Direct Compositionality and Variable-Free Semantics:
The Case of "Principle B"

Drafty outline/handout of material to be presented at the
Workshop on Direct Compositionality

Pauline Jacobson
Brown University

0. The big picture goals:

(A) To argue that direct compositionality is a much simpler hypothesis about the organization of the grammar than the competing views

It is, therefore, the view we should hope is correct. Apparent challenges to this view should perhaps lead us to rethink some of the underlying assumptions which give rise to these challenges.

(B) To point out: a consequence of at least a reasonably strong version of direct compositionality is that *structured representations (e.g., trees)* **play no actual role in the grammar: they are not something that grammatical rules/principles can "see".** They are just convenient representation for the lnguist (representing the syntactic proof system and the compositional semantics)

(C) To point out the consequences of this with respect to principles regarding the distribution of pronouns (reflexives, etc.). Standard assumption: these require the grammar to include principles which are stated on *non strictly-local chucks of representation* Hence: the hypothesis of direct compositionality should inspire us to rethink such statements

(D) To discuss a second way in which binding principles have traditionally been taken to challenge direct compositionality (at least in its strong form):

- there are many cases where there are binding type effects (e.g., "binding" of pronouns, "binding" of reflexives to satisfy Principle A, Principle B violations, etc.) but where *on the surface, there is no big enough chunk of representation*

  Typical example : connectivity in specificational sentences

(1) What John is is proud of himself.
*What $John_i$ is is proud of him_i.

- usual strategy: posit silent/deleted elements and/or abstract levels which contain extra structure

(2) What John is is John is proud of himself
*What $John_i$ is is $John_i$ is proud of him_i
(Bach and Peters, 1968; Ross, 1974; Schlenker, to appear, etc.)

- once everything reanalyzed as facts about *local bits of meaning* (as direct compositionality forces us to do) - then these effects don't require positing extra structure

Case study here: "Principle B" effects

- will argue that the non-local, indexing-based, representational view is problematic in any case: suggesting that indices and representational constraints *are the wrong tools*
- will propose a (preliminary!) recasting of these effects in purely local, semantic terms
- will discuss some open questions and potential problems with the account

I. The hypothesis of Direct compositionality (e.g., Montague, 1973):

syntax and semantics work "in tandem" hence:

- syntax "builds" (i.e., proves the well-formedness of) expressions
  (where in many cases larger expressions are built from two or more smaller ones)
- semantic rules work in tandem with the syntax - to supply the meaning of the whole from the meaning of the parts
- *Meaning:* a model-theoretic object (not a representational level)

Any linguistic expression: a triple <[sound]; syntactic category; meaning>

Rules: take one or more triple as input and give one as output

Typically: binary combinatory rules (two expressions as input - phonology just concatenates them (= phrase structure rules))

Footnote: I will assume also Wrap rules - hence the input expressions are not entirely unstructured strings - they must contain an infixation point - but no additional structure necessary beyond this

Will assume also unary rules - rules mapping a single triple into another

(typically: no phonological change, some change of syntactic category, some change of meaning)

(cf., "type-shift" rules)

One consequence: no intermediate representation, e.g., LF

fairly standard competing view

- syntax first builds representations which are "sent" to the semantics
- semantics compositionally (bottom-up) interprets these structures
  - moreover, in usual view, syntax itself works in a 2 (or more) step process:
    - syntax ultimately specifies well-formed surface representations
    - which is then mapped into another level (LF)

some reasons to believe that direct compositionality is right:

(more detailed discussion in Jacobson, "The (dis)organization of the grammar: 25 years, in *Linguistics and Philosophy* ***25.6)***
which can be accessed at:
http://www.kluweronline.com/issn/0165-0157/contents

(A) Under the view where: syntax first builds representations which then are the input to the semantic rules - large parts of the syntactic information are stated twice

- once as output of some syntactic operation(s)
- once as input to semantics

Example: Take a mini-grammar where the syntax is completely simple - just phrase structure rules - which "build" representations, and where the semantics operates on the output of those representations.

Grammar Fragment 1:
syntax: S ---> NP VP
semantics: [[$[_S$ NP VP]]] = [[VP]] ([[NP]])

Compare to direct compositional vuew

Grammar Fragment 2:
S --> NP VP; [[S]] = [[VP]] ([[NP]])

Grammar fragment 1 needs to refer to the local tree $[_S$ NP VP] twice; once as output of syntax , once as input to semantics; this avoided in Grammar fragment 2

NOTE: This is using "construction-specific" rules. One might think that the problem disappears under a "type-driven" system - where the semantics associated with each bit of syntax (e.g., each local tree which inputs the semantics) is predictable.

In fact, the problem is lessened, but it is still there (the statement of the "type-driven principles" themselves will contain the duplication).

(B) Most work in formal semantics assumes that the input to the semantics is indeed *local trees.* (E.g., we are all proud of the fact that one can Curry the meaning of transitive verbs, so as to supply a meaning for a VP, rather than having to state the semantics over more global chunks of representation including subjects and objects.)

But under the view where the syntax builds representations which are then "sent" to the semantics, **this is a complete accident. There is no reason why this should be so (why not have rules interpreting large chunks of trees?)** Under direct compositionality (combined with standard assumptions about the syntax) this is not an accident. “Local trees” are simply representations of (part of the) inputs and

outputs of a syntactic rule. Compositional semantics works directly with syntax. **Necessarily, the semantics will be stated on "local trees" .**

(C) Obvious elegance

Two mistaken arguments <u>against</u> direct compositionality

A. Standard "Montague-type" grammar uses construction specific rules for the semantics, rather than general rules. Hence has lots more specific statements.

<u>Answer</u>: the question of general vs. specific rules is a separate issue. One can have direct compositionality with very general statements (and where the semantic operations are predictable from the semantic ones and/or from the semantic types). Categorial Grammar is the perfect example of this.

Note: some of the original work on "type-driven" semantics was phrased within GPSG, which had direct compositionality (Klein and Sag, 1982). "Type driven" referred to principles which predicted the semantic combinatorics from the syntactic <u>combinatory rules</u> (the generalized phrase structure rules) s combined with the semantic types of the relevant syntactic expressions.

B. Semantics should be able to interpret structures which are <u>not</u> generated by the syntax (since we can understand stuff which is syntactically ill-formed.) Therefore, we want to divorce the two systems - have generalized semantic rules which are not hooked in just to the syntactic rules. (cf, Heim and Kratzer, 1998)

<u>Answer</u>: How does the "syntax builds - semantics then interprets" fare any better here? The input to the semantics is going to have to be <u>some syntactic representation</u> - so *some system* is going to have to predict that the relevant representation is possible in some sense. In other words, to interpret ill-formed stuff, one has to allow for the system to "imagine" a way in which the syntax slipped up to produce this. But given this power of imagination - the direct compositional view works just as well. (Provided that the semantics can be predicted in some way from the syntax.)

<u>A key consequence to direct compositionality (at least in its relatively strong form)</u>:

Representations (i.,e. trees) are not something that the grammar "sees" - they just represent the proof system and the compositional semantics

(modulo operations like Wrap)

Hence: would not expect to find constraints/rules which refer to non-local chunks of representation; there is no way to state such constraints

When there appear to be such principles:

- could be the result of type-shifting constraints on things affecting co-argument slots
- or of the binary combinatorics

II. Relevance of "binding" phenomena

Standard assumptions regarding pronominal binding:

- "Binding" constraints/principles are in part constraints on co-indexation in the syntax
- hence, constraints on two actual NPs/pronouns/traces
- hence generally stated over ***non strictly local chunks of trees***
  (by "strictly local" I mean a mother and daughters)
- co-indexation itself is a syntactic notion, but will have some effect on the construction of and/or interpretation of LFs

typical example - c-command constraint on binding (for Weak Crossover effects)

- pronoun can be "bound" by another NP only if the latter (or its trace) c-commands the pronoun at some relevant syntactic level

(3) a. Every man$_i$ loves his$_i$ mother.
b. *His$_i$ mother loves every man$_i$.

notation:
letter subscripts (i, j) ndicate *readings* only (these do not indicate the actual indices used in the grammar)
arbitrarily chosen integers indicate actual indices in the grammar

(4) a. Every man$_8$ loves his$_8$ mother.
b. *His$_8$ mother loves every man$_8$.

But: how does blocking (2) actually block the reading in (1)? Not in any straightforward way

- The relationship between "binding" in the semantic sense and co-indexation/binding in the syntactic sense is very indirect and not very natural: there is no obvious sense in which *every man* "binds" *his* in the semantics (under any version of binding)

- e.,g,. a modification of account of Heim and Kratzer (1998)
  NOTE: modifications are mostly terminological; a key content modification is in (b) below, but this for expository convenience and has no affect on the central point here)

(a) each constituent at LF has as its semantic value a function from the set of assignment functions to something else
(b) each assignment function is a total function from the set of variables to model-theoretic objects (in relevant cases - from variables to individuals)
NOTE: H&K use partial functions, see below
(c) hence, if an LF constituent C contains no unbound variables within it, it will denote a constant function from the set of assignment functions

Let G be the set of assignment functions
Let G/x be a set of subsets of G, where each member of G/x agrees on the value of all variables except that assigned to x

Definition: x is free within C if, for any G/x, [[C]] differs according to the different members of G/x

Caveat: strictly speaking, this is not true - even taking intensions into account, there can be a variable which we want to define as "free" within C but where the value of C is constant regardless of what value we assign to that variable; this will be the case for Ss with the form of a tautology, for example

x walks and x doesn't walk

This is avoided under the partial function approach

Definition of Sem- Binding: (modified fairly straightforwardly from H&K):

α Sem- binds χ (a pronoun or a trace) iff the sister to α is the smallest subtree in which χ is free

But: this is not enough to get at the notion of "binding" that we want in order to enforce Weak Crossover:

(5) *His_i mother loves every man_i.

S

every man 8 S

his_8 mother VP

loves t_8

NOTE: there are questions about :

(i) whether *every man* itself should have an index (note that the index itself would play no role in the interpretation)

(ii) if so, must that index be 8 or can it be something else?

see Buring's forthcoming book for some extensive discussion

For the moment: the answer to these questions has no consequence but - we will return

the key point: The Sem- binder for *his* is not *every man* nor is it t_8 - it's actually the index 8 whose semantics is to have the effect of peforming λ-abstraction

the problem in a nutshell: under all well-understood ways to do "binding", there really is no actual relationship between *every man* and the pronoun; the idea that there is such a relationship is arguably an artefact of thinking in terms of indices instead of meanings

Definition of LF-Binding: α LF- binds β iff the trace of α and β are sem-bound by the same thing (Note: H&K refer to this as a "derivative notion of semantically bound")

Definition of syn(tactic) binding: α syn-binds β iff
α is co-indexed with β
α c-commands β
... (a few other conditions)

Weak Crossover: If α LF binds β then α must syn-bind β and vice-versa

Reasons why we might be suspicious:
- extremely complex; relationship between co-indexation and true semantic binding is very complex and indirect; constraints need to refer to LF binding (or "derivative sem-binding")

shouldn't this be telling us something? shouldn't we take quite seriously the fact that this view of "binding" is very unnatural - maybe we are on the wrong track in the first place to think in terms of co-indexing, and in terms of a relationship between the "binder" (the DP) and the pronoun - maybe we need to be recasting this

III. Moreover: lots of cases where there are binding effects but where syntactic conditions on co-indexation not obviously met. Hence:

Standard strategy: When syn-binding not hold on the(visible) surface - posit a level (or silent material) in which this does hold
- hence - cases with binding effects but no pronoun (no "bindee")
- or cases where there is a pronoun that seems "bound" but no "binder" in the relevant syntactic configuration
  - posit additional stuff/levels where the bindee or the binder is supplied

Classic example: binding connectivity in copular sentences

(6) The (only) woman that no man$_i$ would forget to invite is his$_i$ mother. (Geach, 1972)

no c-command - so "standard" strategy: posit some extra structure to get this:

(7) the only woman that no man would forget to invite is
no man$_8$ would forget to invite his$_8$ mother

Note: one version of this motivates this semantically as well as syntactically

semantic claim: these are like question/answer pairs - the pre-copular constituent is a concealed question; the post-copular constituent is an answer
(Ross, 1974; den Dikken et al; Schlenker, to appear; Romero, this workshop)

Ross/Schlenker and others:
- if so, the extra structure needed to get the connectivity effects is motivated anyway, since (it is assumed) answers themselves are elliptical for full Ss

Note though: even supposing that the question/answer theory is correct, this leads to the conclusion that the post-copular constituent has a fuller representation at some abstract level only if we also accept the premise that "short" answers are themselves elliptical for full Ss

This is often assumed (see, e.g., Morgan, 1971; Hankamer, 1974, Merchant, 2003) . But, is this correct?

(a) one can give a semantics for "short" answers without making them elliptical for full Ss, see Groenendijk and Stokhof (1984) and others

(b) a reason to believe that these are not elliptical :

(8) Q. Which man who has red hair came to the party?
A. a. John
b. John came to the party

(a) commits the answerer to the claim that *John* has red hair. (b) does not; it is a "best I can do" answer. Note:

(9) Q. Which man who has red hair came to the party?
A. a. *Well, John.
b. Well, John came to the party.

If one maintains that that "short" answers are elliptical for full Ss, one could say that (a) is elliptical for:

(10) John is the man who has red hair who came to the party.

But then, there is no obvious formal "identity" condition between the question (antecedent) and the answer, and so it's not clear what it buys us to posit ellipsis (how would we explain connectivity effects in terms of syntactic conditions?)

Other strategy (no abstract representations; hence easier to implement under direct compositionality): : assume that the relevant effects here (and elsewhere) are semantic and can be stated in purely local terms (Jacobson, 1994; Sharvit, 1999)

Jacobson, 1994; Sharvit, 1999 (see also von Stechow, 1991)

the only function f (with range woman') such that no man is an x who would forget to invite f(x) is the-mother-of function

Jacobson, 1994: framed within variable-free semantics, in which almost all parts of this come "for free" from variable-free apparatus

IV. The proposed revised tools for binding: variable-free semantics (see, e.g., Jacobson 1992, 1999, 2000)

NOTE: I will probably go over this just very quickly

- semantics makes no use of variables or assignment functions
- no indexing conventions in syntax
- no globally stated constraints on representations (no constraints on "co-indexing", no constraints on the relationship between "binders" and "bindees"; in fact, no notion in the grammar of "binders" and "bindees")

(11) a. Every man$_i$ thinks that he$_i$ lost
b. Every man$_i$ loves his$_i$ mother.

A. <u>The meaning of expressions containing pronouns unbound within those expressions</u>

standard view:

[[he-lost]] : a (non-constant) function from assignment functions to propositions
[[his-mother]] : a (non-constant) function from assignment functions to individuals
[[he]] : a (non-constant) function from assignment functions to individuals
$[[he]](g) = g(x_i)$

variable free view:

any expression containing a pronoun "unbound" within that expression (including a bare pronoun itself) denotes a function from <u>individuals</u> to something else

[[he lost]] : a function from individuals to propositions
in particular, it is the function [[lost]] (modulo gender information)
[[his mother]] : a function from individuals to individuals
in particular: the-mother-of function
[[he]] : a function from individuals to individuals
in particular, it is the identity function
(actually, a partial function defined only for the domain of male individuls; ignore gender here)

Question:: how get the result that [[he lost]] = [[lost]]?

more generally, how allow *lost* to take a pronoun as subject?

[[lost]] of type <e,t>

[[he] of type <e,e>

Geach rule: any function of type f <a,b> can shift to one of type <<c,a>,<c.b>

call this **g** (f)

where **g** (f) = λ V $_{\text{of type } <c,a>}$ [$\lambda x_{\text{of type c}}$ [f (V(c))]

NOTE: this is a unary (Curry'ed) version of function composition

thus for any two functions f and h

g(f)(h) = f o h

lost; S_L/NP; lost' -->$_{\mathbf{g}}$ lost; $S^{NP}/_R NP^{NP}$; $\lambda f_{<e,e>}[\lambda x[$ lost' (f(x))]]

he; NP^{NP}; $\lambda y[y]$

he lost; S^{NP}; $\lambda f_{<e,e>}[\lambda x[$ lost' (f(x))]]$(\lambda y[y])$ = lost'

B. Binding:

one standard view:

binder "pulled out"

x_i thinks x_i lost ---> λx_i [thinks'(lost'(x))(x)]

and this taken as argument of the generalized quantifier every-man'

variable-free view

"Binding" is the consequence of a type-shift rule **z** applying to *thinks*

thinks' <t,<e,t>> - i.e. a relation between individuals (the subject) and propositions

z(thinks') - <e,t>,<e,t>> - i.e., a relation between individuals and properties, such that **z**(think) (P) is to be an x who thinks that P(x)

hence: thinks-he-lost' - is the set of individuals who think that self lost

this then taken as argument of the generalized quantifier every-man'

z also has a syntactic side; this can be ignored here

(12) Every man$_i$ loves his$_i$ mother

standard analysis: every-man'(λx[loves'(the-mother-of'(x))(x)])

Variable free: his-mother' - a function from individuals to individuals

(the-mother-of function)

z(loves;) is a relation between individuals and functions of type <e,e> such that to **z**(love) f is to be an x who loves f(x)

hence: loves-his-mother - is the set of individuals who love their own mothers

this is then taken as argument of the generalized quantifier every-man'

Weak Crossover: a consequence of the combinatorics: rule **z** which merges a slot waiting for a pronoun to a higher argument slot; no reverse rule **s**

• Binding is a relationship between *argument slots* - not between two actual NPs/pronouns/traces in an actual representation
• Since it is between argument slots , it is completely local

Binding connectivity in copular sentences:

(13) a. The (only) woman that no man would forget to invite is his mother.
b. The (only) woman that no man invited was his mother.

Jacobson, 1994; see also Sharvit, 1997, 1999:
[[no man invited]] in ordinary case: λx[no-man'(invite'(x))]
here: λf[no-man'(**z**-invite'(f))]]

cf: functional questions and other kinds of functional gaps
Note: functional readings, and the existence of functional gaps comes "for free"; hence functional questions and functional relatives comes (almost) for free and is perfectly expected
Functional readings are just an automatic consequence of the mechanisms for binding in general
No need for complex meaning in a functional gap, or complex indexing (no need for any indexing)

[[woman]] - set of functions with range women
[[woman who know man invited]] - intersection of these two
[[the (only) woman who no man invited]] - the unique function f with range woman and which is such that no man **z** -invited f
[[his mother]] - automatically denotes a function from individuals to individuals (since it contains a pronoun "unbound" within it) - the-mother-of function

A consequence: "c-command" condition between "binder" and pronoun is an illusion (in fact, no real notion of a "binder"). To the extent that there is anything analogous here, the "c-command" condition is a relation between argument *slots* - the **z**-rule "merges" a higher argument slot with a slot that gets occupied with material that contains a pronoun. Hence all is local; no need for any fancy account of connectivity cases.

A note on free pronouns: (14) He lost.

standard view: this is a non-constant function from assignment functions to propositions
• to compute propositional information, this applied to some contextually salient assignment function
variable-free view: this is a function from individuals to propositions
• to compute propostionaly information, this applied to some contextually salient individual

UNANSWERED QUESTIONS:

a. The challenge for direct compositionality: can all non-local constraints on chunks of representation be eliminated (in a simple and natural way)?
b. The challenge for variable-free semantics: can all constraints on co-indexing be stated in some other way (and in a simple and natural way)?
c. The challenge for the purely semantic view of connectivity: can all connectivity effects be accounted for without abstract levels/silent stuff?

V. One persistent thorn: Principle B effects

(15) a. *Bush$_i$ praises him$_i$ (every chance he gets).
b. *Every candidate praises him$_i$

a. generally stated as a constraint on a pronoun and another NP within the same clause (not purely local)
b. generally stated (at least in part) as a constraint on co-indexing

shows typical connectivity effects:

(16) a. *What Bush$_i$ does (at every news conference) is praise him$_i$.
b. *What every candidate$_i$ does is praise him$_i$.

c. how can (16) be ruled out without a co-indexed subject sitting in the post-copular constituent?

VI. Interlude: Claim: there is no satisfactory account of Principle B effects using the usual tools of co-indexation (hence, we should rethink the standard accounts)

(17) Principle B: (standard kind of way to state it): no co-indexation of a pronoun and another NP within a certain domain (e.g., same clause)

Assume:
- α LF-binds β if trace of α and β are sem-bound by the same thing
- if α LF- binds β then α must synt-bind β

(hence the two must be co-indexed, but (17) blocks this)

ultimately: blocks the "bound" reading for (15a) and (15b)

Problem 1:

Well known: this does not block the "accidental" coreferential reading for (15a) (Reinhart, 1983)
(Let them be co-indexed, but where the relevant assignment function assigns the same individual to the two different indices)

Interlude on the interlude: Are we sure about this? Could we try to actually get both cases by a single constraint on co-indexation - where we assume:

• for each assignment function, there is a one-to-one correspondence between the variable names and the individuals. In other words, for any assignment function g, if $m \neq n$ then $g(m) \neq g(n)$.

(18) $Bush_i$ loves his_i mother.

We still want two representations for this (among other things, to allow for both strict and sloppy readings in *Bush loves his mother, and Dick does too.)*

This remains possible: both involve co-indexation of *Bush* with *his*. But sloppy reading is derived by QR'ing *Bush*, while strict reading involves no QR.

Hence, sloppy reading has access to a VP (or something else) with meaning
λ8 [8 loves 8's mother]

while for the strict reading, the meaning of the VP is just loves 8's mother .

One reason this won't work: this is not a tenable constraint on assignment functions. It gives rise to a class of problems, which can be exemplified by the following:

(19) $Bill_i$ thinks that no one at the moment is living in his_i house.

Assume for the sake of illustration that *no one* has to undergo QR here. Suppose that *Bill* and *no one* have the same index. Then we don't get the reading we want, since *no one* will end up binding *his*.

Assume that *Bill* and *no one* do not have the same index. Then this will end up meaning that Bill thinks that no one other than himself is living in his house.

> Let *no one* have the index 6. Then, on assignment function g, the argument of [[no one]] is the function which maps each individual n into the proposition
> g'(6) lives in Bill's house
> where g' is the assignment function just like g except that n is assigned to 6
>
> However, there is no g' just like g except where Bill is assigned to 6. Hence, the argument of *no one* has to be a partial function, undefined for Bill.

But (19) lacks this reading. Sometimes *no one* can be understood as "excepting" other NPs mentioned in the sentence or discourse. But in this case (19) has no interpretation in which *Bill* is excepted.

NOTE: My proposal also ends up with a specialized version of this type of problem. But as far as I can tell, in the cases where this problem emerges under my proposal

the facts are actually murky (I could probably be convinced either way). In the case here, though, the "except for Bill" prediction seems quite pernicious.

An attempted way out: don't assume *no one* is in the QR'ed position. (Then its index won't matter for anything; it can be co-indexed with *Bill* but still won't get to mistakenly bind the pronoun *his*.)

We fix this by fancying up the example; add in another pronoun which must be bound by *no one*:

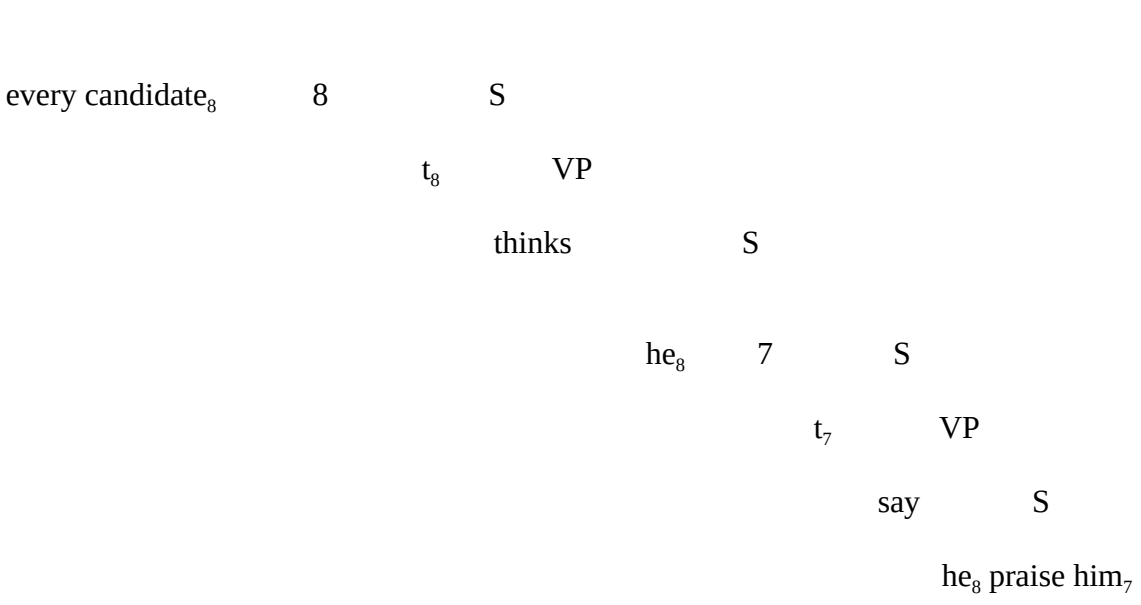
(20) $Bill_i$ thinks that no man_j could possibly live in his_i apartment unless he_j can stand cockroaches.
(no reading which excepts Bill)

bottom line: in addition to a constraint on co-indexation, we need some way to block the accidental coreference reading, where there is no co-indexation

Problem 2: (this type of problem discussed in, e.g., Heim, 1993, Reinhart, 1997; Fox, 2000; Buring, to appear; etc.): Need something extra to block relevant reading in (21) which can be represented as in (22) (the two pronouns here not co-indexed) (call this the "fancy binding" case)

(21) *Every $candidate_i$ thinks that he_i should say that he_i will praise him_i.

(22)



Interlude again: are we sure that this is a problem? Are we sure that a structure like this should exist, and that it can be a representation for (21)? (thanks to Klaus Abels for bringing this kind of question to my notice)

The answer will be: yes, we need to allow for representations of this general type and so we'll need an extra principle to rule out this particular one. But showing this takes a bit of work.

What's at issue is whether the index on the binder must always match the index which performs the semantics of λ-abstraction. In other words, is the following tree bit well-formed:

he_8 [7 S]

Two standard arguments that representations of this general type should be allowed (see, e.g., Fox, 2000):

A. They come "for free". But: this isn't really true. As an LF, it would require some extra work to disallow these. But one could imagine that - although this is a well-formed LF - it can't be derived from the relevant surface structure. How?

• assume everything - all full NPs, quantified NPs, and pronouns have an index, and that this index is preserved under QR. Then

he_8 [7 t_z say he_8 praise him_7]

isn't possible; I've underlined the offending portion.

It is true that the indices on full NPs and quantified NPs do no work, but it's easy enough to ensure that they match the index on the "binder". (For relevant discussion, see Buring, to appear)

B. We need to allow for this general type of structure to get one of the readings for (23):

(23) Every man_i only thinks that HE_i should give a present to his_i mother.

The relevant reading is:

(24) every man only thinks that HE (bound) should give a present to his (bound, but non-sloppy) mother

i.e., this reading can be represented as:

(25) every man_8 [8 [t_8 thinks HE_8 [7 [t_7 should give a present to his_8mother]]]]

This makes the alternatives within the focussed domain be things like:

he_6 [7 [t_7 should give a present to his_8mother]]

7 is bound all the way through, so the alternatives are about present-givers (as we want), but the mother stays constant (as we want)

However: we don't really need to allow for this kind of representation to get this reading. We can also get it simply by not applying QR to HE, so tht the relevant LF would be:

(26) every man$_8$ [8 [t$_8$ thinks HE$_8$ should give a present to his$_8$ mother]]

since HE$_8$ is what is focussed, the relevant alternatives are he$_7$, he$_9$ etc.

and so we get alternatives like
he$_9$ should give a present to his$_8$ mother
which is what we want.

Nonetheless: we can construct fancier versions of this case where the way out in (26) is unavailable. The strategy: make sure that HE must be QR'ed by having it bind a different pronoun, one which is interpreted sloppily

(27) Scenario: There was a bad accident involving a bus full of elderly women from a nursing home. Several were injured, and brought to the hospital. Of those, many had sons who came to the hospital to see them - and each one had several other relatives who also came. As is typical in hospitals, all the sons and the relatives were waiting around for hours - not being able to see their mother/relative and not even being told when they would be able to see them.

Each man$_i$ hoped that at least HE$_i$ would soon be told when he'd$_i$ be allowed to see his$_i$ mother.

The relevant (and very salient) reading:

Each man hoped that at least HE (bound) would soon be told when he'd (bound, sloppy) be allowed to see his (bound, strict) mother

Hence, the alternatives in each man's mind concern other relatives being told when they might be able to see that man's mother.

Hence: to get this reading, we have to allow for the indices on a pronoun (or on a binder in general) to not be the same as the index on the pronouns it binds.

The consequence: we can get "fancy binding" where two pronouns are ultimately bound by the same thing (via a chain of binding) even though they aren't co-indexed.

the bottom line: Principle B effects in (21) can't be ruled out just by a constraint on co-indexing.

• Attempted solution 1 - the Gricean account: (Dowty, 1980; Reinhart, 1983; Sadock, 1983):

The effect is purely Gricean, and there is no Principle B in the grammar. The violating configurations are just those in which a reflexive could have been used. Since a pronoun is ambiguous and a reflexive isn't, use of the reflexive is more informative. Hence, by Gricean priniples, if a speaker did not use a reflexive, the listener will assume that non-coreference (or, non-binding) was intended.

Problem:

(a) reflexives and pronouns are not actually in complementary distribution, as this predicts

(28) John very much liked the picture of himself$_i$/him$_i$ that appeared in the yearbook.

(b) **This resists any kind of augmentation, cancelling, etc. which is always allowed in the case of truly Gricean effects.**

Note: the normal scalar implicature cancellation tests are probably irrelevant here, since this is not a case where the *meaning* of *him* is "less informative" - rather, it's a case of ambiguity

(29) a. Barbara praised him - that is to say, George.
b. Barbara praised him - that is to say, she praised George.

(30) a. I went to the bank, that is to say, the river bank
b. I went to the bank, that is to say, I went to the river bank.
(31) a. *Bush praised him, that is to say - himself.
b. *Bush praised him, that is to say - he praised himself.

Note though: if a Gricean solution does turn out to be correct, then there is no problem for direct compositionality

Attempted Solution 2; two principles, one "transderivational" (Grodzinsky and Reinhart, 1993; Reinhart, 1998; Fox, 2000; Buring, to appear, etc.)

There are two separate constraints here. One is "core Principle B" (e.g., as given in e.g. (17)). The other is a second - transderivational - principle which blocks the accidental coreference reading *in virtue of the fact that - were it not for principle B - the bound reading would have resulted in the same interpretation.* Similarly, we will need the constraint to rule out the "fancy binding" possibility for (22) and allow in only the representation in which the two pronouns are co-indexed. (Hence "core Principle B" will rule out that representation.)

various versions; we'll consider the version in Heim, 1993(?), Fox, 2000, Buring (to appear)

Rule H (this essentially the version in Fox, 2000 and Buring, to appear):

(32) For any two NPs α and β, if α could bind (i.e., if it c-commands β and β is not bound in α's c-command domain already), α must bind β, **unless that changes the interpretation** (use of boldfaced mine, PJ)

Problems with this type of strategy:

(a) "could bind" - *has to overlook "Core Principle B" violations* - that is, Principle B says that one Bush *can't bind* him in (15a). (32) rules out non-coindexation (non-binding) in virtue of the fact that co-indexation (binding) would have been possible and would have yielded the same interpretation. But actually, co-indexation would not have been possible. So we have to rule out a bad case, even though the competing analysis itself is bad.

Put differently: the non-bound interpretation of (15a) is ruled out because there is a competing binding possibility *which gives rise to the same interpretation.* Hence, the competing bound representation has to be able to be interpreted (otherwise, there would be no way to determine that it gives the same "interpretation"). So the Principle B violation does not "throw out" the representation; it is completely interpretable.

Competition effects where one representation/interpretation pair is blocked on the basis of a competing bad representation/interpretation pair are unprecedented.

(b) What is (32) a constraint on??? It appears in the end that it is actually going to need to be stated in two separate places.

• accidental coreference blockage: this has to be a constraint on the processor (and not on the grammar/compositional semantics)

• have seen: this is not a constraint on co-indexing; the two NPs need not be co-indexed

• free pronouns (in standard theory): the meaning of the S is a non-constant function from assignment functions to propositions; the listener applies this to some contextually salient assignment function

• hence: (32) is a direction to the listener/processor (not the grammar). Take a representation of (15) in which the pronoun is not bound. Pick an assignment function. Check the interpretation. Compute a competing LF (with more local binding). Check its interpretation. See if they are the same. If so, throw out the interpretation you computed.

• But take the case of (22). In this case (32) operates as a constraint *on the grammar (compositional semantics).* It says to the grammar: look at the interpretation of a structure in which there is a pronoun. Compute a competing

representation where the pronoun is (more) locally bound. If the interpretation is the same, throw out the representation with the less local binding.

Hence it looks like (32) will really have to be two separate principles. If so, we would have three constraints to account for Principle B effects.

(c) transderivational principles like this: extremely complex - one form/meaning (or, form/interpretation) pair blocked in virtue of existence of another form/meaning pair (consider the complexity of actually formalizing this)

NOTE: Gricean principles are also of this nature - so arguably competition classes are computed. These, however, seem to be principles adhered to by speakers and listeners, not principles within the grammar.

VII. The (tentative) proposal under direct compositionality and variable-free semantics

Categorial Grammar syntax:

- syntactic categories fully encode distribution
  let $A/_R B$ mean an expression which takes a sister of category B to the right to yield a mother of category A
  $A/_L B$ takes a B to its left to give a mother of category A
  - hence verbs specify the syntactic category of the object that they will take
- the meaning of any expression of category A/B is a function from B type meanings to A type meanings (ignoring intensions)
- hence, when an A/B combines with a B to give an A, the associated semantics is functional application
- *in addition:* let an expression with the general distribution of A-type expressions but which contains within it an unbound pronoun of category have as its category A^C
- crucially: this means that a pronoun (e.g., *he)* is of category NP^{NP}
  so is an expression like *his mother* (which contains a pronoun unbound within it)
- semantic shifts: **g** and **z** are accompanied by syntactic category change rules
  g: input expression of category B/A shifts to B^C/A^C
  this just says that when the function takes as argument an expression which has an unbound pronoun in it, it passes up the information that there is an (as-yet) unbound pronoun
  cf. feature passing conventions of G/HPSG

  z: input expression of category (A/NP)/B shifts to $(A/NP)/B^{NP}$

(so, both **g** and **z** give an output expression which says: give me an argument which is or contains a pronoun)

Principle B idea; begin with a minor revision on the syntax

- above: assumed that an NP which contains a pronoun unbound within it and a pronoun itself are both of the category NP^{NP}

- Refine this: a pronoun is of category $NP[p]^{NP}$
  -- NP[p] ≠ NP
  -- NP[p] cannot fill ordinary NP slots; occurs only where explicitly allowed

the informal idea:

praise: listed in lexicon as a 2-place verb - hence of category (S/NP)/NP with meaning of type <e,<e,t>> and where [[praise]] characterizes the set of all order pairs of praiser-praisees (including reflexive pairs)

However: this verb cannot directly take a pronoun in object position *because a pronoun has a special category, and cannot occupy ordinary NP argument slots*

technical detail: the verb in any case couldn't take a pronoun - a pronoun is of category $NP[p]^{NP}$ and has as its meaning a function of type <e,e> , not an individual

however: *his mother* is also of category NP^{NP} and has a meaning of type <e,e>

how let material *containing a pronoun* combine?

- by **z** rule - which will bind the pronoun slot to the subject position

or, by **g** rule - which will allow the pronoun slot to remain free and be bound from higher up

z(praise); $(S/NP)/NP^{NP}$; $\lambda f[\lambda x[praise'(f(x))(x)]]$ (wants a function f of type <e,e>, and returns the set of individuals x which stand in the [[praise]]-relation to f(x)

g(praise); $(S/NP)^{NP}/NP^{NP}$; $\lambda f[\lambda x[praise'(f(x))]]$ =

$\lambda f[\lambda x[\lambda y[$ y praises f(x)$]]]$
(wants a function f of type <e,e> and returns the 2-place relation holding between y's and x's such that y praises f(x)

Thus: bare pronouns and expressions containing pronouns unbound within that expression are not of <u>exactly</u> the same category; the former of category $NP[p]^{NP}$ and so can

occur only where explicitly allowed. The verbs that do allow this category in object position have meanings which are undefined for reflexive pairs.

what allows bare pronouns in object position? - all transitive verbs map to a more fully augmented form - which : syntactically, explicitly allows pronouns in object position, and semantically: *is defined only for the non-reflexive pairs in the original relation*

praise; (S/NP)/NP; praise' ---> $praise_{irr}$; (S/NP)/NP[p]; this is a function of type <e,<e,t>> such that $praise_{irr}$:(x)(y) = 1 if praise' (x)(y) = 1 and x≠ y;
$praise_{irr}$' (x)(y) = 0 if praise'(x)(y) = 0 and x ≠ y; and
$praise_{irr}$' (x)(y) is undefined if x = y

will notate as: $[[praise_{irr}]] = \lambda x[\lambda y_{(y \neq x)} [praise' (x)(y)]$

how this accounts for all of the Principle B effects:

A. The bound case:

(32) *Every $candidate_i$ praises him_i.

him is trying to be bound; this would have to involve **z** on praise:

praise; (S/NP)/NP; praise' ---> $praise_{irr}$; (S/NP)/NP[p]; $\lambda x[\lambda y_{(y \neq x)} [praise' (x)(y)]$

this still can't combine with a pronoun (since a pronoun is $NP[p]^{NP}$ so it has to undergo an additional operation; either **z** or **g**

if **z**:
---> $z(praise_{irr})$; $(S/NP)/NP[p]^{NP}$; $\lambda f[\lambda z[\lambda x[\lambda y_{(y \neq x\_} [praise' (x)(y)]] (f(z))(z)]] =$
$\lambda f[\lambda z_{(z \neq f(z))} [praise'(f(z))(z)]]$

ordinary meaning of pronoun: identity function over individuals; hence if this applies to the identity function over individuals, we will get back the set of z's who self-praise and are not equal to themselves

z-$praise_{irr}$ him; S/NP; $\lambda z_{(z \neq z)} [praise'(z)(z)]$

NOTE: Jacobson (2000) argues that paycheck pronouns are derived from ordinary pronouns via the Geach rule; hence a paycheck pronoun is an identity function over functions. The above does allow for binding 'into" paycheck pronouns, as in:

(33) Every Democratic $candidate_i$ tried to distance himself from his_i father; while every Republican $candidate_i$ copied $him_{f(i)}$

B. The free ("accidental coreference") case:

(34) *Bush$_i$ praises him$_i$.

if **g**:
praise; (S/NP)/NP; praise' ---> praise$_{irr}$; (S/NP)/NP[p]; $\lambda x[\lambda y_{(y \neq x)}[$ praise' $(x)(y)]$
--> g(praise$_{irr}$); $(S/NP)^{NP}/NP[p]^{NP}$; $\lambda f[\lambda z[\lambda x[\lambda y_{(y \neq x)}$ [praise' $(x)(y)]]$ $(f(z))]]$ =
$\lambda f[\lambda z[\lambda y_{(y \neq f(z))}$ [praise'$(f(z))(y)]$

let this apply to a pronoun, whose meaning is the identity function, and we get as meaning for

g(praise$_{irr}$)him = $\lambda z[\lambda y_{(y \neq z)}$ [praise'$(z)(y)]]$

note: this is just the same as the meaning of praise$_{irr}$; it's just the 2-place praise relation holding between pairs, minus the reflexive pairs

(35) Bush praises him; $\lambda z_{(z \neq b)}$[praise'$(z)(b)$]
details require type-lifting the subject and "geaching" on it (see Jacobson, 1999 for full details of this in general):

so: will pick out the set of individuals that b praised and who are distinct from b

Advantages (aside from the fact that this is purely direct compositional):

- no need for two (or three) separate principles
- no "transderivational" mechanisms
- all done locally, with ordinary model-theoretic objects

Open question:

well-known that the "bound case" is much worse than the "accidental coreference" case (and possibly acquisition evidence which suggests these should not be completely collapsed)

Is that evidence that there *should* be two separate principles? Can the "unified" account here explain the difference>

Tentative speculation:

here the bound case is nonsense entirely within the meaning of the VP. I.e., the VP denotes a function which is undefined for everyone. The other case results in an undefined situation only when the function which is the meaning of the entire S is applied to the contextually salient individual b. Could this be the difference?

C. The "fancy" case:

(36) *Every candidate$_i$ thinks that he$_i$ said that he$_i$ praised him$_i$.

Note: there isn't an exact analogue to (22), because in the variable-free account there is nothing analogous to a case of two pronouns "being co-indexed" (i.e., corresponding to the same variable).
When two pronouns within an expression C are bound by the same thing, this is the result of two applications of **z** on tyhe function which takes C as argument.

closest analogue to (22):

he praise him: = praise$_{irr}$' $=\lambda x[\lambda y_{(y \neq x)}[$ praise' $(x)(y)]]$

said: subject slot binds one of these; other slot remains free and is passed up for higher binding; this happens by application of both **z** and **g** on say

z(g(say)) = $\lambda R[\lambda w[\lambda z[z$ say $R(w)(z)]]$

said he praised him: = **z**(**g**(say)) (he-praise-him') =
$\lambda R[\lambda w[\lambda z[z$ say $R(w)(z)]]$ $(\lambda x[\lambda y_{(y \neq x)}[$ praise' $(x)(y)]])$ =
$\lambda w[\lambda z[z$ say $\lambda x[\lambda y_{(y \neq x)}[$ praise' $(x)(y)]]$ $(w)(z)]$ =
$\lambda w[\lambda z[z$ say $\lambda y_{(y \neq w)}[$ praise' $(w)(y)])(z)]$

note: some possible question here about just how to "project" the presupposition

he said that he praised him : same as above

the "free" pronouns correspond to the slot which is the subject of *say* and the subject of *praise* - both of these slots are bound by the subject slot of *think,* and this happens via two applications of **z** on *think:*

z(z(think)) = $\lambda R[\lambda x[x$ thinks $R(x)(x)]]$

$\lambda R[\lambda x[x$ thinks $R(x)(x)]](\lambda w[\lambda z[z$ say $\lambda y_{(y \neq w)}[$ praise' $(w)(y)])(z)])$ =
$\lambda x[x$ thinks $\lambda w[\lambda z[z$ say $\lambda y_{(y \neq w)}[$ praise' $(w)(y)])(z)]$ $(x)(x)]$: =
$\lambda x[x$ thinks x say $\lambda y_{(y \neq x)}[$ praise' $(x)(y)])(x)]]$
i.e., the set of x's who said x praised x for $x \neq x$

every-man'$(\lambda x[x$ thinks that x said that x (for $x \neq x$) praised x])

Caveat: there are other ways that one could get attempt to get the relevant reading; to show that this is "nonsense" in all cases one needs to show this for all derivations. Hence, a promissory note, but the semantics should come out the same however one tries it.

Prediction: to the extent that we can tease apart the "bad in the semantics" cases from the "bad in the supplying of values to 'free pronouns'" cases; we predict this to be as bad as the bound cases - this is bad "in the semantics"; this seems correct

Conclusions:

- have a single principle for all of these cases, although where things go wrong with the "free" case and the bound case are in "different places" (one in the semantics, one in the application of the function to a contextually salient individual - i.e., one in the grammar, the other via the listener)

- No need for a global constraint on co-indexation; this is all done via direct compositionality (and without variables)

- Might be able to mimic this account via constraints on assignment functions, but then not clear that the assignment functions doing any actual work

- Arguably, better than a constraint on co-indexation, since that doesn't work anyway

- N o need for extra stuff in connectivity cases:

(26) *The only thing that everyone$_i$ did was praise him$_i$.

(badness is all bound up within the meaning of the VP)

VIII. Why tentative?

A. Interaction with focus

(37) ?No one praised Bush today. Cheney didn't praise him, Rumsfeld didn't praise him, even Bush himself didn't praise him.

(38) ?Almost no one praised Bush today. Cheney won't praise him, Rumsfeld won't praise him, only Bush himself will praise him.

B. Needs to be generalized for the case of 3-place verbs

C. Syntax needs to be worked out: distribution of [p] feature - why can this always occur in subject position?

D. A version of the "exception" problem noted above emerges here

(39) Every man believes that no man praised him.
(compatible with the situation of each man believing that he himself praised himself - it';s hard to know if that's a good prediction or a bad one)

here: the problem surfaces only when there are bare pronouns in object position - in all of these cases it is not so clear that there isn't a reading in which the subject is excepted to the extent that the subject is excepted, this could just be a fact about the pragmatics; but it *might* also be a fact about the semantics - hard to tell

IX. Additional inadequacies that a defender of the co-indexation representational account might point to:

A. Rule H above, or something like it, is designed (at least in Fox) to also account for Dahl's puzzle

(40) George$_i$ said he$_i$ voted for his$_i$ father, and Al did too.
readings: sloppy throughout
strict throughout
sloppy - strict (Al said Al voted for George's father)
no strict - sloppy (*Al said George voted for Al's father)

this accounted for by Rule H (see Fox); I have no account of this.

B. Various versions of Principle B and/or its supplements are also designed to account for (at least some) Principle C effects. I have no account of these. I assume that Principle C has nothing to do with the semantics (or even "grammar' strictly speaking), but that these (mild and variable) effects arise from principles of discourse, etc. - a la proposals by Kuno and others.