Just Pair-merge: Are adjuncts syntactically defined by operations or representations?

Ken Safir, Rutgers University

In current minimalist theory, there are two structure-building operations, simple merge, which forms sets, and pair-merge, which forms ordered pairs. Chomsky (2004) proposes that syntactic adjuncts are so-defined by the fact that they are always introduced by pair-merge.

However, it is intuitively appealing and consistent with minimalist aspirations to see if linguistic theory can do with just one rather than two merge operations. I will posit that only pair-merge is necessary to generate the full range of syntactic configurations that grammar requires.

A direct consequence of this approach is that all relations are introduced by the adjunction operation, pair-merge, so there is no distinction in the operation that produces syntactic objects that distinguishes complements and adjuncts. Insofar as a binary distinction between adjuncts and arguments is useful, it must be defined on either on syntactic representations or semantic distinctions or both.

Today's message:

1A A theory with just pair-merge and no simple (set) merge is a simpler theory more consistent with minimalist assumptions.

1B A theory with just pair-merge and no simple merge is just as adequate to generate all relevant syntactic relations.

1C Syntactic relations, including the adjunct/argument distinction, is stated on representations and is not a consequence of how a structure-building operation applies.

1D Apparent advantages of the two merge operations theory, including the labeling algorithm, are illusory. **(The existence of agreement is not derived, the EPP is not derived).**

1E Only Just-Pair-merge theory derives the inherent asymmetry of c-command.

The claim that all merge is pair-merge has been proposed by a number of others before, most notably, Kayne (1994), though I crucially do not adopt the claim that c-command translates into any commitment as to linear ordering of the pair members. If there is time, I can talk about other "just pair-merge" proposals in the literature, especially Zwart (…) and how they differ from mine.

1. What is simpler?

A) Two Merge Operations Theory (TMOT) - Chomsky's 2013, 2015 System

Simple Merge forms sets, e.g., X and Y merged form {X, Y} with no linear order specified.

Pair-merge form ordered pairs, e.g., X and Y merged form <X, Y> with no linear order specified.
  The asymmetry of pair-merge is evident in the labeling of the output:
  Pair-merge Labeling: If X is adjoined to Y, then $<X, Y>_Y$
However, Chomsky (2004) assumes that adjuncts introduced by pair-merge are invisible to further syntactic operations. They are made visible in the semantic component by **SIMPL** applying after syntactic derivation.

Labeling Algorithm (LA)
  If X is a head and Y is a non-terminal, then $\{X, Y\}_X$
  If both X and Y are non-terminals, then they only project a label if they "agree"
  Lower copies are invisible to LA
  Adjuncts are invisible to LA
  Labeling Filter: Every constituent must have a label

Because adjuncts are invisible to LA, ***Pair-merge labeling is derived***.

B) Just Pair Merge Theory (JPMT) (what I am advocating for today)

Pair-merge forms ordered pairs, e.g., X and Y merged form <X, Y> with no linear order specified.
  The asymmetry of pair-merge is evident in the labeling of the output:
  Pair-merge Labeling: If X is adjoined to Y, then $<X, Y>_Y$
  Labeling Filter: Every constituent must have a label.
  Adjuncts are visible to subsequent syntactic operations.

 C) Primitives of TMOT Not in JPMT
  Lower copies do not have to be stipulated to be invisible
  There is no need for **SIMPL**
  Simple Merge does not exist
  LA does not exist

D) Primitives of JPMT not in TMOT
  Pair-merge labeling
  Adjuncts, like all constituents, are visible to syntactic operations.

Given (C) and (D), neither theory is a subset of the other, unless it can be shown that adjuncts must be visible to syntactic operations. If it can be shown that <u>TMOT requires Pair-merge Labeling</u> in addition to the other assumptions about LA, ***then JPMT principles are a subset of those of TMOT and JPMT is simpler.***

1.1 Evidence that adjuncts are visible to syntactic operations

There is ample evidence, however, that at least some adjoined constituents are indeed visible to syntax in that their contents are accessible for internal merge to adjunct external positions. For example, as Truswell (2007) and Oseki (2013) show, the restrictions on extraction from adjuncts is not as absolute as Chomsky's different-plane theory would predict.[1] Consider the following sorts of cases of extraction by questions, relatives and topicalizations.

 1a) Which tax bills did John leave town without paying?
  b) Which tax bills did John leave town before taking care of?
  c) Which of his relatives did she cross the room (in order) to avoid?
  d) Those are just a few of the responsibilities he left town without/before taking care of
  e) Those responsibilities I am sure he will leave town without/before taking care of.
  f) I know that guy who she crossed the room (in order) to avoid.
If adjuncts are not susceptible to further operations after they are pair-merged to structure, then (1a-f) should be altogether impossible.

Cinque (1990) argued that extractions of the sort in (1) might be cases where there is A'-licensing of a pro inside the adjunct in order to capture asymmetries between NP and PP extraction. I don't have an answer for blocked PP extraction, however, it is generally assumed that pro does not allow reconstruction.

2a) Which of his$_i$ responsibilities did every congressman$_i$ leave town without/before taking care of?
  b) How many of his$_i$ incompletes can any student$_i$ graduate from here without taking care of?
  c) Which of his creditors would any gambler skip town (in order) to avoid
Reconstruction shows that at least these cases must involve internal merge leaving a copy, in current minimalist terms.

Biskup and Šimuk (forthcoming?) catalogue a number of papers showing that there is extraction from some adjuncts, including from tensed adjuncts in Czech – we will hear more about this tomorrow.

Similarly, if multiple wh-questions can access the position of *what* in (3a), which not everyone accepts, as well as the discourse-linked cases in (3b), which are far more acceptable, then once again, adjuncts are not invisible to further operations. This is especially clear for theories, like that of Bobaljik (2002), where wh-in-situ is a case of a lower copy being pronounced in what is otherwise a case of internal merge in syntax.

3a) Who left town without paying what?

---

[1] See Truswell (2007) argues that adjunct extraction is successful depending on how the adjunct is related to the matrix event, but such an account is to no avail if pair-merge simply renders extraction from adjuncts underivable. See Oseki (2013) for an argument that only 'adjuncts' introduced by simple merge allow extraction and he analyzes examples like (2a) as such a case. His account does not extend to (2d).

b) Which of the congressmen left town before talking to which portion of their constituents?

c) Who crossed the room (in order) to avoid who?

In other words, the absolute restriction on access to adjuncts on the grounds that they are invisible to further syntactic operations is too strong to account for the relative accessibility of adjuncts and is thus a poor account of that restriction.

If adjuncts are visible to syntax, then there is no reason to say that they are not visible to Chomsky's LA. The purported invisibility of adjuncts to just labeling theory, not merge, is then just an additional stipulation, equivalent to Pair-merge Labeling, to insure that the label projected after adjunction is that of the adjoined-to constituent.

- This makes the adjunct assumption like the trace assumption: Lower copies and adjuncts are visible to syntax but just invisible to labeling.
- Only TMOT requires these enabling assumptions.

Thus the only mismatched assumptions between TMOT and JPMT are now removed and the assumptions of JPMT are now a subset of the assumptions required by TMOT.

UPSHOT: JPMT IS A SIMPLER THEORY.

It is also worth pointing out that Chomsky's theory of wh-extraction proceeding by movement to intermediate nodes at the edge of phases (cycles) is not formulable as adjunction if his approach to pair-merge renders intermediate movements invisible to further operations.

Thus the empirical evidence that is supposed to motivate the invisibility of adjuncts in fact gives evidence for the opposite conclusion: Adjoined constituents must be visible to syntax. Once it is established that adjuncts must be visible to syntax, Chomsky's LA is not sufficient to provide the right labels when pair-merge applies to adjoin wh-XP to vP, since no label indeterminacy would be generated. LA would need to block intermediate movement by adjunction by some additional assumption. No such complication arises in JPMT, as all node labels are generated based on the asymmetry of pair-merge.

## 2. JPMT is <u>adequate</u> to generate all relevant syntactic relations

Relevant syntactic relations:

4a) Categorization – combination with a root and a labeled head

b) Complementation – combination of a labeled head and a non-terminal

c) Non-complement argument assignment – "specifiers"

d) Adjunct relations

e) Head-head relations (and morphological complexity)

The first four relevant syntactic relations as generated in JPMT (subscript indicates label)

5a) Categorization <root, X>$_X$
   b) Complementation < YP <root, X>$_X$>$_X$
   c) Specifier <EA, <YP, <root, X>$_X$>$_X$>$_X$
   d) Adjunct <ZP <EA, <YP, <root, X>$_X$>$_X$>$_X$>$_X$

(tree abstracts away from linear relations)
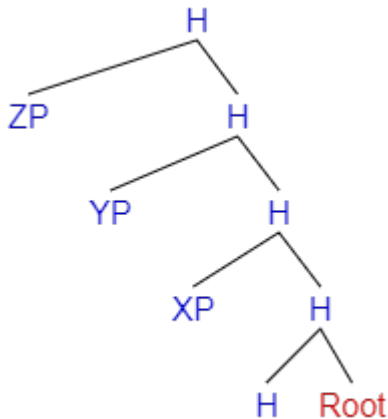
Categorization - Since I have a labeling filter and the root lacks a label, the root must adjoin to a categorized head X to produce a categorized head.

A **complement** YP pair-merged to a categorized head yields a complementation relation. If the categorized head were pair-merged to the complement, the label for the whole would be YP, which is not a complementation relation to categorized head X

What we call a **"specifier"** is just the first constituent to adjoin to a categorized head X and its complement (if X has a complement. The output bears the label of the pair <complement, categorized head> and it is where I assume an external argument (EA) can be assigned.

An **adjunct** constituent is just any ZP that adjoins subsequent to EA, if there is one.

All of these structures generated by JMOT are also generated by TMOT, though they are rarely appealed to. The dominance relations achieved, as represented by tree structure, are completely familiar.

TMOT can generate (8) employing pair-merge, or it can generate the same dominance relations by simple merge.

6a) Categorization {root, X}$_X$
  b) Complementation {YP {root, X}$_X$}$_X$
  c) Specifier {EA, {YP, {root, X}$_X$}$_X$}$_X$
  d) Adjunct {ZP {EA, {YP, {root, X}$_X$}$_X$}$_X$}$_X$

However, TMOT also allows mixed trees that achieve identical dominance relations, as demonstrated in (7) for complements.

7a) <complement, {root, v}$_v^o$>

  b) {complement, <root, v^o>$_v^o$ }

8)
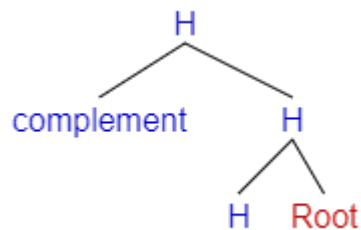
Altogether there are four ways to generate (8) in TMOT, but only one in JPMT.
Is there any reason we need this much generating power if we can already generate (8) with just pair-merge?

INTERNAL Pair-merge

In keeping with Chomsky's treatment of displacement structures, pair-merge allows for internal as well as external merge.

Internal pair-merge of a non-terminal to another nonterminal is now the description of all such relations, including wh-movement and A-movement, traditionally treated as substitutions (generated in Chomsky's system by simple merge) as well as VP fronting, heavy NP shift, clausal extraposition, extraposition from NP etc. which are traditionally treated as adjunctions. Iitalics indicate the unpronounced copy left in the position from which the constituent was internally pair-merged).

 9a) Which books did Mary read *which books*?
  b) The cat seems *the cat* to be out of the bag.
  c) Angry at Alfie though I was *angry at Alfie*, I still invited him.
  d) Selma will consider *any attempt to stop her from posting on Instagram* (to be) outrageous any attempt to stop her from posting on Instagram.
  e) It was apparent to us *that John would not be coming today* yesterday that John would not be coming today.
  f) We spoke to a Democrat *who was from Nashville* yesterday, who was from Nashville.

For (9a), pair-merge has the terms <[which books], [did Mary read *which books*]>
For (9b), pair merge has the terms <[the cat], [seems *the cat* to be out of the bag]>
For (9c), pair-merge has the terms <[angry at Alfie], [though I was *angry at Alfie*]>

For (9d), pair-merge has the terms <[any attempt to stop her posting on Instagram], [*any attempt to stop her posting on Instagram* to be outrageous>
(**Recall that neither pair-merge nor Pair-merge labeling makes any prediction about word order**)
For (9e), pair-merge has the terms <[that John would not be coming today], [was apparent to us *that John would not be coming today* yesterday]>
For (9f), pair-merge has the terms <[who was from Nashville], [spoke to a Democrat *who was from Nashville* yesterday]

9e Diagram
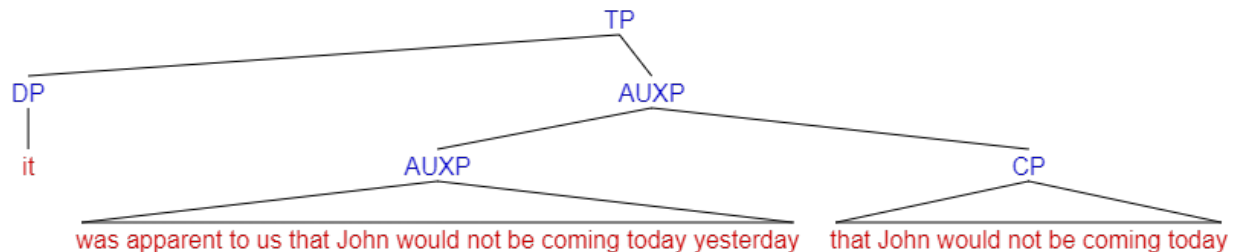
While it is unclear as to exactly where the (apparently) rightward-moved constituents land, it is always assumed that they are adjoined where they land in the clausal architecture (i.e., higher than the temporal adverb).

Notice also, that some adjunctions are to the left and others to the right, which is expected to be possible since pair-merge as posited here does not determine linear order (nor does it in Chomsky, 2004).

The instances of pair-merge in (9) are the only way to generate internal merge outputs in JPMT. TMOT can also generate all of (9a-f) with pair-merge, but it could also generate (9a-f) with set merge.

Not Demonstrated: Every theory will need pair-merge (even if simple merge is assumed).

- This would lead me into a discussion of head-head relations and head-movement, which we can discuss in the question period.

UPSHOT of Section 2: JPMT has the power to generate all of the domination relations that TMOT can generate, including the output of internal merge.

Simple Merge is not necessary for the relevant dominance relations to be generated.

3. What is lost if we give up TMOT?

At least three advantages have been claimed to be consequences of TMOT with LA.

10a) The EPP can be derived

b) Label indeterminacy explains why intermediate movement does not halt short of its destination (the halting problem).

    c) The existence of agreement is motivated as a solution to label indeterminacy

In this section I show that these advantages are only apparent. There is no principled explanation of them in TMOT.

The Halting Problem – Is it a problem?

Suppose we accept that movement is not feature-driven, as suggested by Chomsky (2008) and Safir (2010).

All structure-building is free (free Merge), and that includes internal merge. However, movement outputs fail if they do not reach positions where they can be interpreted.

Rizzi's criterial positions are those destinations. This is translated in labeling theory as the position where appropriate agreement can take place.

But the consequence of stopping in an uninterpretable position means only derivations that move to criterial positions will be successful. There is no halting problem.

Deriving *Who did John see*?

- *Who* originates as object of *see*
- *Who* moves to the edge of vP phase
- *Who* cannot be interpreted as an interrogative unless it reaches Spec CP
- Derivations where *who* stops at the vP phase edge will fail at LF
- Derivations where *who* lands in Spec CP will reach criterial positions and be acceptable.

An additional issue with the "halting solution"

Wh-movement from object position stops at the vP phase edge *as an adjunction*. Adjoined positions are invisible to further syntactic operations by TMOT hypothesis. **An internal pair-merge derivation for vP adjunction is not possible in TMOT.**

This means that wh-movement must proceed by set merge to a position higher than vP. But if VP is a phase and wh-movement moves beyond vP, then vP is not a phase. Alternatively, adjuncts are not defined because they are introduced by pair-merge, so the definition of adjunct is based on representations, not derivations.

For the structures in (9d,e,f) to be generated, then, internal pair-merge must be able to apply, since set merge cannot generate them (or they would never land in positions where LA could label them by agreement).

Labeling as the motivation for the existence of agreement is dubious.

Gerunds, infinitives and small clauses lack phi-feature agreement with their subjects, but both

require the establishment of subject-like arrangements.

Both infinitives and gerunds can have *there* subjects, which is usually the mark of an EPP effect.

11a) [PRO eating fish] didn't bother me

   b) [PRO to be late] would be unfortunate

   c) [Workers angry about the pay] is not what we expected when we agreed to this.

Appeal to abstract agreement between the PRO subject of gerunds and infinitive can be posited To save the day for the labeling theory, but how effectively?

   -   Even if one accepts *-ing* as the T equivalent in gerunds to account for (11a), in the small clause case subject example (11c) there is no candidate for T and still no visible agreement.

It is always possible to create additional abstract agreement to render a theory symmetrical,[2] and sometimes such idealizations lead to insight, but in this case, positing agreement relations that license positions where they can appear on the surface serves no other function – it is a stipulation to save a theory.

There are additional agreement problems that relate to what EPP is supposed to account for.

12a) Under the bed seems to be a good place to hide.

   b) Has under the bed been a good place to hide?

   c) Under the floorboards is/?*are a good place to hide.

It seems that in certain circumstances, PPs can be subjects in that they raise and invert with auxiliaries (see Safir, 1983), but it is not obvious that there is anything in PP to agree with, including the prepositional object. How is weak T propped up by agreement here?

The latter article also shows that small clauses must act as subjects (e.g. *Why shouldn't workers ready to bargain be the goal of our tactics?*). The same questions can now be asked about gerunds and infinitives as subjects as in (13), which can also raise and invert.

13a) Why would eating fish ever seem to be dangerous?

   b) To eat fish would seem to be a good solution

   c) Couldn't workers angry at the pay seem to show our incompetence in the negotiations?

Gerund and infinitive subjects have no overt phi-features that agree with T (at least in English). With less reason to believe that Case is necessary for a nominal (see, e.g., McFadden, 2004, among many others), there is less reason to assume that gerunds and infinitives are nominal subjects (at least in English).[3] LA

---

[2] For example, Abe (2018), building on Sato (2008) proposes that *there* "and its associate form a constituent underlyingly" and then *there* moves to subject position where, unlike other DPs, *there*'s uninterpretable and unvalued phi-features are then satisfied within the assumptions of reasoning about Agree. Goto (2017) (written after Abe's paper) attempts to adapt the Sato-Abe approach to Chomsky's (2015) LA, but to do so Goto invents a special head-head labeling relation in addition to LA.

[3] Saito (2016) argues that these problems are resolved by assigning shared Case between T and the element acting as a subject. This might apply to gerunds, if the possessive marker *'s* is a determiner heading a gerund, perhaps but it would not obviously extend to PP subjects. The agreement in Case would have to extend to contexts where Nominative is not involved but the gerund acts as a subject, e.g., *For treating this malady with antibodies to be*

was designed to account for why subjects must agree, but when they have no apparent features, it is hard to see how LA makes an accurate prediction without a slew of ancillary assumptions. In short, LA does not account for the EPP as it stands, so JPMT loses nothing in comparison with TMOT by failing to derive it.

Moreover, the most glaring problem for the justification of agreement was never solved.

14a) There are many men in the room.
   b) *Are many men in the room. (Declarative)
   c) *There is many men in the room.
- Agreement on T is not sufficient to license a null subject.
- Agreement does not capture why the notional subject can remain where it does, e.g. *there are [[three men] [in the room]]* (as Epstein et. al. 2014: 466, fn.13 point out and set aside), especially as agreement with T does not involve the common features of *there* and T.
- For such cases, some agreement relation must also license the low subject position, but once again there is no overt morphological motivation for an additional agreement relation with the expletive subject.
- *Agree* **is necessary independent of labeling, and the "weak T" stipulation does not help resolve the EPP problem either.**

Moreover, gerunds and infinitives can have *there* subjects as well, e.g, *there being no one in the room would bother Bill, For there to be only two people in the room would be unfortunate.* Abstract agreement would have to hold between *there* and the infinitive/gerund independent of the postverbal subject.

Upshot of section 3: TMOT does not offer any solution to the existence of agreement or EPP that follows naturally from TMOT mechanisms. Paradigm cases of EPP effects are not accounted for.

4.0 An advantage of JPMT

The derivational definition of c-command

     Preserving simple merge alongside pair-merge introduces a whole set of alternative derivations that can result in the same hierarchical structures that pair-merge generates, Consider the derivational definition of c-command, first introduced by Epstein (1999: 329).

15) Derivational c-command: X c-commands all and only the terms of the category Y with which X was merged in the course of a derivation.[4]

At the time (the article was circulated in 1994), Merge and Move were separate operations and, as Hasegawa (1996) points out, Epstein did not even consider additional issues that arise for adjunction

---

effective… and *John considers arresting the demonstrators to be a poor solution*. What is agreeing to license the relation between two non-terminals as sisters in these constructions?

[4] This is Epstein's preliminary definition, but the differences between this and his final version do not concern us here.

(before Chomsky, 2004 reformulated adjunction as pair-merge), or for movement that results in adjunctions (which Hasegawa also noted and set aside).
- Since adjunctions also create c-command relations, a theory with both set merge and pair-merge must create a cover term "Merge" that encompasses both set merge and pair-merge.
- These complications reduce to the following definition once simple merge is eliminated.

16) C-command: If X is pair-merged to Y then X c-commands Y and everything Y dominates.

No special provisos for other operations creating c-command relations need be referred to.

Something to gain – The asymmetry of pair-merge

It is not obvious that there is any situation where fully symmetric (mutual) c-command relations are empirically needed, yet that is what set merge generates. See Brody's (2002: 28) objection to derivational c-command based on the absence of asymmetry in the operation of structure-building

JPMT explains why: Sisters created by pair-merge are, by definition, never symmetrical (see also Zwart, 2011).
Is there ever an empirical reason to assume, for example, that a predicate must c-command something inside what it is predicated of? ***From the perspective of pair-merge, only the adjoiner can c-command something in the adjoinee.***

17a) [[every man's mother] [praises his new spouse]]
   b)*[[his new spouse] [praises every man's mother]]
   c)
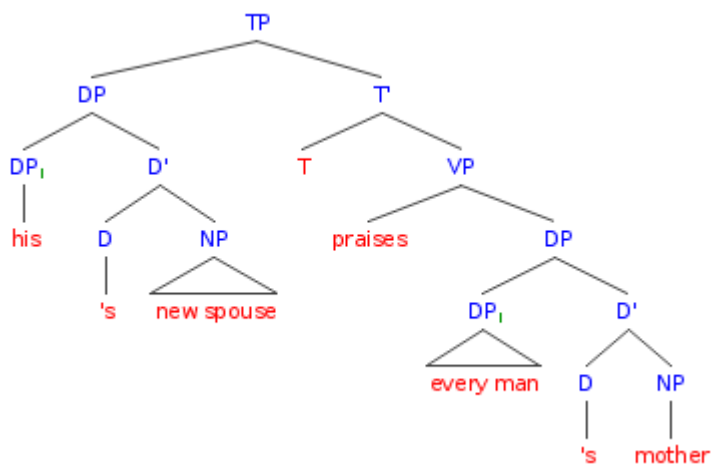
In a theory that allows [praises x's spouse/mother] to c-command the contents of the subject (a symmetric theory of c-command), the two sentences (17a,b) have the same violation, unless it is stipulated that the pronoun cannot be contained in a *nominal* that c-commands the quantifier that binds it. This is because a quantifier is embedded in a constituent (T' in (17c)) that c-commands a bound pronoun and a bound pronoun is embedded in a constituent (the subject DP in (17c)) that c-commands its quantifier. The same is true for (17a) except the roles of the subject DP and T' are reversed. In the JPM theory (and in the other ordered pair output theories), the subject DP is adjoined to [T XP], so it asymmetrically c-commands the contents [T XP] in (17c), (without further stipulation limiting c-

command to nominals[5]). In order to properly capture this weak crossover contrast, the weak crossover configuration can now be reduced, given asymmetric c-command as in (16), to (18).

18) A pronoun cannot be dominated by a constituent that c-commands the quantifier that binds it.[6]

This improvement is a direct result of restricting all merge to pair-merge, which is inherently asymmetric.

5. Conclusions

A syntactic theory with only pair-merge and not simple/set merge is simpler than the Two Merge Operations Theory and much closer to the minimalist ideal.

The Just Pair-merge Theory is adequate to generate all the syntactic relations that are known to be necessary.

Purported advantages of TMOT do not provide adequate explanations for phenomena that JMOT does not address.

JMOT provides an inherently asymmetric derivational c-command definition, which set/simple merge must stipulate.

(And by the way, the adjunct-argument distinction can no longer be defined operationally, i.e., pair-merge vs. set merge – Instead the syntactic aspects of the adjunct-argument distinction must be defined on representations).

---

[5] See Safir (2004: 52) where this detail is observed and stipulated.
[6] Safir (2004) treats this as a correlate of his Independence Principle. As stated, (18) abstracts away from reconstruction phenomena, though most theories of reconstruction based on copy theory end up with structural analyses for which (18) is true.