-like Direct Compositionality*. By this I mean something like the model proposed in Generative Semantics (see, e.g., Bach, 1968; McCawley, 1970; Lakoff, 1971) supplemented with apparatus to supply a model-theoretic to the Logical Forms. Thus, Generative Semantics assumed that deep structure was the same as Logical Form - which means that a series of phrase structure rules and/or rule schemata serve to define a well-formed structure. This was supposed to "represent" the semantics, and in fact much work within Generative Semantics didn't worry about supplying an actual interpretation to these structures. (For that matter, this is equally true of some of the initial work within the "modern" surface structure to LF view; see, e.g., Chomsky (1976).) But it is easy enough to embed the general idea into a more sophisticated theory of semantics with a model-theoretic component. Simply have the "building" and interpretation of the Logical Forms be as in the Strong Direct compositional approach: each local tree is specified as well-formed by the syntactic rules and – in tandem – is provided a model-theoretic interpretation by the semantic part of the rules. A key difference between this and Strong Direct Compositionality is that this view contains an additional set of transformational rules which map the Logical Forms to surface structures. A concomitant difference is that the rules "building" syntactic structures must keep track of whatever structure is used as the input to the transformational rules; presumably then these rules are building trees rather than strings. Again, though the base rules can be seen as mappings from triples to triples.

The treatment of quantifier scopes within this general view is well-known. First, unlike the Quantifying-In rule above, we have an actual level of representation at which quantified NPs are in the tree, but are in a raised position rather than being in their ultimate surface positions. The difference between their deep and surface positions is handled by a quantifier lowering rule. If we take the lexical meaning of a transitive verb like *read* to be of type <e,<e.t>> then the appearance of a quantified NP in object

position will always be the result of Quantifier Lowering. Scope ambiguities are handled in the obvious way: since each local tree is interpreted as it is "built" by the phrase structure rules, the obvious formulation of the semantic rules will assign different scopes according to the initial height of the quantified NPs.

Again it is worth spelling this out explicitly. Let the rules build deep structure expressions such he_1 *read* he_2 as in the Weak Direct Compositional approach shown above. Assume further that this is assigned the meaning and category as above. Further, assume the following two phrase-structure rule semantic rule pairs; these mirror the rules in (2)-(3):

(4) Λ [IND: X – i] ---> S [IND: X, where i ε X]; $[[\Lambda]]^g$ is that function from individuals to propositions which assigns to each individual a in D $[[S]]^{g[a/x(i)]}$

(5) S ---> DP Λ; $[[S]]^g = [[DP]]^g([[\Lambda]]^g)$

Finally, this is supplemented by one transformation, as follows:

(6) $[_S$ DP $[_\Lambda$ A he_i B]] ===> $[_S$ A DP B]

(Again, one can restrict the rule so that the occurrence of he_i which is analyzed as meeting the SD of the rule is leftmost occurrence of he_i (see, e.g., Jacobson, 1977), or one can equally well restrict the rule to apply to the highest such occurrence, to account for Weak Crossover.)

D. *Surface to LF*. Which brings us to the increasingly popular view. This is that there is a level of LF which receives a model-theoretic interpretation, and which is derived from surface structures by some set of rules. There are actually two possible versions of this. One is that the surface structures are given directly by the compositional syntactic rules (i.e., the phrase structure rules or their equivalents) and these are then mapped into LFs. The second, and more standard view, is that the surface structures themselves are the end-product of a mapping from underlying structures. In terms of the treatment of quantifier scopes this makes little difference; it does make a difference when we turn to *wh* questions. The actual proposals cast within D generally do presuppose the existence of transformational operations in the syntax - this is because many of the arguments for D rely on similarities between the "overt" transformational operations (mapping from deep to surface structures) and "covert" operations (mapping from surface to LF).

The treatment here of scope ambiguities is also well-known; it is essentially the same as that given above under C, except that the direction is reversed. We first start out with a series of rules which ultimately define a well-formed surface structure at which the quantified material is *in situ*. Then there are rules mapping this into a structure like the Generative Semantics deep structure, and then the compositional semantic rules will work from the bottom up to interpret this.

Thus assume again that the lexical meaning of *read* is of type <e,<e,t>>, and assume a set of phrase structure rules which allow DPs like *every book* to appear in characteristic "argument" positions (even though they are not of the right type to semantically combine with, e.g., a transitive verb). (No syntactic transformational operations of relevance apply in this case.) We thus build a structure like *some man read every book* in the syntax, but initially with no interpretation. The interpretive part can be accomplished by a combination of one transformation-like rule – Quantifier Raising (May, 1977) which is essentially the inverse of the Quantifier Lowering rule in (6)) and two rules interpreting the relevant structures. The relevant rules are given in (7) – (9). Incidentally, in the formulation that I give here in (7) QR is not the exact inverse of (6) - this is simply because I formulated this to be more or less in keeping with the treatment given in Heim and Kratzer (1998). But one could formulate the rules to be exact inverses. Either way, the rules are formulated so as to keep track of the index on the variable which is λ-abstracted over in the interpretation of the category labeled Λ. This can be done by having it be the index which is different on the mother's index feature set and on the daughter's feature set (which is encoded above via the rule in (4)), or by actually putting that index as a terminal sister of S. (The difference is part and parcel of the question mentioned above about the use of type-shift operations vs. silent operators; it is difficult to find any empirical consequence to the choice of one or the other).

(7) $[_S\ A\quad DP_i\quad B]\ \Longrightarrow\ [_S\ DP\ [_\Lambda\ i\ [_S\ A\ t_i\ B]]]$

The lowest S will be interpreted by the kind of interpretation rules that anyone needs; note that t_i here is interpreted as x_i. The additional rules of interpretation are pretty much just the inverse of what we have already seen:

(8) $[[\ [_\Lambda\ i\ S]\ ]]^g$ is that function which assigns to each individual x in D the value $[[S]]^{g[a/x(i)]}$

(9) $[[\ [_S DP\ \Lambda]\ ]]^g = [[DP]]^g\,([[\Lambda]])^g$

1.2. Why each view is weaker than the one before it

To fully evaluate the advantages or disadvantages of each of these views of the organization of the grammar one would need actual proposals for much larger fragments than are given here. But all other things being equal, it seems obvious that each position in this list represents a more complex view of the organization of the grammar than does the position above it.

In particular, I want to focus on the fact that there is a major cut between A-C on the one hand and D on the other. A, B, and C all have in common the fact that the syntax "builds" in conjunction with the model-theoretic interpretation, and this is discarded in D. But there are certain complications which arise the minute one moves away from the "running in tandem" of the syntax and semantics.

The first objection to the divorcing of the syntax and the semantics might be purely subjective (although I don't really think it is). This is that there is a clear elegance to a system in which the grammar builds (i.e., proves as well-formed) syntactic objects in parallel with assigning them an interpretation, an elegance which is lost if the grammar contains two entirely separate systems, one of which must "run" first because the other (the semantics) works on the output of the first (the syntax).. But elegance aside, there are two other questionable aspects to divorcing the two combinatory systems. The first is that under the conception in D there is no explanation as to *why* these systems work on such similar objects, and the second (related) problem is that D requires a duplication of information not required in A-C.

We turn to the first problem. Regardless of the question of whether there are transformations in addition to phrase-structure rules (or rule schemata), just about all theories agree on something like phrase-structure operations each of which specifies the well-formedness (at some level) of a *local tree*. (As noted above, under strong Direct Compositionality, trees are not actually anything the grammar ever needs or gets to see, so under this view a tree is really a metaphor. Under Weak Compositionality this is also not quite the appropriate terminology. If the syntax allows operations like Wrap, then a "tree" may also be not the right formal object to represent what the syntax does. However, I think that this terminology will do no harm here.) Thus at the end of the day, all theories contain rules which can be seen as proofs of the well-formedness of some kind (or level) of structure, where these "proofs" work bottom up in that they specify the well-formedness of larger expressions (at some level of representation) in terms of the well-formedness of smaller ones. ("Bottom up" is of course also a metaphor here, and one that is not common in discussions of syntax. But once one thinks of the syntactic system as supplying a proof of the well-formedness of some expression then this metaphor is appropriate - larger expressions are well-formed if they are composed in certain ways from smaller well-formed expressions.) Moreover, semantic theories generally agree that the semantics also works on *local trees* to give the meaning of the mother in terms of the meaning of the daughters. And, like the syntax, it also must work "bottom up" - supplying the meaning of larger expressions from the meanings of the expressions which compose them. Given this, it would seem to be extremely surprising to find that the two systems don't work in tandem. If the semantics is divorced from the syntactic combinatory rules, then why <u>should</u> it be the case that it too works on local trees? Why *not* find rules taking large chunks of trees and providing an interpretation for these?[1]

[1] I am oversimplifying, in that I am not sure we have any actual evidence that the semantics *does* necessarily work on local trees - this is pretty much just what everyone assumes (and it is, of course, an assumption which is forced under the view that the syntax builds in tandem with the semantics interpreting). Maybe a fairer way to state the situation, then, would be to say that no evidence has been found to suggest that the semantics *requires* bigger objects as its input, but if the syntax and the semantics were really divorced then there'd be no reason not to find such cases.

Let me push this point a bit: consider the discussion of how the semantics works in Heim and Kratzer (1998). On p. 29 they make explicit the following assumption - one which is pretty much accepted in most modern work in formal semantics: "***Locality.*** Semantic interpretation rules are local: the denotation of any non-terminal node is computed from the denotation of its daughter nodes." Now if mothers and daughters are merely what rules refer to and if the semantics works in tandem with the syntax then this of course follows without further ado. But in the program in which trees are first computed and then sent to the semantics, this principle needs to be stipulated - and as such there is no explanation for why it exists. If one believes that the input to semantic interpretation are trees which are computed by the syntax, why should we all be so proud of the fact that we can Curry (aka Shonfinkelize) the meaning of a transitive verb? Why not state the meaning of *Mitka licked Kolya* on a chunk of representation that includes both subject and object? The fact that the syntax and the semantics work on similar objects is a complete mystery under the view that they are divorced from each other.

Second, there is a clear cost to the move of divorcing the syntactic combinatory rules from the semantic rules. The point is easiest to illustrate by a comparison of theory C to theory D, since these are otherwise most alike. Both theories contain an additional set of rules effecting a mapping between surface structure and LF; they disagree on the direction of the mapping. Crucially, in theory D, the semantic combinatory rules cannot be stated in tandem with the phrase structure rules (or their equivalents), and this means that the syntactic side of things must be stated twice: once as output of the syntax and once as input to the semantics. As illustration, take a case where there happens to be no rule of interest mapping between surface structure and LF (the two are thus the same). So consider the syntactic composition and semantic interpretation of a very simple case like *John walks*. Here is a "fragment" in C to do this, and a "fragment" in D:

(10) C. S ---> NP VP; $[[S]]^g = [[VP]]^g([[NP]]^g)$
D. syntactic rule: S ---> NP VP
semantic rule: $[[\ [_S\ NP\ \ VP]\ ]]^g = [[VP]]^g([[NP]]^g)$

Reference to the local tree $[_S$ NP VP] is required twice in D but only once in C. The same point can be made by a comparison of the C-fragment given in (4)-(6) to the D-fragment given in (7)-(9). In the D-fragment, the two rules (8) and (9) repeat large parts of the output of the transformational rule (7) which creates the appropriate structure to serve as their input.

One might think that this objection disappear once one moves away from "construction-specific" statement of the semantic rules. But actually, it doesn't: restating the semantic rules as more general schemata certainly ameliorates the situation, but it does not entirely eliminate the problem. Regardless of how general one makes these, one still needs semantic combinatory statements which provide interpretations for classes of local trees. Hence the semantic rules still need to refer to a set of local trees - even if in highly generalized forms - both in the input to the semantics and as output to the syntax. The fewer the rules, the less duplication there will be, but there will still be some and ths remains suspicious if there is an alternative theory which avoids this. And, in

fact, any theory which states the rules together can avoid this. There is one important caveat here: a theory with direct compositionality and/or with deep compositionality will find itself in the same "pickle" if it contains a number of specific syntactic rules and general semantic rules stated separately. Such, for example, was the case in some versions of early "type-driven" GPSG, where the syntax contained various phrase structure rule schemata combined with a general principle as follows: for any rule A ---> B C, if the semantic type of B is a function from C-type things to A-type things then the associated semantic rule will be functional application. Such a theory also contains duplication. (Note though that even this kind of theory is not vulnerable to the other criticism of D, which was it is an accident that the output of the syntax is the same general type of objects as is the input to the semantics. Even if the syntactic rules are stated in a particular format and the semantic rules are stated separately as generalizations over the syntactic rules, there is no mystery as to *why* the two involve the same kind of objects.) But while there are versions of A-C which suffer from the problem of duplicating information, there are also versions (e.g., some versions of Categorial Grammar) where the syntactic schemata can be stated in forms as general as the semantic combinatory schemata, avoiding this.[2]

The only way to demonstrate that D avoids unwanted duplication would be to demonstrate that the syntactic rules and the semantic rules *should* actually be stated in terms of very different objects or very different kinds of objects. This, for example, would be the case if the semantic rules interpreted chunks of non-local trees. Or, this would be the case if the semantic rules looked only at linear strings and not at syntactic structures. But, as mentioned above, no theory seems to maintain this and no one (to my knowledge) has found evidence that we need rules of this type. The claim that the rules operate on different objects could also be substantiated if one could show that the semantic rules took as their input a much larger or more general set of local trees than the syntactic rules give as output. If one could really make such a case, then divorcing the output of the syntax from the input to the semantics would be exactly the right move, but there has been little (if any) real arguments to this effect. (I've heard people argue for divorcing the syntax and the semantics on the grounds that we can interpret sentences which the syntax doesn't allow, and a version of this argument is found in Heim and Kratzer (1998). But when one thinks it through, this argues only for having certain aspects of the semantics predictable from the syntax. If the semantics can interpret some (ungrammatical) sentence, then it has to be the case that one can "imagine" some way that the syntax tripped up and allowed it to be proven well-formed (or, in the representational view of semantics - some way in which the syntax tripped up and assigned it a representation). But if one can imagine just how the syntax "goofed" to assign a sentence a representation in a non-direct compositional view of things, then one

---

[2] To be fair, the statement of syntactic and semantic rules in classical Montague grammar also contained a certain amount of duplication. While these were assumed to be paired, they were usually stated within a notational format where a schematized version of the output of the syntactic rule was given as the input to the semantic rule. But this was really a notational oddity: the general program of having the syntactic and the semantic combinatorics linked does not *require* this notational awkwardness, while this duplication is essential in a theory like D.

can equally well imagine - in a direct compsoitional view - the syntax "goofing" to predict that something is well-formed. Either way, one needs a way to have a predictable semantics to go along with the goofy syntax. But this doesn't bear on the question of direct compsoitionality. See Jacobson, 2002 for a more detailed discussion of this point.)

1.3. Direct compositionality (in the broad sense): is there any reason why not?

I have tried to show above that - all other things being equal - Direct compositionality (in any of its versions) is simpler, more elegant, and more explanatory than a non-direct compositional view. Are there any real arguments for D over its competitors? In Jacobson (2002 - *L&P 25)* I tried to show in some detail that the answer is no: arguments which have been put forth in the literature for D at best argue only against Strong Direct compositionality. Unfortunately, much of the modern discussion has pitted these two positions against each other - completely ignoring any of the intermediate positions such as the Weak Direct compositionality of classical Montague grammar.

Space precludes a detailed discussion of this here, but let me briefly mention a few points of relevance. One of the classical "textbook" arguments for the existence of LF is the need to provide a treatment for wide scope quantification in general. But we have seen above that any of the Direct Compositional theories can do just that. The next sort of "textbook" argument that one finds over and over concerns the commonly accepted lore that both wide scope quantification and *wh* movement are subject to island constraints - and so both must involve movement. Now in the first place, it's not clear to me that common bit of wisdom is really true. Granted, wide scope quantification is generally difficult out of islands, but it is also difficult to get wide scope readings in a variety of embedded contexts. In that respect, wide scope quantification seems to actually obey much stronger constraints than the constraints on *wh* movement. For example, recent researchers seem to agree that a *no book* cannot scope over the matrix subject in (11a), while *wh* movement is of course possible:

(11) a. Some man believes no book is lousy.
b. What does some man believe is lousy?

Of course merely pointing out that QR is subject to additional constraints does not in itself invalidate the claim that it is also subject to island constraints, but it is perfectly possible that the island effects will simply follow from the more general constraints at work here. On the other side of the coin, it has often been observed that scoping out of, e.g., relative clauses is quite difficult but not crashingly bad the way corresponding *wh* movement cases can be. Thus the correlation between *wh* movement constraints and quantifiers scopes is sufficiently imperfect as to make one suspect that something else is going on here.

But for the sake of argument, suppose that the conventional wisdom is true. Does this necessitate acceptance of D over B or C? Certainly not. Generative semanticists, in fact, very loudly proclaimed this very observation as an argument for a movement rule of

Quantifier Lowering. Their battlecry in the late 60's and early 70's was that since quantifier scopes obeyed the same constraints as movement, it must be a syntactic movement rule. For a clear statement of just this (appearing in an accessible and "mainstream" journal), see Postal (1974), especially p. 383.

This way of putting it of course assumes the classical Generative Semantics view - according to which *wh* constructions do indeed involve movement. Can the constraints on *wh* constructions and on quantifier scopes also be collapsed in Weak Direct Compositionality? The answer is of course; relevant discussion of just this point is given in Rodman (1976) (Rodman's discussion focuses on relative clauses rather than questions, but the extension to questions is obvious). Of course to fully demonstrate this we need to give a Weak Direct compositional account of *wh* questions; space precludes doing this here (see Karttunen, 1979 for one such account, but with this in hand it is perfectly easy to state the constraints so as to collapse both *wh* constructions and constructions with quantifiers.

To make an argument for the approach in D as opposed to these alternatives, one would need to show two things. The first is that island effects are most simply accounted for in terms of constraints on *movement*. (This is still compatible with the Generative Semantics type solution, but would rule out the non-movement accounts of *wh* constructions taken in Karttunen.) Actually, there was considerable discussion of just this question in the syntactic literature in the mid and late 1970's, but it was inconclusive. See, for example, Bresnan and Grimshaw (1978), who show that even a Subjacency-inspired approach can be recast so as to constrain deletion as well as movement. Second, one would need to show that the constraint will apply in the right ways only in raising and not in lowering situations (or only if both wide scope quantification and *wh* movement have to involve movement in the same direction). Such a demonstration would mean that quantified NPs (like *wh* NPs) must move up the tree. This in turn means that scopes must be given by QR, and that in turn entails a model like D. But I know of no demonstration of this.

There are, of course, other kinds of arguments which have been given for D (and in some cases specifically also against Strong Direct compositionality), but the logic of these arguments often goes as follows. It is <u>assumed</u> that there is some constraint in the grammar which refers to global chunks of representation. We then see that the relevant representation is not found on the surface. Hence we posit an LF with the relevant representational properties. But - as disucssed above - under strong direct compositionality (and in fact under various weaker versions too) we would not expect to find constraints of this type in the first place, suggesting that the original analysis was incorrect. So, in the remainder of this paper I want to turn to some phenomena of just this type: phenomena which are assumed to require representational statements and which in turn seem to require abstract levels like LF at which the appropriate properties are found. What I will argue with this case study is that thinking in representational terms only gets in the way of formulating the compositional semantics.

2. A Case Study: "Reconstruction", Pronoun Binding, and Condition C effects

We thus now turn to some phenomena that have been used in support of some abstract representation (such as LF and/or silent elements at surface structure) which input the model-theoretic semantics and in support of principles stated on representations. These center on the notion of "reconstruction": the claim that material which is in an extracted position plays a role for the purposes of interpretation as if it were in the position of the gap. It should be noted that the earlier work on this assumed that things moved and then were put back for the purposes of interpretation (it seems to me that any theory which needs to posit that something moves and then moves back should really be embarassed and wonder if it is not missing something), while more recent work assumes that movement leaves a silent copy; I will basically be focussing on this view of "reconstruction". But before turning to the arguments themselves, I should point out that even if these arguments did go through, they would still be perfectly compatible with certain versions of direct compositionality. For example, if "moved" material really played a role in interpretation in its pre-movement position, wouldn't we be tempted to adopt a Generative semantics view of things - where interpretation happened *before* movement? Moreover, I can imagine versions of the "silent copy" view which could be framed in Weak Direct compositional terms. However, I will not attempt here to look at all logical possibilities - rather I want to claim that the arguments for a "reconstruction" view of any type really do not go through, nor do the arguments for representational statements.

2.1. "Binding" into extracted constituents

The first phenomenon to be considered here concerns the well-known fact there are cases where there is a pronoun within an extracted constituent which appears to be "bound" by material that c-commands the extraction site but not the surface position of the fronted material. Note first that this goes hand-in-hand with the fact that these induce functional readings (see especially Engdahl, 1986):

(12) a. Which of his_i relatives does every man_i love the most?
b. Which woman that he_i loves would no man_i forget to invite to his wedding?

Conventional wisdom is that in order for a pronoun to be "bound" by a quantified DP, it must be c-commanded by that DP at the level of LF. (The c-command condition on binding is actually cited in two contexts. One centers on the assumption that the account of Weak Crossover effects require a "binder" (or perhaps a trace of a binder) to c-command a bindee at some level of representation (see, e.g., Reinhart, 1983). I am not concerned with this use of a c-command condition here; see Jacobson (1992, 1999) for detailed discussion of a non-representational account of WCO effects. Rather I am concerned here with the common assumption that there is no way for a pronoun to be "bound" by a quantified DP - i.e., no way for the compositional semantics to yield the relevant interpretation - unless it were c-commanded by that DP at LF.) Given this, the conventional reconstruction story is again that there must be a copy of the fronted material at the level of LF.

Because the interpretation of questions involves additional complications concerning the semantics of the *wh*-expression, I will turn instead to the parallel facts with respect to relative clauses. Thus we get this kind of unexpected "binding" of a pronoun in head position not only in functional questions, but also in DPs which allow for functional interpretations - as, for example, in copular sentences:

(13) a. The relative of his_i that every man_i loves the most is his_i mother.
b. The woman that he_i loves that no man_i would forget to invite to his wedding is his_i mother.

(I am assuming in (13) that *woman* is first modified by *that he loves,* and this - which I will call a (complex) N - in turn is modified by *that no man would forget to invite to his wedding* - hence here *woman that he loves* is a complex head of the larger complex N.) There are, in fact, a variety of connectivity effects found between the heads of relative clauses and the gap position; these have been discussed since at least as early as Vergnaud (1974). Accounts of these facts differ, but for the sake of discussion I'll focus on a head raising analysis whereby the head raises from the position of the gap and - under copy movement - thus also leaves a copy of itself in the position of the gap. This makes the analysis of connectivity effects in *wh* questions and in heads of relative clauses parallel. The point, then, is that paper after paper contains the following chain of reasoning: we know that a pronoun must be c-commanded at LF by its binder. In (13b), *no man* binds *his*. So *his* must be in the position of the gap at LF - one way to accomplish this would be by copy movement whereby there is a copy of the entire moved head *woman that he loves* in the position which is the object of *invite*.

But the slogan that a "binder" must c-command a "bindee" at LF has no empirical teeth. Rather than worry about notions like binders, bindees, and representational properties like c-command, we should be worried about getting the semantics right - and indeed this can be done at least as easily (arguably, more so) without reference to representational properties and difficult to define notions like "binders". After all, the notion of a "bound pronoun" is just a shorthand for a way to talk about a particular meaning for a sentence involving a pronoun (and I will not spell out this meaning in any detail, since we all know what is meant by the "bound" meaning). So what would it mean to say that a binder must c-command a bindee at LF? Assuming that we have successfully defined these notions (to which I return below), this could mean two things: (a) that this is not logically necessary but is a constraint on LFs and hence is stated as a principle in the grammar, or (b) this is logically necessary - once we have an appropriate definition of "binder", there could be no other way to get the compositional semantics to work out right. I think that the usual assumption is that (b) is the case. So what I want to show in the discussion below is three things. . First, it is not the case that there is no way to get the compositional semantics right without the c-command condition. This fact is true regardless of whether one buys into the variable-free program which I have advocated in a good deal of previous work or whether one adopts the more standard view of things. Second- even under the standard views of how

pronouns are interpreted - there is no really natural notion of a "binder" such that *every man* in (14) "binds" *his:*

(14) Every man$_i$ loves his$_i$ mother.

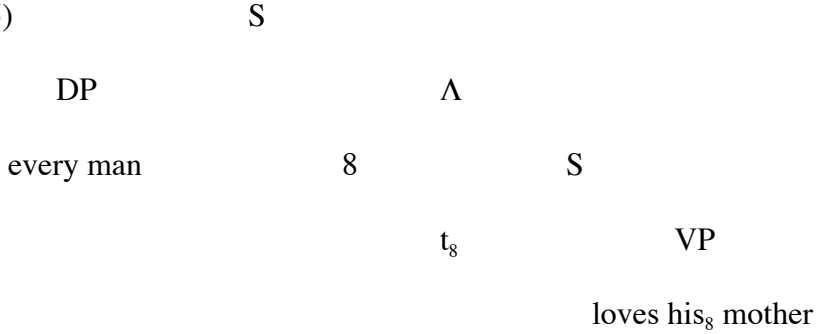
(Here and throughout I am using indices just to indicated intended readings rather than as devices that actually occur in the grammar.) Of course one can define a notion of "binding" such that *every man* "binds" *his*, but this notion plays no useful role in the semantic composition and hence does no empirical work. And third, positing a reconstruction analysis for (13) actually is not a particularly simple way to get the compositional semantics to come out right - just "copying" the head back into the position of the gap is not enough, much more work needs to be done. The bottom line is that phrasing things in purely representational terms (binders must c-command bindees) rather than working on getting the semantics right leads us astray.

A. Defining "binders"

Let me begin with the second point first: that there is no obvious sense in which *every man* binds *his* - and hence no reason to care about whether it c-commands it. (One can define things in such a way that a "binder" c-commands a bindee at LF by definition, but then, who cares? That would just mean that in (13a), *every man* is <u>not</u> the binder for *his*. As long as we can give a meaning for (13), that result would be completely unproblematic.)

I think that this point can be brought out by a consideration of the definition of "binding" given in Heim and Kratzer (1998). I pick their exposition only because they have actually gone to the trouble to define this notion and - once one looks at how complex the definition is - one can question whether its a notion used in the grammar at all. (I use their notation, but one can pick any other reasonable notation and I think the same points will hold.) Thus HK posit an LF like (15) for (14):

(15)



What does it mean in semantic terms to say that *his* is bound? Following the standard story with variables, take each expression to have as its semantic value a function from assignment functions to something else. (HK use partial assignment functions and there is a good reason for this, but to simplify the exposition I will use only total assignment functions.) What this means, then, is that when a pronoun is <u>unbound</u> within some

expression C, the value of C can be a non-constant function from assignment functions. When all pronouns are bound, then the value of C must be a constant function. (Note that it is not the case that whenever we have a pronoun that intuitively we would like to call "unbound" we necessarily have a non-constant function - the monkey wrench in this concerns tautologies and contradictions. Consider for example *he_8 lost or he_8 didn't lose*. This will assign all assignment functions the same value. This is the advantage of having assignment functions be partial: it does give a nicer way to define "unbound" pronouns. But again we won't go down that road here.)

Let G be the set of assignment functions, and let G/i be any subset that agrees on the assignment to all variables except i. Then, as noted above, a pronoun *he_i* is "bound" within an expression C if - for all G/i - [[C]] is the same for each member of G/i. Thus the "binding" of a pronoun *he_i* is really the process of shifting a meaning from a (in general) non-constant function on a G/i to a constant function (for all G/i). Indeed this is essentially the definition used by HK (modulo the fact that they are using partial functions); hence they first define what it is for a pronoun to be "free" within an expression C, and then define "sem-binding" as follows:

(16) An expression α sem-binds β if α is the sister of the smallest subtree in which β is free.[3]

So *every man* is <u>not</u> what "binds" *his* - it's the integer 8 whose semantic function is to perform λ-abstraction (and thus there are G/8 such that the sister to 8 is a non-constant function from members of G/8 to truth values, while the meaning of the expression I have labelled Λ is a constant function from the sets of assignment functions in G/8 (for any G/8) to meanings of type <e,t>). Let me stress that the fact that it is "8" rather than *every man* which "binds" *his* is not some quirk of some perverse definition on the part of HK: the natural definition of "binding" in semantic terms will end up with this result, regardless of the details and of the exact notation. But of course HK do want to be able to define a relationship between *every man* and *his* in order to state the usual "binding constraints", and so they then give a second definition of binding - what they call "a derivative notion of sem-binding", as follows:

(17) α "derivatively sem-binds" β if β and the trace of α are sem-bound by the same thing.

But - given a few other standard assumptions, it follows that if α derivatively sem-binds β then it c-commands β. Strictly speaking, this isn't true by definition but given all the usual other assumptions it could not be any other way. Thus the only thing that gets to sem-bind anything in this system is an integer (the integer whose semantics is what

[3] Note that one should in any case be suspicious of a theory which needs to make definitions like this. The reason is that they require reference to the output of the compositional semantics. But such a definition is useful only if it plays a role somewhere else in the grammar - and so if it is to be used anywhere else in the grammar we will need simultaneous reference both to the output of the compositional semantics and to whatever other level makes use of this notion.

performs λ-abstraction), and that integer necessarily c-commands both the trace of α and β. Moreover, this integer is always going to be the sister of the DP whose trace is at issue, so that DP will necessarily c-command β. So yes, a "derivative sem-binder" of a pronoun must c-command it at LF, but we have gone to an awful lot of definitional work to make this so.

Have we accomplished anything empirical with these defintions? No - the only empirical point that would be relevant here would be the discovery that there is no other way to get a compositional semantics for sentences like (13) without a reconstruction structure (i.e., without a level at which the pronoun - or some copy of it - is "derivatively sem-bound" by the quantified DP).[4] So we should now turn to that question - is it really true that there is no way to give a compositional semantics for (13) without making sure that *every/no man* c-commands *his* at LF?

The answer is: absolutely not. There are perfectly good ways to get the right meaning for (13) without positing c-command; Engdahl (1986) - using a standard semantics with variables - provides one; Jacobson (2002) provides another which can be seen (very roughly) as a translation of Engdahl's into a variable-free semantics (with the advantage that what seemed very awkward and stipulative in Engdahl's analysis becomes completely natural[5] in the variable-free translation); and Winter (2002) provides another variable-free analysis which turns out to be the inverse of Jacobson's (and is also very natural). Since I have become a native speaker of variable-free, I will use my own system to exposit the point, but one should keep in mind that the whole matter can be recast with variables (Engdahl's solution, for example, uses variables).

Before continuing, a very short crash course in the variable-free program developed in, e.g., Jacobson (1999):

A. There are no assignment functions. The meaning of any expression with a pronoun which is "unbound" within that expression is a function from individuals to something else. (If there are two unbound pronouns, its a function from two individuals.)

B. The meaning of a pronoun is thus - as one would expect - a function from individuals to individuals. In particular, it is the identity function.

---

[4] I don't mean to suggest that HK go to all this trouble for nothing. They are assuming that other "binding constraints" in the grammar - e.g., Principle B - will need to refer to these defintions. In other work, I have tried to show that the account of Principle B effects using configurational definitions like this is extremely complex, but that these effects can, instead, be accounted for (without indices and without representational constraints) in direct compositional terms.

[5] I am admittedly using this term in a kind of gut-feeling way and it would be wonderful to have a real definition of it. However, I think that everyone would agree that there is something "natural" about the operation used in this analysis: it is an operation which maps a 2-place relation (a set of ordered pairs) into a set of functions and it is the minimal and obvious way to accomplish this mapping. Take any set of ordered pairs R and then take its power set. Take the members of $\mathcal{P}$(R) which are functions, and voila - this is the set of functions which is needed here.

C. Since an expression with a pronoun within it will be a function of type <e,a> - how does this get to combine with functions looking for a-type arguments? For example: *he* is a function of type <e,e> (the identity function). *Lost* is of type <e,t>. How do we get *he lost?* Answer, the "Geach" rule, which shifts the meaning of any function:

Let f be a function of type <a,b>. Then **g**(f) is a function of type <<c,a>,<c.b>> such that **g**(f) = $\lambda h_{<c,a>}$ [λC_c [f(h(C))]]. So *lost* can shift to be an <<e,e>,<e,t>>; it combines with *he* and the result is of type <e,t> and is the function from individuals to the proposition that they lost. (Modulo the contribution of gender, it has the same meaning as *lost). his mother* will combine up in such a way that this is the-mother-of function; and *his mother lost* is a function from individuals to the proposition that that individual's mother lost. So, whenever we have unbound pronouns, this means we keep "Geach'ing" on functions all the way up (I also allow free type lifting, so that any would-be argument can switch to become the function.)

D. How do we ever get the effect of "binding"? By one more type-shift rule, which I call **z** which is as follows:

Let f be a function of type <a,<e,b>>. Then **z**(f) is a function of type <<e,a>,<e,b>>, where **z**(f) = $\lambda h_{<e,a>}$ [λx_e [f(h(x))(x)]]. NOTE: Any theory needs something to do "binding". The usual way is to do λ-abstraction, which converts an open (on members of some G/i) function from assignment functions to propositions to a closed (on all members of any G/i) function from assignment functions to properties. So there is no extra rule here - just a different one. The only "extra" bit of apparatus is the Geach rule - which is (a) arguably quite simple and natural; and (b) has lots of other payoffs, such as allowing us to get paycheck readings of pronouns for free (see Jacobson, 2000). I've argued in my other papers that there are lots of other benefits to this view of things.

So, both to illustrate the program a bit more and to provide some necessary background to the main task at hand (giving a compositional semantics for the binding of a pronoun within a head), let's first take a case of a functional copular sentence that doesn't have a pronoun in the head, such as:

(18) The woman that every man loves the most is his mother.

In Jacobson (1994) I argued that *woman* can shift its meaning to denote the set of functions whose range is women, and that - with this - everything else about this sentence comes for free from the apparatus discussed above for "binding" in general. *loves* here undergoes **z** and then function-composes with *every man* and so the relative clause is the set of functions f such that every man z-loves f (which is the set of functions f such that every man is an x who loves f(x)). This intersects with the meaning of the head - so now we have the set of functions f whose range is women and which are such that every man z-loves f. The post-copular constituent automatically means the-mother-of function, and so the full S says that that function is the unique function satisfying the pre-copular description. Incidentally, many of the points here are independent of the variable-free implementation; von Stechow (1990) proposed the analogous analysis

within a variable-full framework; showing that most of the bits come for free just from an extension of the analysis of functional questions in Groenendijk and Stokhof (1983) and Engdahl (1986). The same point is made in considerable detail in Sharvit (1999). I believe that any theory will have to give essentially this kind of analysis of (18) - that is one will treat this as an equation among functions; and any theory with essentially the Groenendijk and Stokhof/ Engdahl analysis of functional questions will get most of the pieces of this directly from that analysis. The reason for phrasing this in variable-free terms is that functional phenomena in general are completely unsurprising here: the fact that the "gap" after *loves* is a "function" gap requires no extra machinery. (Thus, for example, in the standard view we need to posit that there are simple traces - which correspond to individuals - and complex functional traces - which correspond to variables over functions of type <e,e> applied to a variable over individuals.) Here the very mechanism used for binding in general automatically gives us the functional gap: *loves* can undergo the **z** rule such that it is "expecting" an object argument of type <e,e>.

But unfortunately, the binding of a pronoun in the head is not automatic under this analysis (nor is it automatic in the variable-full analogues; a point which emerges in detail from Engdahl's discussion). The relative clause *who he loves* will denote just a two-place relation; both the pronoun and the relative pronoun and/or gap will contribute an argument slot). It happens that the particular two-place relation we end up with is the "is-loved-by" relation (for discussion of this point, see Jacobson, 2002), that is the relation $\lambda x[\lambda y[x \text{ loves } y]]$. Thus take something like (19), which is a simplification of (13b) just to strip away some irrelevant bits of the sentence:

(19) *the woman that he loves that no man invites*

(that) no man invites has already been discussed: it is the set of functions f such that no-man z-invites f. This should intersect with the head, which should denote a set of functions. We have already hypothesized that *woman* can denote the set of functions f with range women. But what do we do about *that he loves?* It's a function of type <e,<e,t>> but to fold easily into the semantic composition we want it to also be a set of functions.

But it turns out that there is a perfectly simple and "natural' way to map a function of type <e,<e,t>> into a set of functions. Just de-Curry it to a set of ordered pairs R, take the power set of R, and then take all members of $\mathscr{P}$(R) which are functions, and we have exactly what we need. Thus *that he loves* will now denote the set of functions f such that each x loves f(x). I will call this shift **m**; it is spelled out more generally in (20)

(20) Let F be a function of type <b,<a,t>>. Then **m**(F) is a function of type <<b,a>,t> such that $\mathbf{m}(F) = \lambda h_{<b.a>} [\forall x_{\text{ in the domain of h}} [F(x)(h(x))]]$, where h is a partial function from b to a.

Notice that this means that *he* in *he loves* is not in any straightforward way "connected" to or "bound by" *no man* and some people have thus referred to this strategy as allowing

he to be "pseudo-bound". But I think that's somewhat misleading - none of the pronouns are "connected to" or "bound by" things like *no man* in this way of thinking; the "binding" of a pronoun by a DP is not a particularly useful way to think of things. But, as I've argued above, the same is true in the standard view too.

Of course a reconstructionist could immediately object as follows: "You've managed to get the compositional semantics to yield the right meaning - but at a cost - you've had to add yet another type-shift rule." This is true - but now let us see whether the "reconstruction" view fares any better: does having a copy of the pronoun in the gap position automatically get us the right meaning? No. In fact there are a number of complications that arise in the reconstruction analysis, so let us now turn to this.

First, we need to say something about the interpretation of traces with lexical material in them; I follow especially Fox (2000) and others on this. Thus the idea is that the lexical material restricts the assignment function and restricts interpretation only to those assignments g for which the variable in question is a member of the set denoted by the head noun (on g):

(21) $[[\ t_{i\text{-}N}\ ]]^g$ = g(i) if g(i) ε g(N), and undefined otherwise

So what happens if the head noun contains a pronoun? Suppose for example we have something like *the woman he loves that every man invited* where the trace in the object of *invite* is $t_{i\ \text{-woman he(j) invited t(i) (woman)}}$ I don't think we need to worry about the interpretation of the most deeply embedded trace there; we can take the value of the whole thing on g to be g(i) if g(i) is a woman that g(j) invited (and undefined otherwise). At the end of the day, this will not give us the right reading: this is not the functional reading that we want for (13), and rather is the unique individual who is a woman such every man invited her and every man loves her.

Problem 1: blocking an individual reading. Thus we need to do something else to get the functional meaning, but first let me point out that a problem arises right away: do we ever want to allow the meaning shown above as a possible meaning for *the woman he loves that no man invited*? I don't think so; it would make it a mystery as to why these couldn't stand in for ordinary individuals in ordinary individual positions like:[6]

(22) a. I like the woman he loves that no man invited.
b. I like the relative of his that every man invited.

(a), for example, cannot mean that I like the woman that no man invited that no man loves. Thus unless there is some independent explanation for the lack of the individual reading here, this is a serious problem for the reconstruction analysis.

(Possible) problem 2: getting the functional reading

[6] At the Workshop on Direct Compositionality held at Brown in 2003, Danny Fox answered this argument by claiming that one can tell a pragmatic story for why this reading is absent for (22). I unfortunately can't reconstruct his story; so hopefully he will reconstruct it for us at this workshop.

So we still need to get the functional reading for the full DP *the woman he loves that no man invited* (the same remarks will hold for the functional question case). For starters, let's take Chierchia's (1991) trace-based implementation of Engdahl's analysis: Chierchia's mechanics involve giving the trace two indices, one of which corresponds to an indexed variable over functions of type <e,e> and one which corresponds to a variable over individuals, which is the argument of the function. Let's notate a trace of this sort as $t_{f(-i)\,(j)}$. f-i is the ith variable of type <e,e> and j is the jth variable of type e. We still need to fold in the linguistic material from the head into this kind of a complex trace (and show how it is interpreted).[7]

But before doing this, recall that Engdahl showed that in functional questions, there can (to use her terminology) actually be variables over n-place functions from individuals applied to n variables over individuals. (The observations giving rise to this follow immediately in the variable-free analogue; we simply have multiple applications of the **z** rule, see Jacobson (1999) for discussion.) For example, a case of a function of type <e,<e,e>> applied to two individual variables is shown in the functional question in (23) (to make life easy, imagine a world in which all students are male and all phonology professors are female)

(23) Which paper that he handed in for her class did every student want every phonology professor to like? (Ans: the first one that he handed in)

We can recast this into the relative clause setting:

(24) The paper that he handed in for her class that every student wanted every phonology professor to like the best was the first one that he handed in.

Since the trace can therefore be arbitrarily complex - consisting of an indexed variable over any n-place functions into individuals and any number of indices corresponding to the individuals (which serve as argument of the function variable), it seems to me that we will want to (and will have to) approach the combination of the lexical material and the indices compositionally. That is, take a simple functional case like the DP in

(25) The woman that every man loves is his mother.

We have said that the trace after *loves* is something like $t_{f\text{-}i\ (j)}$. What I'm suggesting is that this can all be seen as a shorthand for something with more structure and where the lexical material associates with the index *f-i*; if we build into the syntax of the representation some way of showing that this is a function which will take the j-th

---

[7] The discussion here assumes the Groenendijk and Stokhof/Engdahl analysis of functional questions and considers the implications of this for the interpretation of a complex trace. In recent work Sauerland (2003) explores a different tack in which complex traces involve a choice function as part of their interpretation. I will not have space to discuss this proposal here.

variable as argument we can have a recursive procedure which will interpret traces of arbitrary complexity. Thus the trace is actually $t_{\textit{f-i [woman] (j).}}$ If this is right, the interpretation for the function-variable part of the trace is:

(26) $[[\textit{f-i - [N]}]]^g$ = g(f-i) provided that the range of g(f-i) is [[woman]]; undefined otherwise.

In other words, I am using a translation of the view above whereby the semantic contribution of *woman* is to restrict the range of the function. This does require an extra rule (in other words, this contribution doesn't follow from the semantics given for traces in general), but I see no way around having an extra rule for the interpretation of these structured traces. The rest should follow once we have a general way to notate in the syntax that this part is a function and the other index (or, indices) are arguments, and so it will follow that: $[[\text{f-i [woman] (j) }]]^g$ = g(f-i)(g(j)), provided that g(f-i) has the set of women as its range.

All well and good - but this does not give the right semantics for the bound pronoun case! Take *the woman* he_j *loves that every man invited*. The idea is that *woman* he_j *loves* will appear (as a copy) in the position following *invite*; this will restrict the range of the function picked out by the function variable (according to the assignment function in question), and he_j will correspond to a variable which is "bound" in the normal way - a structure like $\underline{t_i \text{ invite } t}_{\text{f-i -[woman that he(j) loves] (j)}}$ will undergo λ-abstraction over j. But what this semantics does is restrict the range of f-i in such a way that for all x, f must map x into a woman that g(j) loves. But that's not what we want - we want the function to map each x into a woman that x loves. So - at least if we adopt this semantics - putting the pronoun "back" into the gap position doesn't help at all.

Now one might object that this problem is just an artefact of how I set up the semantics: I tried to make the semantic contribution of the head be such that it is a restriction on the range of the function (which corresponds to one of the pieces of the trace). But one could have done it differently: which is to give a global interpretation of the complex trace. To make the syntax also match better, let's put the lexical content as part of the entire thing (rather than associating it with the function index):

(27) $[[t_{\text{f-i (j) - [N]}}]]^g$ = g(f-i)(g(j)) if g(f-i)(g(j)) ε [[N]], and undefined otherwise

This works, since now the interpretation of the trace $t_{\textit{f-i woman he-j loves (j)}}$ on some g will be g(f(i)) applied to g(j) just in case that result is a woman that g(j) loves. (So, for any assignment function, for each x we will be concerned with some woman that x loves.)

I might be being overly picky, but it seems to me that the rule in (27) is not a particularly happy rule to end up with. The problem is that it doesn't immediately generalize to the case of n-place functions applied to n individual variables, and the reason is that it is not compositional. The meaning of the entire complex trace is not gotten in a predictable way from some meaning associated with each of the indices. In fact, given that this is embedded in a theory in which the semantic composition is, in

general, predictable from the syntax, one would probably want the complex trace to have a structured representation, where there is a trace portion corresponding to a function and an argument portion as its sister. For a complex trace of, e.g., type <e,<e,e>> applied to two individual variables, one would assume a structure of the form [[$t_{f\text{-}i\ <e,<e,e>>}$- $t_{j\ ,<e>}$] $t_{k\ <e>}$] - etc. But, as we see above, if we try to let the lexical head just make a contribution to the function part of the trace we don't get the pronoun bound in the right way, and so there is no way to get the right result unless we have a rule like (27) which globally interprets all of the pieces of a complex trace at once. In either case, though, we should note that "reconstruction" certainly has not automatically gotten us a meaning in which the pronoun is "bound" - we still need to add in conventions for these complex (functional) traces and conventions for interpreting them.[8]

Problem 3: What is the head doing upstairs?

Now comes another question about the entire reconstruction analysis: what is the semantic contribution of the head upstairs? Do we ignore it? Do we let it make its contribution twice? Notice that the question would not arise if we were to adopt a Generative Semantics-like view of this (combined with head raising). Let's suppose that the head starts out in the embedded position, is a restriction on some variable or pronoun there (or, perhaps, an indexed determiner *the)* and that interpretation happens completely at that level, before any head raising applies. But this view incompatible with the general theory under consideration here.

So if we are looking at a view in which a copy of the head is in the trace position but the head itself (even when complex) is also in the upstairs position, it would, I think, be fairly complex to formulate a compositional semantics in which the upstairs head is just ignored. (I may be wrong on this, but I'd like to see it all spelled out.) Normally it will do no harm to interpret the head in both places. Take:

(28) *the woman who Bill loves* $t_{i\text{-}[woman]}$

The contribution of *woman* in the trace position is such that *who Bill loves* will denote the set of x's loved by Bill such that x is a woman. If this intersects with [[woman]] then no harm is done. Essentially, having the head intersect with the set denoted by the relative clause is just the same as having that set restricted in the first place to individuals satisfying the head description. So we don't have to worry about the fact that we might be interpreting *woman* twice.

But in the case where there is a pronoun within a complex head, we would be in trouble. Putting it into the trace position allowed us a way to "bind" that pronoun (provided we are happy enough to live with the interpretation rule in (27)), but we also

[8] I suppose that one might try to collapse (27) with (21) (the rule for interpreting plain old individual traces) by something like the "trick" that Engdahl used to collapse functional and individual questions: invoke a notion of 0-place functions and say that the interpetation of an ordinary individual trace is just a generalization of (xx) for the case of a 0-place function. Some people seem perfectly happy with this notion, but I find the idea of a 0-place function to be a terminological trick.

end up with a free pronoun upstairs. (Moreover, in the functional case in question, we still need a way to shift the head into a set of functions where the head denotes the range - this is exactly analogous to the extra step that we need in the variable-free view.) The upshot would be the following. Given an assignment function g, *the woman he$_j$ loves that every man invited* should denote the unique function f which is such that: every man invited f(x) and for all x (f(x)) is a woman that x loves and f(x) is a woman that g(j) loves. As far as I can tell, this is not a happy result - the bottom line, then, is that we have to make sure to ignore the semantic contribution of the head in its "upstairs" position. Perhaps that is not a problem - but in the absence of a full-blown theory of how the semantics interprets these structures, we should certainly be suspicious.

Problem 4: Multiple pronouns/Multiple binders

Finally, what I think is probably the most serious problem of all is that the reconstruction story just can't give a meaning for the full range of cases (at least not without severe complications). As discussed in detail in Jacobson (2002), the "binder" and the pronoun can be in either order within the case of stacked relatives:

(29) a. The woman he loves the most that every man invited is his mother
b. The woman that every man loves that he invited is his mother.

The case we have been considering is (a), but (b) is fine too (I thank Chris Barker for this observation). Yet if *woman that every man loves* is the head here, then "reconstructing" this into the gap following *invite* won't do any good to get *he* bound.

There is a potential way out: suppose that the structure for (29b) does not involve stacking but that rather *that he invited* is an extraposed relative, extraposed from the trace position following *loves*. Then *every man* can happily "bind" *he*. Of course we also have to allow for the stacking (plus reconstruction) analysis for the (a) case - but I see no reason not to think that both are possible, so both binding patterns could occur. But this alleviates the problem only for a brief moment. For we can also get both binding patterns within a single sentence:

(30) The paper that every student$_i$ handed in to her$_j$ that every phonology professor$_j$ most praised him$_i$ for is the one that he handed in to her the first week of class.

(The variable-free non-reconstruction analysis of binding into heads that I gave above also doesn't immediately give an analysis of this, but it is straightforward to extend it in such a way that it does; see Jacobson, 2002 for the relevant extension.) I see no way to get (30) under reconstruction. If *that every student handed in to her* is part of the head and has thus left a copy after *for* we can bind *her* - but this leaves us with no way to get *him* (the object of *praise)* bound. If we say that *that every phonology professor most praised him for* is an extraposed relative and is extraposed from the position which is the object of *handed in,* then we can get *him* bound, but not *her*. Unless I am missing something, (30) is at least close to a death knell for the reconstruction account.

2.2. The rest of the Story: Enter "Condition C"

The intent of the above discussion was to cast doubt on a reconstruction story about the "binding" of pronouns in heads (and in *wh* constituents), and on the representational assumptions ("we need c-command") that underlie this story. But I have told only a part of the story: there is a series of interesting facts developed in Lebeaux (1990), Fox (2000) and others concerning the interaction of pronoun binding (in extracted constituents) and Condition C effects. The story is intricate and intriguing - and just might make one tempted to believe the reconstruction account of pronoun binding despite the problems above. Thus I will conclude this paper with some discussion about these facts - and about just what they do and do not show.

By way of background, the account first assumes the following (the basic claims here are due originally to Friedin, 1985 and Lebeaux, 1988, 1990):

(1) It is claimed that when extracted material consists of a head noun + complements, the entire material behaves as if it is in the position of the gap with respect to Condition C effects. Thus the "received" judgments are that the following is bad (the * here indicates that judgment, not mine):

(31) *Which picture of John does he like the best?

(2) Moreover, in contrast to this, there is supposedly no Condition C reconstruction effect if the extracted material contains an adjunct). (Note that the grammaticality of cases like (32) was noted at least as early as Postal, 1970. used by him to argue that "Condition C" (known under a different label at that time) held post-cyclically.[9])

(32) Which picture that John took does he like the best?

Before continuing, I would like to point out that I strongly believe that the conventional wisdom here is a fiction, and I have found informants over the years to be very consistent about this. Most students approaching this material without a previous theoretical position on it (but students who are perfectly sophisticated about what it means for something to be bad) are quite surprised to hear that (31) is supposed to be bad, and even more surprised to hear that there is supposed to be a significant contrast between (31) and (32).[10] I do think there is something slightly "off" about (31), but the

---

[9] The way he actually put this was that he assumed that the "Principle C" effect derived from a condition on a pronominalization transformation, and he thus argued that this transformation applied post-cyclically, after the application of *wh* movement).

[10] There is no doubt that the contrasts originally cited in the literature are correct - these contrasts were essentially:

(1) a. *Which claim that John was a fool did he dispute?
b. Which claim that John made did he later dispute?

violation is extremely mild, and certainly not at all like the violation found in *?*He likes the picture of John that was on the mantle*.[11] Note that the claimed argument/adjunct asymmetry predicts that there should be a significant contrast between the following two:

(33) a. *Which article about John's mother did he find the most infuriating?
b. Which article that was about John's mother did he find the most infuriating?

I have never encountered a single informant without a prior theoretical commitment who found any strong contrast here, and I have asked several people about these (probably somewhere between 12 and 20) over the years. What is interesting is that - when asked if there is a contrast between (33a) and (33b) - informants very often do report finding a mild contrast. But when asked which sentence is better, the judgments seem to be at best random (if anything, my impression is that more people find (b) worse, but I suspect that this if for irrelevant reasons). Granted, though, this is all based on unsystematic data collection and on impressions - some systematic and trustworthy informant research is sorely needed here.

In any case, the main discussion in this section focusses not on the contrast between (31) and (32) but between (32) and (34), contrasts which seem more firm to me. It should be noted, though, that the reconstruction story quite crucially does predict that (31) is bad, and if this is not true there is a real problem. Thus the common wisdom regarding the contrast between (31) and (32) goes roughly like this. Assume the copy theory of movement. Assume that condition C holds at LF. Assume that adjuncts can be "merged" with what they modify late in the derivation. Then in the case of a fronted DP which contains an argument of the head noun (as in (31)) the argument material will have to be there before movement. It will thus leave a copy, this copy is interpreted at LF, and hence we see a Condition C violation. In the case of adjuncts, though, the adjunct material can be added as a modifier to the head after movement, and hence there will be no copy of the ajdunct in the position of the gap at LF. I realize that there are slight variations on this story, but I think these will not effect the main point.

But recent literature has stayed away from these for obvious reasons: of course (1a) is bad - equally bad is (1c):

c. *Which claim that he was a fool did he dispute?

A claim of X is a claim of X , and one can't differentiate them and ask which one was disputed. More recent work (at least some of the more recent work) has taken care to avoid this problem, but then the facts become quite unclear.

[11] Even these remarks are somewhat inappropriate, since I think that conditions on where to use pronouns and where to use full NPs really cannot be discussed in isolation of discourse context; these are well-known to be highly discourse-dependent. However, we can put it this way: It is extremely difficult - perhaps impossible - to find a context in which *He likes the picture of John which was on the mantle* is good (on the obvious relevant understanding) - this holds even in contexts where John is under discussion and where *John* has been mentioned in a previous S (see Lasnik, 1972)). But it is much easier to dream up contexts in which (31) is good.

With this background, we can turn to the main event: the interaction of pronoun binding and Condition C. Thus Lebeaux (1990) (and see also Fox (2000) point out an interesting correlation between these effects: they note that the following shows a Condition C violation even though the offending name is in an adjunct:

(34) *Which paper that he_i handed in to $Joan_j$ did she_j ask every $student_i$ to revise?
(intended kind of answer: the first one that he handed in to her)

I do agree that this contrasts with (32) and that at least part of the problem has to do with the relationship between the pronoun and the name, since the following is (I think) a bit better:

(35) ??Which paper that he_i handed in to her_j did $Joan_j$ ask every $student_i$ to revise?

(The contrast is difficult to hone in on since there is so much else going on in these sentences, and I am not very secure that there is any significant contrast here, but I am willing to believe that there is.)

Since I've been concerned with heads of relative clauses, let me first translate the entire domain into the relative clause setting. Thus to the extent that there is something slightly "off" about (31) the same "offness" holds for (36) (again the * does not indicate my judgment; it indicates that I see no difference between this and (31)):

(36) *The picture of John that he liked the best was the one of him eating sausages.

The run-of-the-mill case where the head is complex and consists of a lexical head + adjunct (another relative clause) is, according to everyone, fine:

(37) The picture that John took that he liked the best was the one he took on Wednesday.

And finally, the badness (or "offness") reemerges in the functional relative case where there is a "bound" pronoun in the head (I've slightly changed the example just because I think it brings out the functional reading more easily)

(38) a. *The (only) assignment that he_i handed into $Joan_j$ that she_j didn't ask any $student_i$ to revise was the first one he handed in.
b. ??The (only) assignment that he_i handed in to her_j that $Joan_j$ didn't ask any $student_i$ to revise was the first one he handed in.

The idea, then, is that while adjuncts *can* be inserted late, they don't have to be. Assume the reconstruction story about binding developed in Sec. 2.1: the compositional semantics requires the pronoun to be in the position of the gap and thus to be part of the lexical copy of the trace and thus part of the restriction on the interpretation of the complex trace. Then the only way to get the appropriate reading for (38a) would be to have had the adjunct inserted early, where it leaves a copy. But in that case we have a

Condition C violation - since *Joan* will also be in the position of the gap (after *revise)* and will thus be c-commanded by the pronoun *she*.

As intriguing as this story is, there are a number of problems with it. The first is representative of the kind of line I have been trying to push throughout this paper: The story simply takes as an axiomatic assumption that Condition C is stated on representations (and requires reference to c-command). But do we really know that Condition C effects can be accounted for by this kind of constraint - that is, by a constraint on (syntactic, albeit LF) representations?

Answer: not only do we not know that, but we know that they *can't* be accounted for simply in this way. This has been discussed since at least as early as Reinhart (1983). The problem is that a constraint on representations could only access co-indexing - and a constraint on co-indexing is not sufficient to rule out Condition C violating cases. Thus consider a run-of-the-mill case like (39) (keep in mind that my use of indices here is meant to indicate the relevant understanding, not an actual device in the grammar):

(39) *He_i thinks that $John_i$ should be elected.

(Once again discussions of these cases should really put them in context - but, as pointed out in Lasnik (1972), even in a context where John is under discussion and has been mentioned prior - even in the very sentence before - these are still somewhat bad.) But - regardless of how one treats indices, binding, etc. - there is nothing in the grammar itself which would stop a representation for (39) in which *he* is a free pronoun which happens to pick out *John* - and yet we still don't like this understanding of (39). So no purely representational constraint is going to do the trick.

In fact then, it seems that Condition C effects are really poorly understood (although perhaps people who work on Centering Theory, discourse and information packaging have a better understanding of these). Until we have a better understanding of what underlies these effects, there is no reason at all to think that c-command (or other purely configurational properties) will play any role in the statement of these effects and hence no reason to believe that extending the account to (38a) will require reconstruction. All we know is that the effects *often* happen to line up with c-command (which probably has to do with fine-grained facts about the semantics of the sorts of configurations we tend to look at), but since we don't have know how to make a grammatical statement which specifically refers to c-command, we can hardly conclude that c-command must hiddenly be there in the cases where it doesn't seem to be. (I am not suggesting that explanations appealing to abstract things that we don't see can't be right - I am only suggesting that those make sense only when we have reason to believe that the particular properties that we don't see are actually <u>necessary</u> for the explanation.)

Nonetheless, let me pursue one sort of line which has been taken on these which does make a correlation with c-command relevant. This line (a version of which is found originally in Reinhart, 1983) is that Condition C effects arise as a result of a competition from another sentence in which the pronoun *can* be bound by the name (or other DP).

Thus the relevant understanding of (39) is blocked because the same understanding can be expressed by a sentence like (40) and, moreover, in (40) the pronoun can be a bound pronoun:

(40) John$_i$ thinks that he$_i$ should be elected.

Of course to actually formulate this we need to decide what this principle is a principle about. (Is it, for example, a constraint on processing?) We also need a definition of "bound" pronoun; for the purposes of this discussion we can take the HK definition discussed above. If something like this is correct, then the relevant condition itself doesn't actually refer directly to c-command, but it will line up with c-command: the competing sentence with a bound pronoun will be possible only when there is c-command (because of the definition of binding). Hence it would appear that the basic story told above still goes through. (38a) is blocked on the relevant reading by the competiton-based account of Condition C. Assume again that adjunct can but don't have to be inserted early (before movement). The reconstruction account of pronoun binding says that the functional reading of (38a) (where the pronoun *he* is "bound") is possible only if the adjunct had been inserted early (so that the entire material can leave a copy in the trace position and the index on *he* can be relevant to restricting the interpretation of the complex trace). But then consider the pronoun *she*. It should be able to be free and happen to pick up Joan, but there is a competing version of this sentence (the one in (38b)) where there is instead a pronoun "bound" by *Joan* and which gives the same meaning.

But wait. Then why should the normal case with an adjunct (where there isn't also another bound pronoun) be good? Why should this not also be bad? After all, adjuncts <u>can</u> be inserted early (the whole story crucially relies on this). So take (37). There is a version of this sentence, whose surface form is in (41), where the pronoun *he* is "bound" by *John:*

(41) The picture that he took that John liked the best was the one he took on Wednesday.

The relevant derivation of this sentence would involve inserting the adjunct early, having its copy sit in the trace position, and having *he* bound by *John* in the way that we get bound pronouns in heads in general. So the existence of this sentence and this derivation should compete with (37), and the latter should show a Condition C violation. I suppose that one can define the relevant competition set so that the derivation of (41) which involves early insertion of the adjunct doesn't matter (it doesn't compete with (37)), but the plausibility of the full account will depend on just how plausible and/or complex this turns out to be.

There are other problems. The facts are very delicate, but it seems to me that there is no obvious contrast between (38a) and cases like the following:

(42) *The (only) assignment that was handed into Joan$_j$ that she$_j$ didn't ask any student$_i$ to revise was the first one (he) handed in. (bad on the functional reading)

(43) *The only assignment that he$_i$ handed in that was for Joan's$_j$ class that she$_j$ didn't ask any student$_i$ to revise was the first one he handed in.

The first involves an interaction of functional readings and Condition C effects but contains no pronoun needing to be bound. The above account would extend to these only if we have some reason to believe that these functional readings always require the full head (including its adjunct) to be in the position of the trace, but I don't know what this would follow from. (42) is even more problematic - for here the pronoun that wants to be bound *he* is in a different adjunct from the offending name - so there's no reason why *that he handed in* couldn't be inserted early, with *that was for Joan's class* being inserted late (after the head has raised).

Finally, the Lebeaux/Fox account can be properly evaluated only when combined with a full and explicit semantics for the binding of pronouns in disclocated constituents (a task which I argued above is not easy), the way in which the semantics is put together with "late merger", etc. (For example, here's one question. Relative clauses are adjuncts. If we assume head raising but also assume "late merger" - then where does the internal head come from? Or, where does it go? Isn't the point of late merger that the relative clause itself is combined with the head only late in the derivation, after that head might have moved from some other position? But then how does the relative clause itself get constructed?)[12] In other words, we would need to see a full fragment with the semantics of all of the pieces worked out. Until then, it is difficult to conclude that these facts provide convincing evidence for reconstruction, nor that they provide convincing evidence for representationally stated constraints. I would thus argue that these cannot be taken as a serious challenge to direct compositionality. Rather, a faith in direct compositionality should inspire us to look for a more explanatory account of things like Condition C effects.

[12] I've used the Head Raising analysis for the whole discussion here. An answer to these questions might be to assume instead a "matching" analysis - the head itself does not raise, but rather the relative pronoun has lexical content associated with it which is identical to that of the head (see, e.g., Sauerland, 2003). But there are still many pieces of this kind of analysis which to my knowledge have not been spelled out; suffice it to say again that until we have a full fragment, we cannot really evaluate this proposal.