

## There is no need for Late Merger

Carlos Muñoz Pérez

Universidad de Buenos Aires & CONICET

There are two ways in which a set of constituents can be identified as a chain under Copy Theory: (i) either they are marked as non-distinct by an operation (e.g., Copy, Select), or (ii) they are recognized as non-distinct by some inspection of their properties (i.e., their features). In this squib, I argue that the second approach allows for a simpler and more elegant explanation of the reconstruction asymmetries that are standardly captured by assuming counter-cyclical operations as Late Merger and Wholesale Late Merger.

### 1. Introduction

In its simpler form, *Copy Theory* (Chomsky 1993) predicts that every gap of syntactic movement should be able to reconstruct. However, there are systematic exceptions to this prediction: (i) adjuncts inside wh-phrases do not violate Condition C (cf. (1)), while (ii) complements inside A-moved elements may disallow reconstruction in some cases (cf. (2a)) or apparently reconstruct in other contexts (cf. (2b)).

- (1) a. [Which argument [that John<sub>i</sub> made]]<sub>j</sub> did he<sub>i</sub> believe <sub>t<sub>j</sub></sub> ?  
b. \*[Which argument [that John<sub>i</sub> is a genius]]<sub>j</sub> did he<sub>i</sub> believe <sub>t<sub>j</sub></sub> ?
- (2) a. [The claim that John<sub>i</sub> was asleep]<sub>j</sub> seems to him<sub>i</sub> to be correct <sub>t<sub>j</sub></sub>  
b. [His<sub>i</sub> picture of [the president]<sub>k</sub>]<sub>j</sub> seems to [every man]<sub>i</sub> <sub>t<sub>j</sub></sub> to be seen by him<sub>k</sub> <sub>t<sub>j</sub></sub> to be an intrusion.

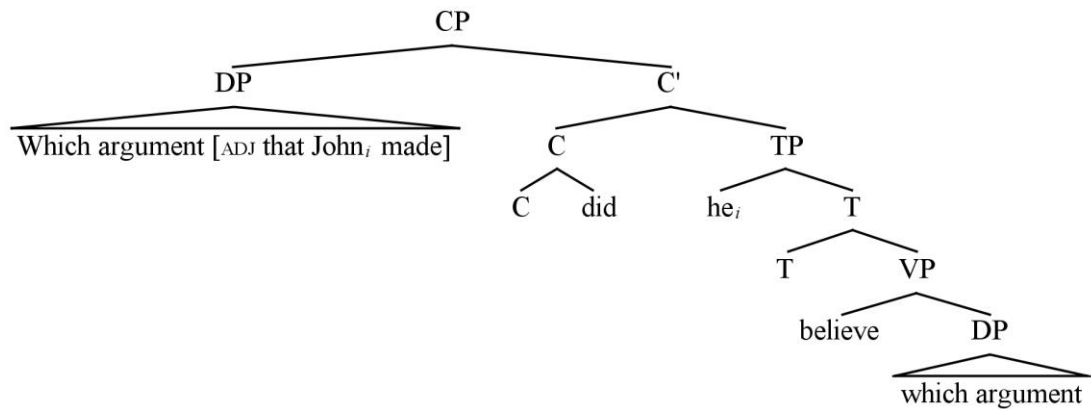
According to Lebeaux (1988), cases of *anti-reconstruction* as (1a) do not involve a real violation of Condition C. I will refer to his approach as the *Lebeauxian Approach to Anti-Reconstruction* (LATAR).

### (3) *Lebeauxian Approach to Anti-Reconstruction* (LATAR)

Apparent violations of Condition C follow from the absence of the constituent containing the relevant R-expression in some links of the movement chain.

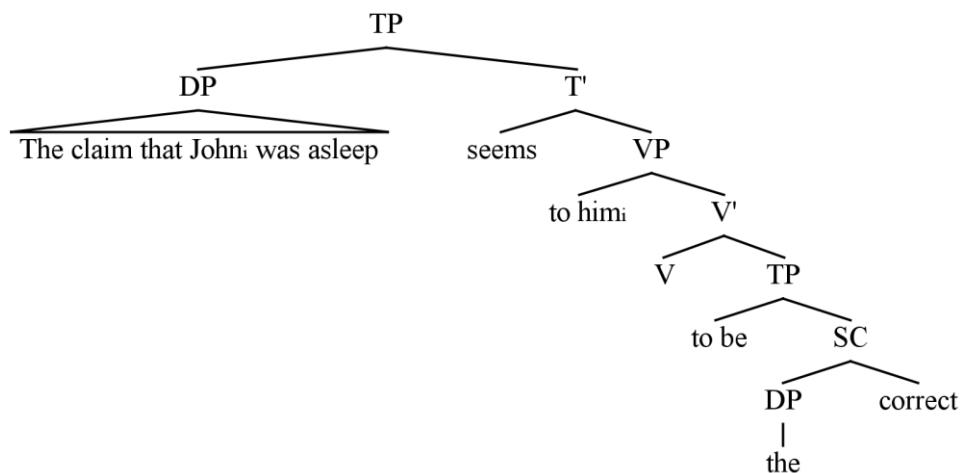
According to LATAR, there is no occurrence of the adjunct *that John made* inside the structure of the lower link of the movement chain in (1a). Therefore, the pronoun *he* cannot trigger a Condition C violation by being coindexed with *John*.

(4)



Takahashi & Hulsey (2009) extended LATAR to capture cases of A-movement as (2a). In this sentence, the complement NP of the determiner *the* appears only in the head of the movement chain.<sup>1</sup> Consequently, *him* cannot cause a Condition C violation.

(5)

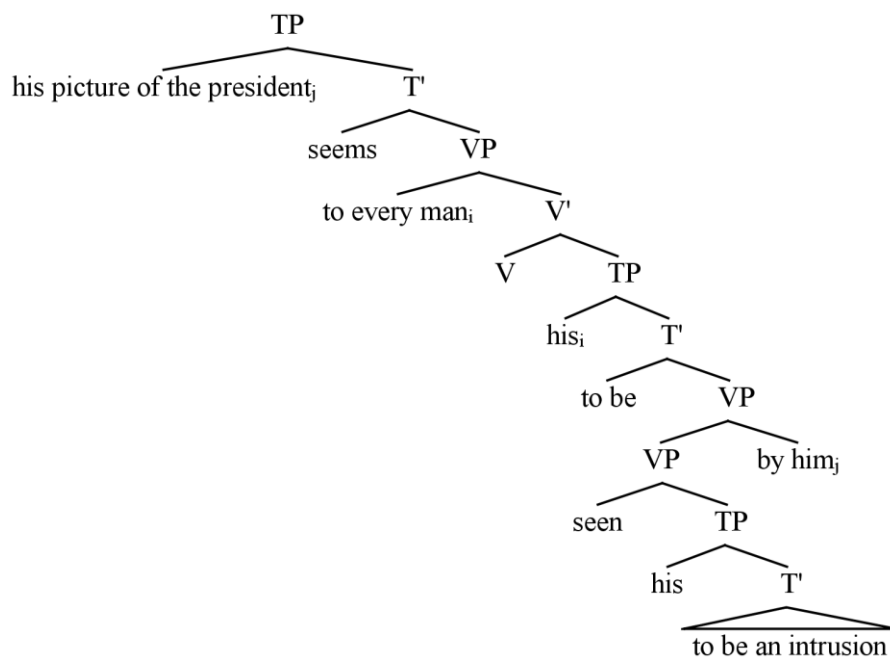



---

<sup>1</sup> For convergent evidence to assume the presence of bare determiners in low links of movement chains, see Stanton (to appear).

Sentence (2b) is also captured by this approach. While the NP *picture of the president* does not appear in the lower occurrences of the chain, the pronominal determiner *his* is present throughout the entire derivation, so it can be bound by *every man*.

(6)



An implicit feature of LATAR is that it allows chain formation between non-isomorphic syntactic objects. As it can be seen in (7), such an explanation for anti-reconstruction relies on movement chains being formed by constituents that may or may not contain a R-expression as part of their structure.

- |     |    |  |         |
|-----|----|--|---------|
| (7) | a. | CH = (which argument that John made, which argument) | cf. (4) |
|     | b. | CH = (the claim that John was asleep, the)           | cf. (5) |
|     | c. | CH = (his picture of the president, his, his)        | cf. (6) |

A priori, this idea clashes with the assumption under Copy Theory that movement chains are sets of non-distinct syntactic objects. Particularly, it seems that not every definition of Non-Distinctiveness may predict straightforwardly the formation of the chains in (7). The main aim of this paper is showing that some particular definition of Non-Distinctiveness deals in a more parsimonious way with LATAR.

The structure of the paper is as follows. Section 2 discusses the “standard” approach to Non-Distinctiveness and the way it deals with LATAR. Section 3 introduces the main proposal: following Muñoz Pérez (2016), Non-Distinctiveness is assumed to be computed as an inclusion relation between the features of syntactic objects. Anti-reconstruction effects as the ones observed in (1) and (2) are shown to be predicted by such a definition. Section 4 presents some mechanisms required by any implementation of LATAR, showing that the present account does not require additional assumptions. Finally, section 5 contains the conclusions.

## 2. Indexical Sameness and Late Merger

A “standard” approach to Non-Distinctiveness of copies assumes that two or more syntactic objects are “the same” if they are marked as such by a grammatical operation (cf. Chomsky 1995, Nunes 1995, 2004, Bošković & Nunes 2007). Thus, for example, the constituent *John* in (8a) receives an index by the operation *Select* (cf. Chomsky 1995). Given that the Copy operation also replicates this index (cf. 8b), the newly formed occurrence of *John* will also carry the same index. Since both syntactic objects carry the same marking, they are identified as non-distinct in (8c).

- (8) a. *Derivational step a*  
 $K = [_{TP} \text{ was } [_{VP} \text{ kissed John}_1]]$
- b. *Copy*  
 $K = [_{TP} \text{ was } [_{VP} \text{ kissed John}_1]]$   
 $L = \text{John}_1$
- c. *Derivational step a+1*  
 $[_{TP} \text{ John}_1 [_{T'} \text{ was } [_{VP} \text{ kissed John}_1]]]$

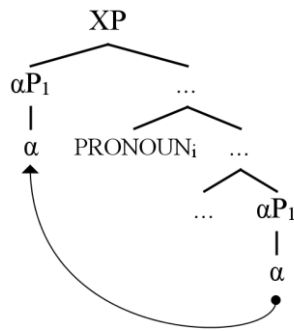
For ease of presentation, I term this type of definition of Non-Distinctiveness the *Indexical Definition of Sameness*, or *Indexical-S* for short.

### (9) *Indexical-S*

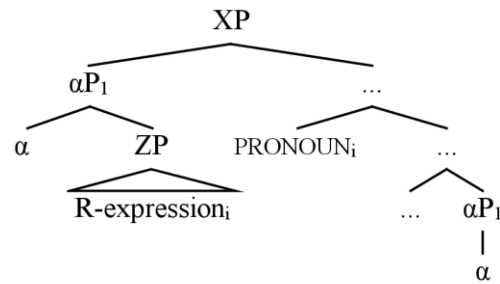
Two constituents  $\alpha$  and  $\beta$  are “the same” if they are assigned the same index by a grammatical device.

Under Indexical-S, forming a chain with two (or more) non-isomorphic constituents to implement LATAR requires two complex steps: (i) generating two (or more) strictly identical copies carrying the same index, as in (10a), and (ii) applying an additional operation on the higher copy to introduce the constituent with the relevant R-expression, as in (10b).

(10) a.



b.



The operations *Late Merger* (Lebeaux 1988) and *Wholesale Late Merger* (Takahashi & Hulsey 2009) are two versions of the type of mechanism depicted in (10b). In a nutshell, both involve merging a constituent inside a derived specifier in a counter-cyclic fashion.

Adopting Indexical-S to implement LATAR introduces three problems. The first one was already mentioned: assuming the derivational steps in (10) involves abandoning strict cyclicity. Thus, for example, both *Late Merger* and *Wholesale Late Merger* violate the *Extension Condition*.

(11) *Extension Condition* (Chomsky 1993)

Syntactic operations must extend the tree at the root.

The second problem involves the so-called *Inclusiveness Condition* (Chomsky 1995), a constraint on introducing non-lexical material into the syntactic derivation. Given that indexes (or any other kind of markings) are not intrinsic properties of any lexical item, adopting Indexical-S violates Inclusiveness.<sup>2</sup>

---

<sup>2</sup> Even more, given that there are no current viable alternatives to Indexical-S, some authors criticize Copy Theory based on this issue: e.g., “Copy Theory by itself does not resolve the tension between Inclusiveness and the displacement property of natural language” (Neeleman & van de Koot 2010: 332).

Third, Indexical-S is just a marking mechanism adopted due to the difficulty of defining Non-Distinctiveness in a principled way. A true theory of why similar syntactic objects are interpreted as some kind of unit, i.e., a chain, must rely on independent linguistic principles, and not on an inductive device proposed to fill the gaps in the theoretical apparatus.

### 3. Inclusion-based Sameness

According to Muñoz Pérez (2016), syntactic features  $\langle \text{Att}, \text{VAL} \rangle$  consisting on a pair of attribute and value are interpreted at the interfaces as privative values  $\langle \text{VAL} \rangle$ . Thus, the set of features in (12a) is “seen” at PF and LF as the set of values in (12b). Given that the Case feature  $\langle \kappa, \_ \rangle$  lacks a value in (12a), it is not represented at the interfaces, being *uninterpretable*.

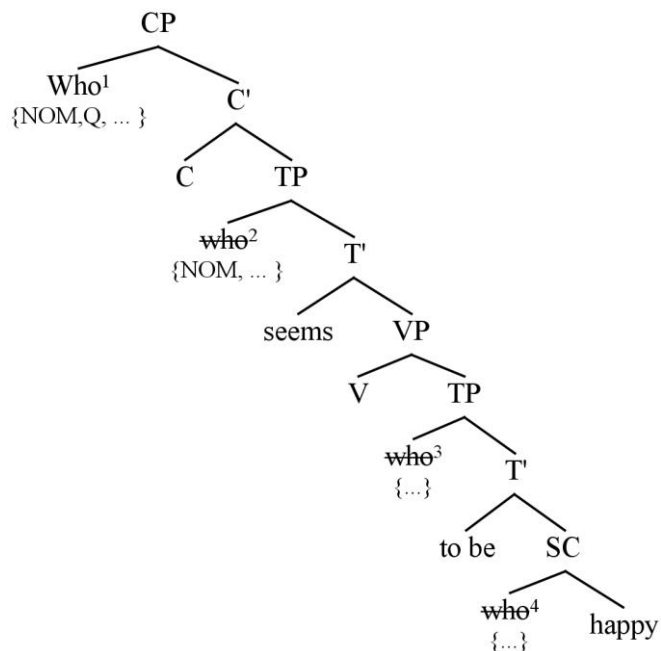
- |      |    |  |            |
|------|----|--|------------|
| (12) | a. | $\{ \langle \text{Cat}, \text{D} \rangle, \langle \text{Gen}, \text{FEM} \rangle, \langle \text{Num}, \text{SG} \rangle, \langle \text{Per}, 3 \rangle, \langle \kappa, \_ \rangle \}$ | SYNTAX     |
|      | b. | $\{ \text{D}, \text{FEM}, \text{SG}, 3 \}$   | INTERFACES |

Consider the derivation in (13). The wh-pronoun *who* carries two unvalued activity features (cf. Chomsky 2001), a Case feature  $\langle \kappa, \_ \rangle$  and a Wh-feature  $\langle \text{Wh}, \_ \rangle$ , that are satisfied in the Spec,T and Spec,C positions respectively.

- |      |    |  |
|------|----|--|
| (13) | a. | $[\text{CP Who}^1 [\text{TP } \text{who}^2 \text{ seems } [\text{TP } \text{who}^3 \text{ to be } [\text{who}^4 \text{ happy}]]]]?$  |
|      | b. | $\text{Who}^1 \{ \langle \kappa, \text{NOM} \rangle, \langle \text{Wh}, \text{Q} \rangle, \dots \} \dots \text{who}^2 \{ \langle \kappa, \text{NOM} \rangle, \langle \text{Wh}, \_ \rangle, \dots \} \dots \text{who}^3 \{ \langle \kappa, \_ \rangle, \langle \text{Wh}, \_ \rangle, \dots \} \dots$<br>$\text{who}^4 \{ \langle \kappa, \_ \rangle, \langle \text{Wh}, \_ \rangle, \dots \}$ |

Given that unvalued features are not visible at the interfaces, the syntactic representation in (13) is interpreted at PF and LF as in (14).

(14) *Representation at the interfaces*



Notice that there is an *inclusion* relation between the (features of the) occurrences of *who*. Such a relation will arise systematically for every new copy of a constituent, so it may be capitalized to define Non-Distinctiveness. For ease of presentation, I will call this definition the *Inclusion-based Definition of Sameness*, or *Inclusion-S* for short.

(15) *Inclusion-S*

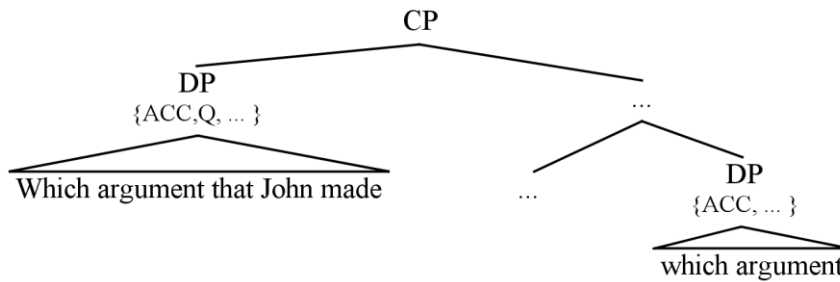
Two constituents  $\alpha$  and  $\beta$  are non-distinct if:

- a.  $\alpha$  c-commands  $\beta$ ,
- b. the features of  $\beta$  are a subset of the features of  $\alpha$ ,
- c. there is no  $\delta$  between  $\alpha$  and  $\beta$  being a proper subset of  $\alpha$  or a proper superset of  $\beta$ .

According to the definition in (15), two syntactic objects will form a chain if the features of the lower element are contained in the set of features of the higher constituent, and there are no intervening elements between them. Such a mechanism of chain formation makes no reference to transformational devices relating the members of a chain (e.g., operations as *Move-a* or *Copy*), so it predicts that two base generated DPs can form a chain as far as they comply with *Inclusion-S*.

Consider the structure of (4) under the following derivation. The DP *which argument* is externally merged in the complement position of the verb *believe*. The derivation proceeds as usual until it is necessary merging a wh-phrase in the Spec,C position; at that moment, a new DP *which argument that John made* is base-generated there, valuing its Wh-feature <Wh,\_\_\_> through Agree with the interrogative complementizer.<sup>3</sup>

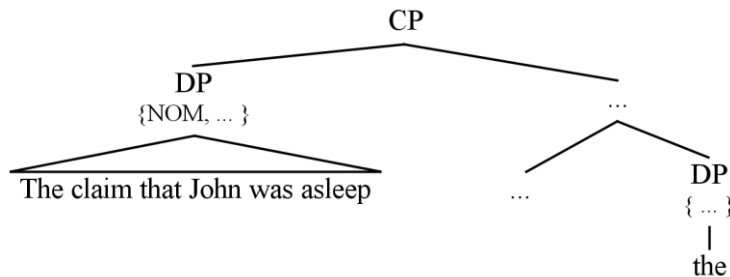
(16)



Given that (i) there is a c-command relation, (ii) the features of the c-commanding element contain the features of the other DP, and (iii) there are no interveners between them, both constituents are computed as non-distinct at the interfaces, forming the chain in (7a) without any additional assumption.

The same kind of derivation allows forming the chains in (7b) and (7c). In both cases, bare determiners projecting DPs form chains with base-generated and Case-marked full-DPs in subject positions.

(17)

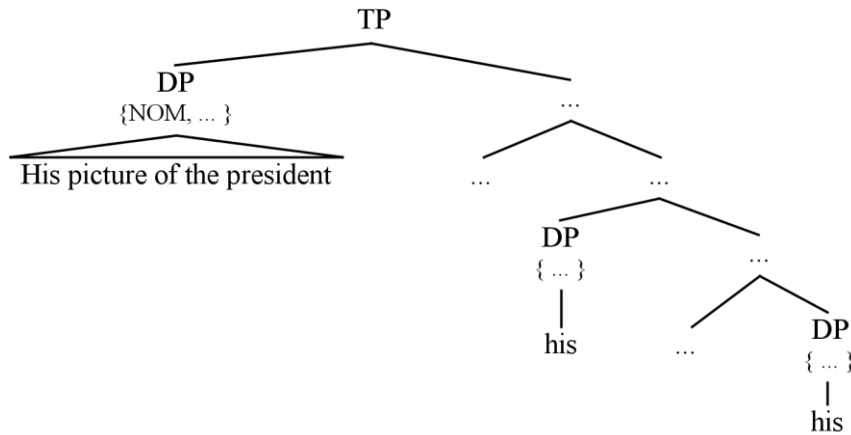



---

<sup>3</sup> Questions may arise regarding the accusative Case feature in the higher DP. This feature may be assumed to be valued through Agree: the DP in Spec,C acts as a Probe looking for a Goal with a matching Case feature, and the only available candidate for such a dependency is the lower DP *which argument* since it carries an unvalued Wh-feature marking it as active.



(18)



The implementation of LATAR based on Inclusion-S allows solving the pointed problems with assuming Indexical-S as definition of Non-Distinctiveness of copies. First, it provides a cyclic derivation for the chains in (7) since no counter-cyclic mechanisms are introduced and the Extension Condition is respected. Second, it complies with the Inclusiveness Condition as it avoids introducing indexes or markings of any sort to account for chain formation. Third, it follows from a principled and independently motivated definition of Non-Distinctiveness of copies.

#### 4. Constraints on LATAR

Any implementation of LATAR (e.g., either assuming Indexical-S or Inclusion-S) involves a rather powerful set of derivational mechanisms that should be constrained. These constraints must be capable of explaining, for instance, why LATAR cannot be applied in a case like (1b), a sentence where the complement of a noun does trigger a Condition C violation.

Takahashi & Hulsey (2009) propose two types of restriction on LATAR: (i) Agreement/Case, and (ii) semantic interpretability. Regarding agreement/Case, the constraint follows straightforwardly from standard assumptions on the structure of DPs. A bare D is nothing but a determiner without an NP complement. As such, bare Ds lack some features that are inherently valued inside the NP projection (e.g., Number, Gender), so they do not carry complete sets of  $\varphi$ -features. The immediate prediction is that bare Ds are not capable of valuing the  $\varphi$ -features on a Case-assigning head and cannot receive Case. The scenario is sketched in (19), where a T head cannot value its features with the complement of the verb.

$$(19) \quad [TP \ T_{\{<\varphi, \_\>\}} \dots [VP \ V \ D_{\{<\kappa, \_\>, \dots\}}]]$$

In order to satisfy T's requirements, it is necessary introducing a full DP with a complete set of  $\varphi$ -features in the checking domain of T (i.e., as late as Spec,T).

$$(20) \quad [TP \ DP_{\{<\varphi, 3P.SG>, <\kappa, NOM>, \dots\}} [T' \ T_{\{<\varphi, 3P.SG>\}} \dots [XP \ D_{\{<\kappa, \_\>, \dots\}}]]]$$

Such a restriction on the insertion of full DPs may be postulated as in (21).

$$(21) \quad \text{CONSTRAINT A}$$

Only unpronounced Caseless links in non-trivial chains (i.e., “traces” of A-movement) may be bare Ds.

Constraint A in (21) is meant to explain the well-known asymmetry between in A and A'-movement regarding reconstruction: traces of A-movement reconstruct in some cases; traces of A'-movement always reconstruct (cf. Takahashi 2010). Given that traces of A'-movement are marked with Case, they are generally expected to consist of full DPs, so reconstruction is predicted.

Regarding interpretability, Takahashi & Hulsey follow Fox (2002) on assuming that *Trace Conversion* applies at LF to the lowest link in an A'-chain to obtain a valid operator-variable relation under Copy Theory.

$$(22) \quad \text{Trace Conversion (Fox 2002: 67)}$$

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

The rule in (22) transforms a wh-phrase into a definite description with anaphoric value. The subpart in (22a) involves introducing a predicate  $<e,t>$  that functions as a variable, and is interpreted compositionally with a complete nominal predicate, another  $<e,t>$  expression, through *Predicate Modification* (cf. Heim & Kratzer 1998). Incomplete nominal predicates (i.e., nouns lacking some argument) are not  $<e,t>$  expressions, so they are not proper inputs for

Trace Conversion. Therefore, the lowest member of an A'-chain must always contain a noun with all its arguments. Consequently, the statement in (23) follows.

(23) *CONSTRAINT B*

Unpronounced traces of A'-movement must satisfy the selectional requirements of the element heading their semantic restrictors.<sup>4</sup>

This constraint is meant to derive that anti-reconstruction effects in A'-movement are restricted to non-arguments of nominal predicates (e.g., Lebeaux 1988, Fox 1999). For example, the pattern in (1) is explained straightforwardly: (1a) is acceptable because *that John made* is an adjunct to the noun *argument*, so it is not required in the lower position in the A'-chain; on the other side, (1b) is unacceptable because *that John is a genius* is a complement of the noun *argument*, so it must appear in the lower link of the A'-chain.

Trace Conversion also allows ruling out some unwanted consequences of Inclusion-S. According to (15), two DPs can form a chain if their features are in an inclusion relation. Consider a case as (24): if a DP as *which girl* is externally merged in the  $\theta$ -position, and a different DP as *which woman* is base-generated in Spec,C, they would be expected to form the chain in (24d) because of their morpho-syntactic features.

- (24) a. Which woman did Mary meet which girl?  
 b. [CP Which woman{<<sub>K</sub>,ACC>,<<sub>QK</sub>,ACC>,<<sub>Q
 c. [CP Which woman{ACC, Q, ...} ... [VP which girl{ACC, ...}]]  
 d. CH = ([DP which woman], [DP which girl])</sub>

This unwanted result is ruled-out because the output of Trace Conversion would be uninterpretable: there is a lack of correspondence between the predicates *woman* and *girl* that cancels the anaphoric interpretation of the definite description.

- (25) [Which woman]  $\lambda x$ . Mary met the girl  $x$   
 ??For *which woman*  $x$ , Mary met the girl  $x$

---

<sup>4</sup> This is, basically, the same constraint proposed by Lebeaux (1988).

If elliptical sites and traces of wh-movement form a natural class (i.e., if they are variables), then it should be expected that both of them are constrained by similar parallelism conditions. If some of these conditions are semantic, then it is necessary to consider them as *complementing* Inclusion-S at LF.

## 5. Conclusion

Anti-reconstruction effects are evidence for the non-isomorphism between members of a movement chain. In this sense, Inclusion-S is a pretty straightforward way of capturing such a lack of isomorphism. Moreover, adopting Inclusion-S as definition of Non-Distinctiveness allows getting rid of counter-cyclic operations as Late Merger and Wholesale Late Merger, a very welcome result from a conceptual point of view.

## References

- Bošković, Ž. & Nunes, J. 2007. The copy theory of movement: A view from PF. In N. Corver & J. Nunes (eds), *The Copy Theory of Movement*. Amsterdam: John Benjamins.
- Chomsky, N. 1993. A minimalist program for linguistic theory. In K. Hale & S.J. Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*. Cambridge: MIT Press.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. 2000. Minimalist inquiries: The framework. In R. Martin, D. Michaels & J. Uriagereka (eds), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. Cambridge, MA: MIT Press.
- Chomsky, N. 2001. Derivation by Phase. M. Kenstowicz (ed.), *Ken Hale. A Life in Language*. 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.
- Halle, M, & Marantz, A. 1993. Distributed Morphology and the pieces of inflection. In K. Hale & S.J. Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*. Cambridge: MIT Press.
- 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.
- Muñoz Pérez, C. 2016. Three explanatory challenges for Copy Theory. In K. Bellamy, E. Karvovskaya, M. Kohlberger & G. Saad (eds.), *ConSOLE XXIII: Proceedings of the 23rd Conference of the Student Organization of Linguistics in Europe (7-9 January 2015, Paris)*. Leiden: Leiden University Centre for Linguistics.
- Neeleman, A. & H. van De Koot. 2010. A local encoding of syntactic dependencies and its consequences for the theory of movement. *Syntax* 13. 331-372.
- Nunes, J. 1995. *The copy theory of movement and the linearization of chains in the minimalist program*. Doctoral Dissertation, University of Maryland.
- Nunes, J. 2004. *Linearization of chains and sideward movement*. Cambridge, MA: MIT Press.
- Stanton, J. To appear. Wholesale Late Merger in A'-movement: evidence from preposition stranding. *Linguistic Inquiry* 47.
- Takahashi, S. 2010. 'Traces or Copies, or Both – Part I: Characterizing Movement Properties. *Language and Linguistics Compass* 4. 1091–1103.
- Takahashi, S. & Hulsey, S. 2009. Wholesale Late Merger: Beyond the A/A' Distinction. *Linguistic Inquiry* 40. 487-526.