# There is no need for Late Merger – SQUIB

*Carlos Muñoz Pérez*
*Universidad de Buenos Aires & Georg-August-Universität Göttingen*

## 1. Countercyclic operations in the explanation of reconstruction patterns

If Copy Theory (Chomsky 1993) is assumed, the existence of an asymmetry between complements and adjuncts regarding Reconstruction is unexpected. However, it has been observed in several occasions (cf. Lebeaux 1988, Fox 1999, among others) that a R-expression inside an adjunct modifying a wh-phrase bleeds Condition C, while the same does not happen with a complement modifying a wh-phrase.

(1) [Which argument [that John$_i$ made]]$_j$ did he$_i$ believe $t_j$ ?

(2) *[Which argument [that John$_i$ is a genius]]$_j$ did he$_i$ believe $t_j$ ?

To explain this kind of pattern, Lebeaux (1988) proposes assuming that adjuncts may be introduced into syntactic structure countercyclically trough *Late Merger*. The basic idea is that the adjunct is merged inside the antecedent of the trace after movement has taken place. In terms of Copy Theory, this means that the derivation for (1) would involve the steps in (3), where *which argument* A'-moves from its thematic position as in (3a) and then the adjunct CP *that John made* is combined with the noun *argument* as in (3b).

(3) a. *A'-movement*
[Which argument] did he believe [which argument]?
b. *Late Merger*
[Which argument [that John$_i$ made]] did he$_i$ believe [which argument]?

The reason why Late Merger is not available for complements as the one in (2) follows, according to Lebeaux, from the *Projection Principle* (Chomsky 1981): since the lexical head N carries a subcategorization feature selecting a complement, merging them countercyclically would not satisfy such requirement (at least immediately). Such a derivational restriction is

restated by Fox (2002) in terms of properties of the final syntactic representation. Fox assumes that *Trace Conversion* is required to apply on the lowest link in a chain to obtain a valid operator-variable representation under Copy Theory:

(4) *Trace Conversion (from Fox 2002: 67)*

a. Variable Insertion: (Det) Pred→ (Det [Pred λy(y=x)]

b. Determiner Replacement: (Det) [Pred λy(y=x)] → the [Pred λy(y=x)]

In few words, Trace Conversion requires the restrictor of a determiner being a predicate of type <e,t> to interpret it compositionally with λy(y=x), also a <e,t> predicate, through *Predicate Modification* (cf. Heim & Kratzer 1998). A complement-taking noun is an element of type <<e,t>,<e,t>>: it requires merging with a <e,t> complement to return a <e,t> type of projection. But since Late Merger only introduces the relevant complement in a higher copy, the lowest link in the chain remains a wrong input to Trace Conversion, being this way uninterpretable. Thus, it is predicted that complements cannot be introduced countercyclically.

The asymmetry attested in (1) and (2) is not the only one that can be found in the realm of reconstruction patterns. In particular, it has been observed that unpronounced copies of A-movement bleed Condition C (cf. Chomsky 1995), in a similar way as R-expressions inside adjuncts in A'-movement.

(5) [The claim that John$_i$ was asleep]$_j$ seems to him$_i$ to be correct $t_j$

The acceptability of (5) is surprising if sentences as (6) are considered, since they show that unpronounced copies of A-movement may be in fact interpreted.

(6) [The best student of his$_i$ class]$_j$ seems to [every professor]$_i$ $t_j$ to be a genius.

This kind of pattern has led to the conclusion that A'-movement obligatorily leaves a contentful copy of the moved constituent, while the same is optional for A-movement (cf. Takahashi 2010).

Takahashi & Hulsey (2009) had proposed an explanation for this generalization based on an operation called *Wholesale Late Merger* (WLM). Their proposal is an extension of Lebeaux's

(1988) analysis since they assume that an NP may be countercyclically merged with its selecting D under certain conditions. Schematically:

(7) a. *A-Movement*

$[_{XP}\ [_{DP}\ D]\ \ldots\ [_{YP}\ [_{DP}\ D]]]$

b. *Wholesale Late Merger*

$[_{XP}\ [_{DP}\ D\ NP]\ldots\ [_{YP}\ [_{DP}\ D]]]$

According to Takahashi & Hulsey, Case plays an important role on constraining the applicability of WLM. Basically, they assume that NPs (besides DPs) require receiving Case, thus they cannot be merged countercyclically outside the c-command domain of its Case assigner. Importantly, if it is considered that A'-movement always takes place from Case-marked positions, they also predict that WLM cannot apply on such type of dependencies. Thus, it is predicted that A'-movement always must leave behind a contentful copy.

This system derives the troubling patterns regarding data as (5) and (6). The acceptability of (5) follows from an application of WLM introducing lately the lexical content of the moved DP. Thus, there is no R-expression in the lowest link in the chain that may cause a Condition C violation.

(8) a. *A-Movement*

[The] seems to him to be correct [the]

b. *Wholesale Late Merger*

[The claim that $John_i$ was asleep] seems to $him_i$ to be correct [the]

Regarding (6), WLM applies on an intermediate copy of the raising D. And since this copy is in the scope of *every man*, a bound lecture of the possessive pronoun is obtained.

(9) a. *Cyclic/Intermediate A-movement*

[The] to be [the] a genius.

b. *Wholesale Late Merger*

[The best student of his class] to be [the] a genius.

c. *A-movement*

[The best student of his$_i$ class] seems to [every professor]$_i$ ~~[the best student of his$_i$ class]~~ to be [the] a genius

In a nutshell, WLM allows explaining the distinction between A and A'-movement regarding Reconstruction, and it also provides a unified theory of countercyclical operations.

**2. A cyclical implementation**

The conceptual drawbacks on assuming WLM are palpable. The idea that syntactic operations must observe strict cyclicity is a theoretical desideratum in generative linguistics since, at least, Chomsky (1965). Thus, it would be highly desirable being able to capture the interpretative patterns derived by WLM within a cyclic approach to syntax.

The system I will propose here takes (10) as a main assumption.

(10) *Assumption on Agree*

If an XP enters in an *Agree* relation with a (non-defective) head $H^0$, $H^0$ assigns a feature [H] to XP.

So, for example, if a DP agrees with an inflected T head, the DP will be assigned a T-feature (nominative Case); or if a wh-phrase agrees with an interrogative complementizer $C_{INT}$, the wh-phrase will be assigned a Q-feature[1] (which, by hypothesis, is related with quantificational force and scope).

Now, suppose an scenario where a non-defective *Probe* $Y^0$ with an EPP-feature enters in an Agree relation with a *Goal* $\alpha_{\{...\}}$, where {…} is used to indicate the set of features of α. The newly generated copy of α in [Spec,$Y^0$] would receive an additional Y-feature, becoming the object $\alpha_{\{Y, ...\}}$.

(11) $[_{YP}\ Y^0_{EPP}\ [_{XP}\ \ldots\ \boldsymbol{\alpha}_{\{...\}}]]$ → $[_{YP}\ \boldsymbol{\alpha}_{\{Y,...\}}\ [_{Y'}\ Y^0_{EPP}\ [_{XP}\ \ldots\ \boldsymbol{\alpha}_{\{...\}}]]]$

[1] I am assuming that A'-dependencies also involve Probe-Goal mechanisms. Thus, intermediate/non-wh complementizers will count as "defective" regarding wh-phrases.

There is an *inclusion* relation between both copies of α: the set {…} is a subset of {Y, …} ($\alpha_{\{...\}} \subseteq \alpha_{\{Y, ...\}}$). This relation between copies is going to arise every time a new copy is generated[2]. Since this inclusion relation is systematic, it is possible to use it as one of the *Condition on Chain Formation*:

(12) *Conditions on Chain Formation*

Two constituents α and β are links in the same chain iff

a. α is a superset of β;

b. α c-commands β;

c. There is no syntactic object γ between α y β such as (i) β is a superset of γ and (ii) γ is not a superset of α.

Before explaining how this system derives the same results than WLM, some considerations regarding the featural content of DPs are in order. The typical set of features composing a DP are *D*, *Person*, *Number* and *Gender*. However, at least two of those features, Number and Gender, are not inherent to the D head but transmitted to it through agreement with the NP. Thus, a *bare D* (a D head without its NP complement) would be composed by the features Person and D. Schematically:

(13) a. *Full DP before agreement between D and NP*

$[_{DP}\ D_{\{D, Per\}}\ [_{NP}\ N_{\{Num, Gen\}}]$

b. *Full DP after agreement between D and NP*

$[_{DP}\ D_{\{D, Per, Num, Gen\}}\ [_{NP}\ N_{\{Num, Gen\}}]$

(14) *Bare D*

$[_{DP}\ D_{\{D, Per\}}]$

Among the functions that a DP complies in the syntactic derivation, a bare D would not be able to check/valuate the uninterpretable morphology of a Case assigner (T or v), since it lacks

[2] Even when the new copy does not receive an additional feature, it will be a superset of the original copy (e.g., $\alpha_{\{...\}} \subseteq \alpha_{\{...\}}$). The notions of subset/superset assumed here must not be confused with the more complex notions of *proper subset/superset* (e.g., A is a Proper Subset of B if $A \subseteq B$ and $A \neq B$).

Number features. And since Case assignment is usually understood as a counterpart of full φ-agreement, bare Ds cannot receive Case. However, a bare D can participate in thematic relations since, as Longobardi (1994) observed, the D-feature allows constituents to receive a θ-role.

Thus, under an approach to chain formation as the one depicted in (12), a full DP and a bare D would form a chain CH if they are in a local configuration.

(15) $[_{YP}\ DP_{\{D, Per, Num, Gen\}} \ldots [_{XP}\ D_{\{D, Per\}}\ ]$ CH = (DP, D)

Notice that there is no transformational relation between these two elements. Nevertheless, if the DP is base generated in a higher and c-commanding position (and there are no interveners), it will form a chain with a lower bare D. Since there is no syntactic relation between these two elements, there is no real reason to assume that such a chain is constructed in the narrow syntax. Thus, it will be proposed as a working hypothesis that the mechanism introduced in (12) applies independently at PF and LF, forming chains in accordance to interface specific requirements.

This system is ruled by the same conditions on Case assignment constraining WLM: since bare Ds cannot check the full set of φ-features of Case assigning heads, it is required that a full DP is merged in their domain to do so. So, for example, a structure like (16) would inevitably lead to a crash since T's Number feature (*u*Num) cannot be valued/checked by (a bare) D.

(16) $*[_{TP}\ D_{\{D, Per\}}\ [_{T'}\ T_{\{uNum, Per\}}\ [_{VP}\ \ldots\ D_{\{D, Per\}}]]]$

However, a derivation where a full DP is base generated in the domain of T would be grammatical since T's φ-features would be valued/checked.

(17) $[_{TP}\ DP_{\{D, Per, Num, Gen\}}\ [_{T'}\ T_{\{Num, Per\}}\ [_{VP}\ \ldots\ D_{\{D, Per\}}]]]$

Since WLM and the inclusion-based system are constrained in exactly the same way, it is expected that both accounts would be extensionally equivalent regarding their predictions. This claim, of course, requires further research and comparison between both proposals. For now, it

will be shown how the asymmetries in reconstruction discussed in the previous section can be derived by the inclusion-based system without using any countercyclical operation.

The conditions in (12) predict that a wh-phrase containing an adjunct will form a chain with a lower wh-phrase in the structure even if the later does not contain an adjunct. Thus, a sentence as (1), repeated for convenience in (18a), would receive the analysis sketched in (18b), where the base generated wh-phrase *which argument that John made* forms a chain with the wh-phrase *which argument* in its thematic position. Since (i) v's features are valued/checked, (ii) Trace Conversion can apply on *which argument* and (iii) there is no copy of *John* in the domain of *he,* the sentence is acceptable.

(18) a. Which argument that $John_i$ made did $he_i$ believe?

b. [Which argument [that $John_i$ made]] did $he_i$ believe [which argument]?

{Q, Acc, Dwh, φ} {Acc, Dwh, φ}

There are two reasons for the sentence in (2), repeated in (19a), being unacceptable. If a movement relation is assumed to underlie this sentence, that would produce a violation of Condition C, as sketched in (19b).

(19) a. Which argument that $John_i$ is a genius did $he_i$ believe?

b. *[Which argument that $John_i$ is a genius] did $he_i$ believe [which argument that $John_i$ is a genius]

On the other hand, if the complement containing the R-expression *John* would be introduced in a higher base-generated version of *which argument*, as in (20a), that would produce an uninterpretable input for Trace Conversion in the lowest link of the resulting chain.

(20) a. [Which argument [that $John_i$ is a genius]] did $he_i$ believe [which argument]

b. *$[_{NP}$ argument$_{<<e,t>,<e,t>>}$ $\lambda y(y=x)_{<e,t>}]$ SINCE PREDICATE MODIFICATION CANNOT BE APPLIED

Hence, the inclusion-based system of chain formation captures the reconstruction asymmetry between complements and adjuncts.

The distinct behavior of A and A'-movement regarding reconstruction is also captured. Much in the same way than in Takahashi & Hulsey's (2009) approach, a bare D cannot be the tail of an A'-chain since A'-movement requires moving a Case assigned constituent, and bare Ds cannot receive Case. Thus, the NP complement of a determiner must be always present for A'-movement, predicting reconstruction.

For the cases as (5), repeated for convenience in (21a), where A-movement bleeds Condition C, a bare D is assumed to occupy the thematic position of the moved constituent. Later, the full DP is base-generated in its Case-marking position. And since the features of the bare D are included in the set of features of the full DP, a chain is formed.

(21) a. The claim that John$_i$ was asleep seems to him$_i$ to be correct.

b. [$_{DP}$ The claim that John$_i$ was asleep] seems to him$_i$ to be correct [$_{DP}$ D]
{T, D, Gen, Num, Per} {D, Per}

Given that there is no R-expression in the c-command domain of the pronoun *his,* no violation of binding principles arises.

Regarding the cases where reconstruction is attested with A-movement, as in sentence (6), repeated in (22a), it is assumed that a bare D occupies a thematic position. Then, a full DP is base-generated in the specifier position of an infinitive T checking its EPP-feature. When finite T is introduced into the derivation, it attracts to its specifier position the closest DP in its c-command domain. Thus, a copy of the full DP valuates/checks the φ-features of T and receives nominative Case (a T-feature).

(22) a. The best student of his$_i$ class seems to [every professor]$_i$ to be a genius.

b. [$_{DP}$ The best student of his$_i$ class] seems to [every professor]$_i$ [$_{DP}$ the best
{T, D, Gen, Num, Per} {D, Gen, Num, Per}
student of his$_i$ class] to be [$_{DP}$ D] a genius.
{D, Per}

In sum, it has been demonstrated that the inclusion-based system can deal with, at least, the basic asymmetries regarding reconstruction without positing countercyclical operations.

## 3. Concluding remarks

In this brief squib, it has been shown how a flexible definition of identity between links in a chain allows for a cyclic treatment of reconstruction asymmetries. Interestingly, this approach seems to be available only if chains are assumed to be constructed on representational grounds at the interfaces. In other words, given a syntactic derivation Δ and its output, the syntactic representation Σ, the constituents α and β in Σ are members of a chain if Σ complies with some conditions, independently of the operations applied to α and β during Δ.

## References


Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge: MIT press.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris

Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (eds.), *The View from Building 20*. Cambridge: MIT Press.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press

Fox, D. (1999). Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30. 53-96.

Fox, D. (2002). Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33. 63-96.

Heim, I. & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford: Blackwell.

Lebeaux, D. (1988). *Language acquisition and the form of grammar*. Doctoral Dissertation, UMass Amherst.

Longobardi, G. (1994). Reference and proper names: a theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25. 609–665.

Takahashi, S. (2010). Traces or Copies, or Both – Part I: Characterizing Movement Properties. Language and Linguistics Compass 4/11. 1091-1103.

Takahashi, S. & Hulsey, S. (2009). Wholesale Late Merger: Beyond the A/A' Distinction. *Linguistic Inquiry* 40. 487-526.