

# The Copy Theory

Jairo Nunes (Universidade de São Paulo)

## 1. Introduction \*

One of the major goals of the generative enterprise has been to capture the so-called “displacement” property of human languages – the fact syntactic constituents may be interpretively associated with positions different from the ones where they are phonetically realized. The GB model captured this property in terms of the trace theory, according to which a movement operation displacing a given element from one structural position to another leaves behind a coindexed trace. Under this view, traces are conceived of as phonetically unrealized categories that inherit the relevant interpretive properties of the moved element, forming with it a discontinuous object – a (nontrivial) chain. The derivation of a sentence such as (1), for instance, is as represented in (2), where *John* moves from the object to the subject position.

(1) John was arrested.

(2) [John<sub>i</sub> was arrested t<sub>i</sub>]

---

\* I would like to thank CNPq (302262/2008-3) and FAPESP (2006/00965-2) for their support during the writing of this chapter.

With the wholesale conceptual evaluation of the GB apparatus that arose with the emergence of the Minimalist Program (Chomsky 1993, 1995), the trace theory of movement became an obvious topic to be examined under minimalist lenses. Chomsky (1993) argues that the trace theory should actually be abandoned in favor of an earlier interpretation of movement as copying. More specifically, he proposes that a movement operation leaves behind a copy of the moved element which gets deleted in the phonological component (in case of overt movement). From this perspective, the derivation of (1) proceeds along the lines of (3), where the crossed material represents lack of phonetic realization at PF.

- (3) a. [was arrested John]
- b. *Copy:*
- John [was arrested John]
- c. *Merge:*
- [John was arrested John]
- d. *Delete:*
- [John was arrested ~~John~~]

Although Chomsky's arguments for incorporating the copy theory of movement into the minimalist framework were largely architectural in nature, the copy theory made it possible to analyze recalcitrant data and opened new avenues to analyze the syntax-PF mapping. By presenting us with a clear case where the simplification of the theoretical apparatus has led to significant gains in empirical coverage, the copy theory has become one of the most stable pillars of minimalist theorizing. As a consequence of its solid status within the framework, the copy theory has also been employed to

empirically distinguish between competing minimalist analyses, as is the case of the PRO-based and movement-based approaches to obligatory control.

This chapter reviews the general features of the copy theory of movement, focusing on some empirical gains prompted by its adoption in the Minimalist Program. The chapter is organized as follows. In section 2 I present Chomsky's (1993) original conceptual arguments for reinterpreting traces as copies. In section 3, I discuss some of the questions that the copy theory pose to the syntax-PF mapping and present Nunes's (1995, 1999, 2004) approach to phonetic realization of copies in terms of linearization and economy computations. In section 4, I present an overview of the kinds of empirical material that may receive a natural account within the copy theory, but remain rather mysterious within the trace theory. In section 5, I discuss how the copy theory makes different predictions for the current debate on how to analyze obligatory control within minimalism. Finally, some concluding remarks are presented in section 6.

## **2. Conceptual Motivations for the Copy Theory**

Sentences such as (4a) and (4b) below are transparent examples of the displacement property of human languages. In (4a) the anaphor requires being interpreted in the object position in order to be c-commanded by the subject, whereas the idiom chunk *the shit* in (4b) must be interpreted in [Spec,vP] in order to form a constituent with *hit the fan*. The representation of these sentences under the trace theory given in (5) must therefore be supplemented with extra assumptions in order to ensure that the required configurations obtain at the point where interpretation applies.

- (4) a. Which picture of himself did John see?  
 b. The shit will hit the fan.

- (5) a. [[which picture of himself]<sub>i</sub> did John see t<sub>i</sub>]  
 b. [[the shit]<sub>i</sub> will [t<sub>i</sub> hit the fan]]

Within GB, two main lines of inquiry were explored to account for facts like these. The interpretation of the anaphor and the idiom was taken to be computed at a level prior to movement (D-Structure) or an operation of reconstruction applying in the LF component was employed to restore the moved material back in its original position. Chomsky (1993) observes that both alternatives are suspect from a minimalist perspective. The first alternative is at odds with one of the major tenets of minimalist downsizing, which is the elimination of non-interface levels of representation such as D-Structure or S-Structure. The second approach circumvents this problem by computing the relevant relations at LF, but at the price of invoking a lowering application of movement (the reconstruction operation).

As Chomsky points out, the copy theory remedies both of these problems. If the sentences in (4) are to be associated to the structures in (6) below, the expected interpretation can be computed at LF via the lower copies. Crucially, the lower copies are deleted in the phonological component, but are available for interpretation at LF.

- (6) a. [[which picture of himself] did John see [~~which picture of himself~~]]  
 b. [[the shit] will [~~the shit~~ hit [the fan]]]

Another important conceptual advantage of the copy theory over the trace theory has to do with the Inclusiveness Condition. Chomsky (1995:228) has proposed that the mapping from the numeration to LF should be subject to an Inclusiveness Condition

requiring that an LF object be built from the features of the lexical items of the numeration. The Inclusiveness Condition can be viewed as a metatheoretical condition that ensures internal coherence within the model. First, it restricts the reference set of derivations that can be compared for economy purposes. If the system could add material that is not present in the numeration in the course of syntactic computations, the role of the numeration in determining the class of comparable derivations would be completely undermined. Thus, given the minimalist assumption that economy matters in the computations from the numeration to LF, something like the Inclusiveness Condition must be enforced in the system. The second important role played by the Inclusiveness Condition is that it ensures that the inventory of syntactic primitives is kept to a minimum, by preventing the syntactic component from creating objects that cannot be defined in terms of the atoms that feed the derivation.

Given this general picture, it is clear that the trace theory within GB was ripe for a minimalist reanalysis, as it is flagrantly incompatible with the underpinnings of the Inclusiveness Condition. Traces are not part of the initial array, but are introduced in the course of the computation. Besides, they are taken to be independent grammatical formatives, with their own properties and licensing conditions. In comparison, the copy theory provides a much more congenial way for movement operations to comply with the Inclusiveness Condition. First, a copy is not a new grammatical formative; it is either a lexical item or a phrase built from lexical items.<sup>1</sup> Second, the copies are built

---

<sup>1</sup> Due to space limitations, here I will not discuss Chomsky's (1995) Move-F approach, according to which the syntactic component can also move/copy the set of formal features of a given lexical item (see Aoun and Nunes 2007 for a comparison with the Agree-approach). However, it should be pointed out that nothing substantially changes

from the material that was present in the numeration. Note that is not the introduction of objects in the course of the derivation *per se* that is problematic, for both traces and copies are introduced in this way. The difference is that the operation Copy, like the structure-building operation Merge, creates an object by manipulating material that is available in the numeration, thus permitting a simple formulation of the reference set for economy computations. By contrast, under the trace theory traces pop up as completely new elements in the computation, thereby requiring that the reference set be further specified with respect to which new elements can or cannot be introduced in the computation out of the blue.

At first sight, this way of satisfying the Inclusiveness Condition may look too costly, as it seems to require the introduction of two operations in the system – Copy and Delete (cf. (3)). Appearances are misleading, though. These operations are in fact independently motivated. Delete, for example, must be invoked in the derivation of ellipsis constructions,<sup>2</sup> regardless of whether it is interpreted as an erasure operation in lexicalist approaches or as a blockage to late insertion in approaches based on Distributed Morphology.<sup>3</sup> As for Copy, standard cases of morphological reduplication provide evidence of its effects elsewhere in the system and as we will see in section 4 below, we may also find unequivocal reflexes of its applications in the syntactic component when more than one chain link gets pronounced. But postponing the

---

regarding the Inclusiveness Condition if the Move-F approach is correct, as the copied set of formal features is a replica of features present in the numeration.

<sup>2</sup> In fact, when Chomsky (1993) proposed adopting the copy theory, he suggested that deletion of traces could be related to deletion in ellipsis constructions (see e.g. Nunes 2004 and Saab 2008 for relevant discussion).

<sup>3</sup> See Saab 2008 for recent illuminating discussion.

presentation of this empirical evidence until section 4, it is worth observing at this point that a Copy-like operation must also be independently resorted to in the mapping from the lexicon to the numeration/derivation.<sup>4</sup> After all, when we say we take an item from the lexicon to form a given numeration, we definitely do not mean that the lexicon has lost one item and is now smaller. Rather, we tacitly assume that numerations are formed by *copying* items from the lexicon.

To summarize. In addition to conforming to the Inclusiveness Condition, the copy theory considerably simplifies the analysis of reconstruction phenomena. Furthermore, by making it possible to treat reconstruction as an LF phenomenon, the copy theory contributes to the attempt to eliminate non-interface levels of representation. Finally, by eliminating traces *qua* grammatical formatives, it reduces the number of theoretical primitives in our inventory: if traces are copies, they are either lexical items or complex objects built from lexical items (see fn. 1).<sup>5</sup>

---

<sup>4</sup> See Hornstein 2001 for discussion of this point.

<sup>5</sup> One question that arises in any version of the copy theory of movement is how the computational system distinguishes copies from elements that happen to have the same set of features. The derivation of (i) below, for instance, should converge if it starts with the numeration  $N_1$  in (iia) (with one instance of *Mary*), but not with the numeration  $N_2$  in (iib) (with two instances).

(i) [Mary [was [hired Mary]]]

(ii) a.  $N_1 = \{\text{Mary}_1, \text{was}_1, \text{hired}_1, \dots\}$

b.  $N_2 = \{\text{Mary}_2, \text{was}_1, \text{hired}_1, \dots\}$

### 3. The Copy Theory and the Syntax-PF Mapping

In his reanalysis of reconstruction in terms of the copy theory, Chomsky (1993) argues that there is more than one possibility for the interpretive systems to read LF objects. He proposes that the ambiguity of a sentence like (7) below, for instance, is due to the different parts of the *wh*-chain that the interpretive systems may compute. If the whole *wh*-phrase is computed upstairs after deletion of the lower copy, as represented in (8a) with the outlined material annotating lack of interpretation at LF, we obtain the reading under which the anaphor is bound by the matrix subject. By contrast, if only the *wh*-

---

Chomsky (1995:227) suggests that two lexical items *l* and *l'* should be marked as distinct if they enter the derivation via different applications of Select. Alternatively, Nunes (1995, 2004) suggests that it is in fact the operation Copy that assigns a nondistinctiveness index; in other words, all elements will be taken to be distinct for purposes of the computational system, unless they are specified as nondistinct by the Copy operation. It is worth pointing out that both suggestions run afoul of the Inclusiveness Condition, as the added (non)distinctiveness markings are not part of the numeration. A possibility that is in consonance with the Inclusiveness Condition (see Nunes 2004:165) is to allow the system to compute (non)distinctiveness by comparing derivational steps. For instance, if two contiguous derivational steps  $\sigma_1$  and  $\sigma_2$  differ in that a new term  $\tau$  is introduced into the computation, two possibilities arise: if from  $\sigma_1$  to  $\sigma_2$  the numeration has been reduced,  $\tau$  is to be interpreted as distinct from all the other syntactic objects available at  $\sigma_2$ ; if the numerations of  $\sigma_1$  and  $\sigma_2$  are the same,  $\tau$  must be a copy of some syntactic object available at  $\sigma_1$ . Whether or not it is desirable that the recognition of copies by the computational system proceeds along these lines remains to be determined.

element is computed upstairs after scattered deletion within the *wh*-chain, as in (8b), we get the embedded subject reading for the anaphor.<sup>6</sup>

(7) John<sub>i</sub> wonders which picture of himself<sub>i/j</sub> Bill<sub>j</sub> saw

(8) a. John wonders [<sub>CP</sub> [which picture of himself] [<sub>IP</sub> Bill saw [which picture of himself]]]

b. John wonders [<sub>CP</sub> [which picture of himself] [<sub>IP</sub> Bill saw [which picture of himself]]]

---

<sup>6</sup> Based on the obligatory reconstruction in sentences such as (ia) below, Chomsky (1993) proposes that in the case of A'-chains, there is actually an economy preference for minimizing operator restrictions in LF, which normally leads to scattered deletion (cf. (ib)/(8b)). To force reconstruction in (ia) while allowing the upstairs reading of *himself* in (7), Chomsky suggests that either the higher or the lower copy of *himself* undergoes anaphor movement covertly. When the lower copy of *himself* moves, deletion along the lines of (8b), which complies with this preference principle, yields a well formed result. By contrast, if the higher copy undergoes anaphor movement, scattered deletion as in (8b) would “break” the anaphor chain, causing the derivation to crash. The system is then allowed to employ deletion as in (8a), for only convergent derivations can compete for purposes of economy.

(i) a. \*Mary wondered which picture of Tom<sub>i</sub> he<sub>i</sub> liked.

b. \*Mary wondered [<sub>CP</sub> [which picture of Tom<sub>i</sub>] he<sub>i</sub> liked [which picture of Tom<sub>i</sub>]]

In addition to these two interpretive alternatives, there also arises the logical possibility that only a lower link gets interpreted after deletion of the higher copy, as is arguably the case of reconstruction of idiom chunks in (9) (cf. (4b)).

(9) ~~[[the shit]]~~ will ~~[[the shit]]~~ hit the fan]]

On the LF side we thus have some choice in deciding how deletion should take place in nontrivial chains (see fn. 6). That being so, questions arise with respect to the syntax-PF mapping. Take the structure in (10) below, for instance, which has been formed after the object DP moved to the subject position. One wonders why the only well formed PF output for (10) is (11a), where only the highest copy is pronounced. The ungrammaticality of (11b), with no deletion, is particularly interesting as it is the most transparent output with respect to the structure that reaches LF.

(10) ~~[[the student]]~~ was arrested ~~[[the student]]~~

(11) *PF outputs:*

- a. ~~[[the student]]~~ was arrested ~~[[the student]]~~
- b. \*~~[[the student]]~~ was arrested ~~[[the student]]~~
- c. \*~~[[the student]]~~ was arrested ~~[[the student]]~~
- d. \*~~[[the student]]~~ was arrested ~~[[the student]]~~

At first sight, the trace theory fares better in this regard as it need not concern with potentially different phonetic outputs for a given nontrivial chain. However, appearances are again illusory. The issue of phonetic realization of chains is mute

within the trace theory because traces are *stipulated* to be phonetically null. A truly explanatory account of movement in terms of traces should provide an appropriate answer for why traces are necessarily devoid of phonetic content. In other words, even under the trace theory, the fact that that only chain heads are phonetically realized is something in need of an explanation. Upon close inspection, this conundrum may in fact provide the copy theory with the upper hand on this issue. Consider why.

In section 2, we discussed conceptual reasons for why the copy theory was a better alternative than the trace theory, given minimalist considerations regarding the mapping to LF. Recall that the argument was not that the trace theory was empirically flawed, but that it required additional assumptions that did not fit snugly within the general architectural features of minimalism. By contrast, the mapping from the syntactic component to PF may offer a deadly empirical argument against the trace theory. Suppose, for instance, that some constructions (in some languages) may allow pronunciation of more than one copy, pronunciation of a lower copy, or scattered deletion within a chain in a way analogous to (11b-d). If such cases do exist, we will then have a very powerful argument for choosing the copy theory over the trace theory. The latter has no room to accommodate facts like these, for it is a defining property of traces that they are phonetically empty.

Section 4 below is devoted to showing that we do indeed find PF outputs parallel to (11b-d). But before we examine such cases, let us first discuss why the PF output represented in (11a), where only the head of the chain is pronounced, is by far the most common pattern found across languages. It is productive to break the puzzle in two different questions: (i) why is it the case that in general a chain cannot surface with all of its links phonetically realized (cf. (11b))?; and (ii) why is it the case that full

pronunciation of the highest copy is in general the only grammatical PF output (cf. (11a) vs. (11c-d))?

By combining a convergence requirement with economy considerations, Nunes (1995, 1999, 2004) provides a general answer to these questions that is able to account for both the standard option illustrated in (11a) and exceptional cases that parallel the options in (11b-d). The convergence ingredient is related to linearization at PF. The gist of the proposal is that copies count as “the same” for purposes of linearization because they are nondistinct elements (i.e. they relate to the same occurrences of lexical items of the numeration; see fn. 5) and this creates problems. Take the structure in (10), for example. Given that the higher occurrence of [*the student*] asymmetrically c-commands *was*, Kayne’s (1994) LCA dictates that both *the* and *student* should precede *was*. Likewise, given that *was* asymmetrically c-commands the lower occurrence of [*the student*], it should precede *the* and *book* in compliance with the LCA. Given that these occurrences of [*the book*] are nondistinct, we reach a contradiction: *was* should precede and be preceded by *the* and *student*. Similarly, given that the higher occurrence of [*the student*] asymmetrically c-commands the lower one, we obtain the undesirable result that *the*, for instance, should be required to precede itself.

To make the same point in a slightly different way, the fact that a nontrivial chain is a discontinuous object that simultaneously occupies different structural positions in the syntactic structure creates an impasse for linearization. On the one hand, a chain cannot be assigned a single slot in the PF linear sequence resulting from the LCA, for it is associated with more than one structural position; on the other hand, the assignment of multiple slots should create contradictory requirements, preventing the whole structure from being linearized. Thus, the reason why a chain cannot (in general)

surface with all of its links phonetically realized (cf. (11b)) under this view is that the structure containing it cannot be linearized.

Nunes (1995, 1999, 2004) argues that deletion comes into play in this scenario as a rescuing strategy to permit the linearization of structures containing chains. More specifically, deletion of the “repeated” material within chains before linearization (*Chain Reduction* in Nunes’s terms) circumvents the problem of linearizing *was* with respect to *the* and *student* in (10). If the material of the chain  $CH = ([the\ student], [the\ student])$  is deleted in any of the ways depicted in (11a,c,d), the structure in (10) can be linearized without any problems. The question now is why only the deletion sketched in (11a) yields an acceptable sentence.

This is the point where economy plays a crucial role. More specifically, economy considerations should ensure that deletion applies as few times as possible. Applying to the DP chain in (10), Chain Reduction may yield the output in (11d), with two applications of deletion, or the outputs in (11a) and (11c), with a single application targeting the whole DP node. Once there is arguably no convergence problem resulting from these reductions, the three derivations are eligible for economy comparison and the derivation yielding (11d) is excluded for employing more operations of deletion than necessary. What is now missing is an explanation for why the actual reduction of the DP chain in (10) must involve the deletion of the lower copy, rather than the head of the chain (cf. (11a) vs. (11c)), despite the fact that both reductions may employ a single operation of deletion targeting the whole DP node. Obviously, we cannot simply say that lower copies must delete. Conceptually, that would amount to reintroducing traces and empirically, it would be just wrong, as we will see in section 4.

The most plausible answer should again be formulated in economy terms<sup>7</sup>. There should be some factor that makes the pronunciation of the highest copy more economical in the general case. One possibility is that such independent factor is feature checking/valuation (see Nunes 1995, 1999, 2004). If the highest copy always has more features checked/valued than the lower copies, it should be the optimal candidate for phonetic realization. Actual implementation of this idea depends on specific assumptions regarding the inner workings of feature checking/valuation and the relation among copies when one of them undergoes feature checking/valuation.<sup>8</sup>

For concreteness, I will here assume Bošković's (2007) proposal that a given element can only have its uninterpretable features valued if it acts as a probe. Under this view, the derivation of (10) proceeds along the lines of (12) below. Given (12a), T probes the structure and has its  $\phi$ -features valued by agreeing with [*the student*], yielding (12b). In order to have its Case-feature valued, the internal argument then moves to [Spec,TP] and from this position, it probes T and values its Case-feature, yielding (12c).

- (12) a. [ $T_{\phi:?}$  be arrested [*the student*]<sub>Case:?</sub>]  
       b. [ $T_{\phi:3SG}$  be arrested [*the student*]<sub>Case:?</sub>]  
       c. [[*the student*]<sub>Case:NOM</sub>  $T_{\phi:3SG}$  be arrested [*the student*]<sub>Case:?</sub>]

The structure in (12c) arguably causes the derivation to crash at LF, as the lower copy does not have its Case feature valued. Let us then assume that once a given element has

---

<sup>7</sup> See Franks 1998 for the seeds of the economy approach to be explored below.

<sup>8</sup> For relevant discussion and alternatives, see e.g. Nunes 1995, 1999, 2004, Kobele 2006, and Bošković and Nunes 2007.

its uninterpretable features valued, it is allowed to probe the structure again and value the features of its lower copies (in a domino fashion if more than one copy is involved). If so, before the structure in (12c) is shipped to the C-I interface, the upper copy values the Case-feature of the lower copy, yielding (13) below. This suggestion captures in a derivational fashion Chomsky's (1995:381, n. 12) proposal that "the features of chain are considered a unit: if one is affected by an operation, all are".

(13) [[the student]<sub>Case:NOM</sub> T<sub>φ:3SG</sub> be arrested [the student]<sub>Case:NOM</sub>]

We now have all the ingredients we need. At the point where the structure in (12c) is assembled, probing by the upper copy is not required for PF purposes. If no such probing is forced to apply before Spell-Out, the structure in (12c) is shipped to the phonological component as is and the probing yielding (13) takes place after Spell-Out. In the phonological component, the choice for pronunciation between (11a) and (11c) is then determined by the copy that has its features valued (cf. (12c)), as this information is needed by Morphology. Thus, (11a) trumps (11c).

To sum up. The combination of a convergence requirement in terms of linearization and economy considerations regarding the number of applications of deletion provides an account for why a chain (in general) does not surface with all of its links phonetically realized (the structure containing such a chain cannot be linearized) and why scattered deletion constructions are uncommon (they employ an unnecessary number of applications of deletion). Finally, an independent asymmetry among copies due to feature checking/valuation establishes a specific economy metric that favors deletion of lower copies.

For the sake of completeness, let us consider how this proposal accounts for the standard output of remnant movement constructions (see section 4.3 below for further discussion). Take the derivation of (14) below, for instance, whose representations under the trace theory and the copy theory are given in (15). The interesting thing about the representation in (15b) (see Gärtner 1998) is that the leftmost copy of *John* gets deleted despite the fact that it doesn't form a chain with either of the other copies (it neither c-commands nor is c-commanded by the other copies, for instance). This potential problem for the copy theory is in fact analogous to the one faced by the trace theory in accounting for how  $t_i$  in (15a) is not c-commanded by its antecedent.

(14) ... and elected, John was.

- (15) a. ... and  $[_{XP} [\text{elected } t_i]_k [_{X'} X [_{TP} \text{John}_i [_{T'} \text{was } t_k]]]]$   
 b. ... and  $[_{XP} [\text{elected } \text{John}] [_{X'} X [_{TP} \text{John} [_{T'} \text{was } \text{elected John}]]]]$

Within the copy theory, there are two possible approaches to this issue. A more representational answer is offered in Nunes (2003, 2004), building on Chomsky's (1995:300) observation that the representation of a chain such as  $CH = (\alpha, \alpha)$  should be seen as a notational abbreviation of  $CH = ((\alpha, K), (\alpha, L))$ , where K and L are each the sister of one occurrence of  $\alpha$ . In other words, the individual links of a chain must be identified not only in terms of their content, but also in terms of their local structural configuration. Hence, movement of *John* in (15b) first forms the chain  $CH_1 = ((\text{John}, T'), (\text{John}, \text{elected}))$  and movement of the remnant VP then forms the chain  $CH_2 = (([\text{elected John}], X'), ([\text{elected John}], \text{was}))$ . Under the assumption that Spell-Out ships the whole structure in (15b) to the phonological component, Chain Reduction inspects

CH<sub>1</sub> and instructs the phonological component to delete the occurrence of *John* that is the sister of *elected*. Interestingly, there are two elements in (15b) that satisfy this description, namely, the leftmost and the rightmost copies of *John*. In fact, these two copies are technically identical: they are nondistinct in terms of the initial numeration, they have participated in no checking relations, and their sisters are nondistinct. Assuming that the phonological component blindly scans the structure to carry out the deletion instructed by Chain Reduction, it ends up deleting the two copies that satisfy the instruction, as represented in (15b); Chain Reduction of CH<sub>2</sub> then deletes the lower copy of VP and the sentence in (14) is derived.

Under a more derivational approach (see Bošković and Nunes 2007), linearization/Chain Reduction applies as the phonological component is fed with spell-out units. From this perspective, the system spells out TP after the structure in (16a) below is built and Chain Reduction deletes the lower copy of *John*. From this point on, the copy of *John* in the object position will be unavailable to any operation of the phonological component. Hence, movement of VP later on in the derivation, as shown in (16b), will be oblivious of this copy. After the whole structure in (16c) is spelled out, deletion of the lower VP copy then yields the sentence in (14).

- (16) a. [<sub>XP</sub> X [<sub>TP</sub> John [<sub>T</sub> was [<sub>VP</sub> elected ~~John~~]]]]  
 b. [<sub>XP</sub> [<sub>VP</sub> elected ~~John~~] [<sub>X'</sub> X [<sub>TP</sub> John [<sub>T</sub> was [<sub>VP</sub> elected ~~John~~]]]]]  
 c. [<sub>XP</sub> [<sub>VP</sub> elected ~~John~~] [<sub>X'</sub> X [<sub>TP</sub> John [<sub>T</sub> was [<sub>VP</sub> elected ~~John~~]]]]]

This is not the place to decide between these alternatives. For our purposes, it suffices that both of them correctly enforce deletion of traces in standard remnant

movement constructions and can also handle the remnant movement constructions involving multiple copies to be discussed in section 4.3 below.

#### **4. Empirical Payoff**

As mentioned in section 3, the most powerful argument for the copy theory should come from the mapping from Spell-Out to PF. If lower copies can somehow be pronounced, we will have a knock-out argument for the copy theory, for under the trace theory traces are phonetically null by definition.

Below we consider several cases that instantiate the possibility that lower copies can be pronounced.<sup>9</sup>

##### *4.1. Phonetic realization of a lower copy*

In section 3, the preference for pronouncing chain heads was taken to follow from an economy condition. Given the derivation sketched in (17) below, where  $\alpha$  moves to value its uninterpretable feature F, the structure in (17b) can be shipped to the phonological component without the additional valuation of the lower copy by the higher one. Such valuation, as sketched in (18), is only required for LF purposes. Once the additional valuation is not required to apply before Spell-Out, (local) economy prevents it from doing so and Spell-Out applies to (17b), yielding the preference for deleting lower copies.

---

<sup>9</sup> For additional examples and general discussion, see e.g. Nunes 1999, 2004, Bošković 2001, Bošković and Nunes 2007, Saab 2008, Kandybowicz 2008, the collection of papers in Corver and Nunes 2007, and references therein.

- (17) a. [H [ ...  $\alpha_{F:?$  ... ]]  
 b. [ $\alpha_{F:\vee}$  H [ ...  $\alpha_{F:?$  ... ]]

- (18) [ $\alpha_{F:\vee}$  H [ ...  $\alpha_{F:\vee}$  ... ]]

Now suppose that in a given derivation, independent convergence requirements of the phonological component ban the pronunciation of the higher copy of  $\alpha$  in (17b). In such circumstances, the system will then be forced to trigger valuation in (18) before Spell-Out in order to ensure convergence. Once (18) is shipped to the phonological component, each copy has its features valued and is eligible for pronunciation. But if the higher copy violates wellformedness conditions of the phonological component, it should be deleted and the lower copy should be pronounced instead, as sketched in (19).

- (19) [ ~~$\alpha_{F:?$~~  H [ ...  $\alpha_{F:\vee}$  ... ]]

The scenario depicted above where a lower copy is pronounced instead of the head of the chain has been increasingly documented in the literature (see fn. 9).

Consider the contrast between (20) and (21) below, for instance. (20) illustrates the well-known fact that Romanian is a multiple *wh*-fronting language; hence the unacceptability of the *wh-in situ* in (20b). (21), on the other hand, seems to be an exception to the paradigm illustrated in (20), in that a *wh*-element *in situ* is allowed.

- (20) *Romanian*:

- a. Cine ce precede?  
*who what precedes*

- b. \*Cine precede ce?  
*who precedes what*  
 ‘Who precedes what?’

(21) *Romanian*:

- a. \*Ce ce precede?  
*what what precedes*
- b. Ce precede ce?  
*what precedes what*  
 ‘What precedes what?’

Bošković (2002) however argues that the appearances here are deceiving. The unacceptability of (21a) is related to a restriction in the phonological component prohibiting adjacent occurrences of *ce* ‘what’. That is, from a syntactic point of view, there is no difference between (20) and (21); we have multiple *wh*-fronting in both cases. It just happens that if the higher copy of the moved object of (21) is realized, it will violate this ban on adjacent identical words, which is found in several languages.<sup>10</sup> The phonological system then deletes the higher copy of the object *ce* ‘what’, as sketched in (22) below, allowing the structure to both be linearized and comply with this adjacency restriction. Bošković provides independent evidence for the deletion sketched in (22) by showing that the object in (21b) patterns like moved *wh*-objects in being able to license a parasitic gap, as shown in (23), something that a truly *in situ wh*-object cannot do.

---

<sup>10</sup> See Golston (1995) for a discussion of many such cases and Richards (2006) for some related issues.

(22) [<sub>ce<sub>SU</sub></sub> [<sub>ce<sub>OB</sub></sub> [<sub>ce<sub>SU</sub></sub> precede ce<sub>OB</sub>]]]

(23) *Romanian*

Ce precede ce fara sa influenteze?  
*what precedes what without SUBF.PRT influence.3.SG*  
 ‘What precedes what<sub>i</sub> without influencing it<sub>i</sub>?’

Another interesting argument for pronunciation of lower copies is provided by Bobaljik’s (1995) account of Holmberg’s (1986) Generalization (see also Bobaljik 2002). Holmberg (1986) has observed that object shift in Scandinavian can take place in matrix main verb V-2 clauses, but not in auxiliary+participle clauses or embedded clauses, which do not involve main verb movement. This can be seen in (24), where *ekki* ‘not’ is taken to mark the VP boundary.

- (24) a. Í gær máluðu stúdentarnir húsið<sub>i</sub> [<sub>VP ekki t<sub>i</sub></sub>] (*Icelandic*)  
 yesterday painted the-students the-house not  
 ‘The students didn’t paint the house yesterday.’
- b. \*at Peter den<sub>i</sub> [<sub>VP læste t<sub>i</sub></sub>] (*Danish*)  
 that Peter it read
- c. at Peter [<sub>VP læste den</sub>]  
 that Peter read it  
 ‘that Peter read it.’
- d. \*Hann hefur bókina<sub>i</sub> [<sub>VP lesið t<sub>i</sub></sub>] (*Icelandic*)  
 he has the-book read

- e. Hann hefur [<sub>VP</sub> lesið bókina]  
 he has read the-book  
 ‘He has read the book.’

Bobaljik argues that in clauses in which V-movement does not take place, the relevant Infl head (finite or participial) must be adjacent to the verbal head in order for them to undergo morphological merger after Spell-Out. Thus, obligatory overt movement of (specific, noncontrastive definite) objects with standard pronunciation of the head of the chain disrupts the adjacency between Infl and V and yields an ungrammatical result (cf. (24b)/(24d)). Bobaljik proposes that in these circumstances, the head of the object shift chain is deleted and its tail is pronounced, as sketched in (25) (cf. (24c)/(24e)), which allows morphological merger between Infl and the verb, as they are now adjacent.

- (25) a. [at [<sub>IP</sub> Peter I [<sub>AgroP</sub> ~~den~~ [<sub>VP</sub> læste den]]]]  
 b. [hann hefur [<sub>PartP</sub> Part [<sub>AgroP</sub> ~~bókina~~ [<sub>VP</sub> lesið bókina]]]]

The possibility of pronouncing lower copies due to independent requirements of the phonological components can also account for some interesting facts concerning V-2 in Northern Norwegian, as argued by Bošković (2001). Rice and Svenonius (1998) have observed that the V-2 requirement in Northern Norwegian is stricter than in other Germanic V-2 languages in that the material preceding the verb must minimally contain one foot (i.e. two syllables), as illustrated by the contrast in (26). Rice and Svenonius further note that (26b) can be saved by using the *wh*-subject-V order, as shown in (27).

(26) *Northern Norwegian:*

- a. Korsen kom ho hit?  
 how came she here  
 ‘How did she get here?’
- b. \*Kor kom du fra?  
 where came you from  
 ‘Where did you come from?’

(27) *Northern Norwegian:*

- Kor du kom fra?  
 where you came from  
 ‘Where did you come from?’

Bošković (2001) argues that a uniform analysis in terms of V-to-C movement in the syntactic component can be maintained for (26) and (27) if it is coupled with the possibility of pronouncing lower copies under PF demands. That is, in order to comply with the stricter prosodic requirements of Northern Norwegian regarding V2-constructions, the head of the of verb chain, which is adjoined to C, is deleted and the lower copy in [Spec,TP] is pronounced instead, as illustrated in (28) (cf. (26b) vs. (27)).

(28) [<sub>CP</sub> kor ~~kom~~ [<sub>IP</sub> du kom fra]]

Recall from the discussion above that the preference for pronouncing the head of the chain is ultimately related to an economy condition. If the highest copy does not need to probe the structure to value the features of lower copies before Spell-Out, it

doesn't. In the case of (26b)/(27), such probing was required because otherwise the lower copy could not be pronounced. In the case of the derivation of (26a), on the other hand, no such probing before Spell-Out is needed to ensure convergence at PF. It is therefore blocked from applying and the highest copy of the verb must be pronounced, as illustrated by the contrast between (26a) and (29a).

(29) *Northern Norwegian:*

- a. \*Korsen ho kom hit?  
     how she came here  
     ‘How did she get here?’
- b. \*<sub>[CP</sub> korsen ~~kom~~ <sub>[IP</sub> ho kom hit]]

Let us finally consider one more argument for lower copy pronunciation triggered by PF considerations, this time based on the relationship between word order and stress assignment in a “free” word order language such as Serbo-Croatian. Stjepanović (1999, 2003, 2007) offers a variety of arguments showing that S, V, IO, and DO all move out of VP overtly in Serbo-Croatian. However, a focused element must surface as the most embedded element of the sentence, as illustrated in (30).

(30) *Serbo-Croatian:*

- a. [*Context*: Who is Petar introducing to Marija?]  
     Petar Mariji predstavlja Marka.  
     Petar Marija-DAT introduces Marko-ACC  
     ‘Petar is introducing Marko to Marija.’

b. [*Context*: Who is Petar introducing Marija to?]

Petar Mariju predstavlja Marku.

Petar Marija-ACC introduces Marko-DAT

‘Petar is introducing Marko to Marija.’

In order to reconcile the evidence showing the verb and its arguments leave VP with the position of focused elements, Stjepanović argues that the lower copy of a moved focus element may be pronounced instead of the head of the chain so that it surfaces in a position where it can receive focus stress. Under this view, Chain Reduction in the derivation of the sentences in (30b), for instance, proceeds along the lines sketched in (31).

(31) [S V ~~IO~~ DO [<sub>VP</sub> ~~S~~ ~~V~~ IO ~~DO~~]]

To summarize. Under the specific implementation of the copy theory reviewed here, standard pronunciation of the head of the chain is more economical as it does not employ probing by the highest copy to value the features of the lower copies before Spell-Out. However, not always does the more economical option lead to a convergent result at PF. In such circumstances, the additional probing is required to apply overtly and a lower copy can be pronounced instead. Notice that by relying on economy, we have an account for why pronunciation of the head of a chain is always preferred all things being equal, while also being able to account for the output when things are not equal, that is, when additional convergence requirements of the phonological component block pronunciation of the highest copy and remove this derivational option from the comparison set, allowing pronunciation of a lower copy. As mentioned early,

by stipulating that traces do not have phonetic content, the trace theory is not so fortunate and cannot account for data such as the ones discussed in this section in a principled manner.

#### 4.2. Scattered deletion

Let us now examine what would be necessary for scattered deletion within a chain, as illustrated in (32), to obtain in the phonological component, allowing different parts of different links to be phonetically realized.

(32)  $[[\alpha \beta] H [ \dots [\epsilon \beta] \dots ]]$

The first requirement necessary for such an output to be derived is that the links that surface have their features valued; otherwise, the derivation would crash. That is, given the derivational steps in (33) below, where the constituent  $[\alpha \beta]$  moves to have its feature F valued, the higher copy of  $[\alpha \beta]$  in (33b) must probe the structure and value the feature F of its lower copy, as represented in (34), before Spell-Out.

(33) a.  $[H [ \dots [\alpha \beta]_{F:?} \dots ]]$

b.  $[[\alpha \beta]_{F:\checkmark} H [ \dots [\alpha \beta]_{F:?} \dots ]]$

(34)  $[[\alpha \beta]_{F:\checkmark} H [ \dots [\alpha \beta]_{F:\checkmark} \dots ]]$

As we saw in detail in section 4.1, overt probing by a higher copy to value a lower copy is not an economical option. To derive the output in (32) from the structure in (34), the system must still resort to an additional non-economical route, namely, two

applications of deletion when just one application targeting the upper or the lower link would suffice to allow the structure to be linearized in accordance with the LCA. In other words, the non-economical alternative in (32) can only be an optimal output if neither alternative employing just one application of deletion converges.

A rather persuasive example of this possibility is provided by Bošković's (2001) in his analysis of the contrast between Macedonian and Bulgarian with respect to their surface location of clitics, as illustrated in (35) and (36).

(35) *Macedonian* (Rudin, Kramer, Billings, and Baerman 1999):

- a. Si      mu      (gi)      dal      li      parite?  
       are him-DAT    them    given    Q    the-money
- b. \*Dal    li    si    mu            (gi)    parite?  
       given    Q    are him-DAT    them    the-money  
       ‘Have you given him the money?’

(36) *Bulgarian* (Rudin, Kramer, Billings, and Baerman 1999):

- a. \*Si      mu            (gi)      dal      li      parite?  
       are    him-DAT    them    given    Q    the-money
- b. Dal      li    si    mu            (gi)      parite?  
       given    Q    are him- DAT    them    the-money  
       ‘Have you given him the money?’

Bošković argues that in both languages the complex head [si+mu+gi+dal] left-adjoins to the interrogative particle *li*, leaving a copy behind, as represented in (37) below.

Deletion of the lower copy of [ si+mu+gi+dal ], as shown in (38), yields a well formed

result in Macedonian (cf. (35a)), because in this language pronominal clitics are proclitic and *li* is enclitic. The unacceptability of (35b) then follows from the general ban on scattered deletion imposed by economy considerations regarding the number of applications of deletion. In Bulgarian, on the other hand, *li* as well as the pronominal clitics are enclitics; thus, deletion of the lower copy of the complex head does not lead to a convergent result (see (36a)). Bošković proposes that the system then resorts to scattered deletion, as shown in (39), allowing the chain to be linearized while at the same time satisfying the additional requirements of the phonological component.

(37) [[ si+mu+gi+dal ]+li ... [ si+mu+gi+dal ] ...]

(38) *Macedonian*:

[[ si+mu+gi+dal ]+li ... ~~[ si+mu+gi+dal ]~~ ...]

(39) *Bulgarian*:

[[ ~~si+mu+gi+dal~~ ]+li ... [ si+mu+gi+dal ] ...]

Under the linearization approach reviewed in section 3, the fact that constructions involving scattered deletion are considerably rare follows from their having to resort to non-economical derivational routes in the mapping from the syntactic component to PF. But to the extent that they do exist,<sup>11</sup> they provide very convincing arguments for the copy theory and against the trace theory.

---

<sup>11</sup> For other examples other constructions that are argued to involve scattered deletion, see e.g. Čavar and Fanselow's (1997) analysis of split constructions in Germanic and Slavic languages and Wilder's (1995) analysis of extraposition.

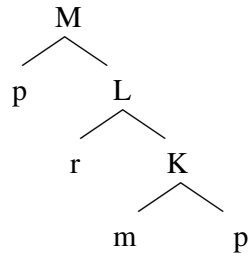
### 4.3. *Phonetic realization of multiple copies*

Let us finally examine cases where different links of a given chain are pronounced, but with the same phonetic material.<sup>12</sup> Recall from section 3 that according to the linearization approach to deletion of copies, a chain cannot surface with more than one link realized with the same phonetic material because the structure containing it cannot be linearized. This line of thinking predicts that if two given copies somehow manage to not interfere with linearization, they should in principle be able to both surface overtly. Nunes (1999, 2004) argues that under certain conditions, this actually happens. Here is the reasoning. Suppose that after the syntactic structure in (40a) below, with two copies of *p*, is spelled out, the morphological component fuses (in the sense Halle and Marantz 1993) the terminals *m* and *p*, yielding the atomic blended terminal #*mp*# (or #*pm*#, for that matter), with no internal structure accessible to further morphological or syntactic computations, as sketched in (40b).

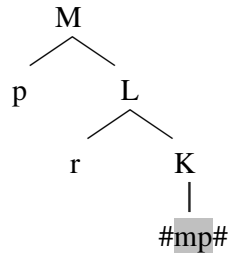
---

<sup>12</sup> Due to space limitations, I will not discuss cases where it has been argued that lower copies are realized as (resumptive) pronouns, reflexives, or partial copies (see e.g. Lidz and Idsardi 1997, Pesetsky 1997, 1998, Hornstein 2001, 2007, Grohmann 2003, Fujii 2007, and Barbiers, Koenenman, and Lakakou 2010 for relevant discussion). For our purposes suffice it to say that to the extent that these lower copies are rendered distinct from the head of the chain, no linearization problem is at stake.

(40) a. *Spelled out structure:*



b. *Fusion in the morphological component:*



The content of #mp# in (40b) cannot be directly linearized with respect to *r* or the upper copy of *p* because it is an inaccessible part of #mp#. From an LCA perspective, for instance, the blended material within #mp# is not accessible to c-command computations. However, it can be indirectly linearized in (40b) in virtue of being an integral part of #mp#: given that the upper copy of *p* asymmetrically c-commands *r* and that *r* asymmetrically c-commands #mp#, we should obtain the linear order  $p > r > \text{\#mp\#}$ . In other words, the material inside #mp# gets linearized in a way analogous to how the phoneme /l/ is indirectly linearized in *John loves Mary* due to its being part of the lexical item *loves*. But, crucially, once the lower copy of *p* in (40b) becomes invisible for standard linearization computations, the linearization problems caused by the presence of multiple copies discussed in section 3 cease to exist. Thus, the structure in (40b) not only can, but must surface with two copies of *p* at PF.

With this in mind, consider verb clefting constructions in Vata, as illustrated in (41) below. Koopman (1984) shows that the two verbal occurrences in (41) cannot be

separated by islands, which indicates that they should be related by movement. The problem from the present perspective is that if these occurrences are to be treated as copies, the structure containing them should not be able to be linearized in accordance with the LCA, as discussed in section 3. Nunes (2004) proposes that this possibility does not in fact arise because the highest copy of the clefted verb gets morphologically fused, thereby evading the purview of the LCA. More precisely, he analyzes verb clefting in *Vata* as involving verb movement to a Focus head, followed by fusion in the morphological component between the moved verb and the Focus head, as represented in (42a). Of the three verbal copies in (42a), the LCA only “sees” the lower two after the highest copy gets fused with  $\text{Foc}^0$ .<sup>13</sup> The lowest copy is then deleted (cf. (42b)) and the structure is linearized as in (41), with two copies of the verb phonetically realized.

(41) *Vata* (Koopman 1984):

*li*    *à*    *li-da*    *zué*    *saká*  
 eat we    eat-PAST    yesterday    rice  
 ‘We ATE rice yesterday’

(42) a. *Fusion*:

$[\text{FocP} \# [\text{Foc}^0 \text{V} [\text{Foc}^0 \text{Foc}^0]] \# [\text{TP} \dots [\text{T}^0 \text{V} [\text{T}^0 \text{T}^0]] [\text{VP} \dots \text{V} \dots]]]$

b. *Deletion of copies*:

$[\text{FocP} \# [\text{Foc}^0 \text{V} [\text{Foc}^0 \text{Foc}^0]] \# [\text{TP} \dots [\text{T}^0 \text{V} [\text{T}^0 \text{T}^0]] [\text{VP} \dots \cancel{\text{V}} \dots]]]$

---

<sup>13</sup> Just to be clear, the point is not that every instance of head movement renders the adjoined element invisible to the LCA, but rather that *fused* elements are not computed by the LCA (cf. (40)).

Nunes (2004) presents two bits of evidence in favor of this account of verb clefting in Vata. The first one relates to Koopman's (1984:158) observation that the restricted set of verbs that cannot undergo clefting in Vata has in common the property that they cannot serve as input for morphological processes that apply to other verbs. If these verbs cannot participate in any morphological process, they certainly should not be able to undergo the morphological fusion with  $\text{Foc}^0$  depicted in (42a) and should not be allowed in predicate clefting constructions. The second piece of evidence is provided by the fact, also observed by Koopman, that the fronted verb in these focus constructions must be morphologically unencumbered; in particular, none of the tense or negative particles that occur with the verb in Infl may appear with the fronted verb, as illustrated in (43) below. This makes sense if these particles render the verb morphologically too complex, thereby preventing the verb from undergoing fusion with the focus head.

(43) *Vata* (Koopman 1984):

- a. (\*na`-)*le*    wa    ná`-*le-ka*  
       NEG eat    they    NEG-eat-FT  
       ‘They will not EAT’
- b. *li*(\*-wa)        wà    *li-wa*    zué  
       eat TP        they    eat-TP yesterday  
       ‘They ATE yesterday’

These restrictions can be interpreted as showing that if the realization of multiple copies is licensed via morphological fusion, it should naturally be very sensitive to

morphological information. The first kind of relevant information regards the feature composition of the elements that are to be fused. After all, not any two elements can get fused, but only the ones that satisfy the morphological requirements of one another. In Vata, for instance, the duplication of focused material only affects verbs and many languages only allow multiple copies of *wh*-elements, as will see below. This may be viewed as a reflex of the morphological (categorical) restrictions a given head may impose on the copy with which it may fuse. The second kind of information concerns morphological complexity. As a rule, the more morphologically complex a given element is, the less likely it is for it to undergo fusion and become part of a terminal. Thus, the addition of specific morphemes (which may vary from language to language) may make the resulting element morphologically “too heavy” to become reanalyzed as part of a word. This seems to be what is going on in (43), with the addition of Infl particles to the fronted verb. Of course, if a given copy is syntactically complex, i.e. it is phrasal, it is also morphologically complex and not a good candidate to undergo morphological fusion.<sup>14</sup>

---

<sup>14</sup> There are languages that allow a fronted predicate to be duplicated, as illustrated by Yoruba in (i). If (i) does involve nondistinctive copies, they should be somehow prevented from being computed “at the same time” for purposes of linearization. See Aboh 2006 and Kobele 2006 for specific suggestions.

(i) *Yoruba* (Kobele 2006):

Rira    adi        ti    Jim    ra        adi  
 buying chicken    TI    Jim    HTS    buy    chicken  
 ‘the fact/way that Jim bought a chicken’

This general approach provides a natural account of *wh*-copying constructions found in many languages, as illustrated by German in (44) below. *Wh*-copying constructions are subject to two intriguing constraints. First, although more than one trace may be phonetically realized (cf. (44)), only intermediate traces can be pronounced, as shown by the ungrammaticality of (45), where the tail of the *wh*-chain is realized, as well. The second pervasive characteristic of *wh*-copying constructions is that, roughly speaking, they can only involve simplex, not complex *wh*-phrases, as illustrated by (46).

(44) *German* (Fanselow and Mahajan 2000):

*Wen* denkst Du *wen* sie meint *wen* Harald liebt?  
*who* think you *who* she believes *who* Harald loves  
 ‘Who do you think that she believes that Harald loves?’

(45) *German*:

\**Wen* glaubt Hans *wen* Jakob *wen* gesehen hat?  
*whom* thinks Hans *whom* Jakob *whom* seen has  
 ‘Who does Hans think Jakob saw?’

(46) *German* (McDaniel 1986):

\**Welche Bücher* glaubst du *welche Bücher* Hans liest?  
*which book* think you *which book* Hans reads  
 ‘Which book do you think Hans is reading?’

Nunes (1999, 2004) argues that this paradigm can be accounted for, if long distance *wh*-movement in languages that allow for *wh*-copying constructions may proceed via head adjunction to C, as illustrated in (47a),<sup>15</sup> and if a [-wh] C fuses with the adjoined *wh*-element in the morphological component, as represented in (47b).

- (47) a. [CP [C<sup>0</sup> WH [C<sup>0</sup> Q]] ... [CP [C<sup>0</sup> WH [C<sup>0</sup> C<sub>[-wh]</sub>]] [TP ... WH ...]]  
 b. [CP [C<sup>0</sup> WH [C<sup>0</sup> Q]] ... [CP #<sub>[C<sup>0</sup> WH [C<sup>0</sup> C<sub>[-wh]</sub>]]#] [TP ... WH ...]]</sub>

The *wh*-chain in (47b) has only two links visible to the LCA, as the intermediate *wh*-copy becomes invisible after it undergoes fusion. The two visible copies should then prevent the structure from being linearized unless Chain Reduction is employed. Thus, the derivation of (45), for instance, cannot converge because the relevant structure

---

<sup>15</sup> For arguments that head adjunction should in general be preferred over movement to specifiers, all things being equal, see Nunes 1998 and Bošković 2001. Suggestive evidence that *wh*-movement in *wh*-copying does indeed involve head adjunction is provided by the fact the *wh*-copying is more restricted than regular *wh*-movement. In particular, it is subject to negative islands even when arguments are moved, as illustrated in (i), which can be accounted for if *wh*-copying involves head-adjunction to Comp and if an intervening Neg head blocks such head movement.

- (i) *German* (Reis 2000):

\**Wen* glaubst du nicht, *wen* sie liebt?  
*whom* believe you not *whom* she loves  
 ‘Who don’t you think that she loves?’

cannot be linearized.<sup>16</sup> Under the assumption that the highest copy in (47b) has more features checked, it should be kept and the lowest copy should be deleted, as discussed in section 3, yielding (48).

(48) [CP [C<sup>0</sup> WH [C<sup>0</sup> Q]] .... [CP # [C<sup>0</sup> WH [C<sup>0</sup> C<sub>[-wh]</sub>]#] [TP ... ~~WH~~ ...]]

We now have an answer for why the tail of the *wh*-chain contrasts with intermediate traces with respect to phonetic realization. There is nothing intrinsic to intermediate traces themselves that allows them to be phonetically realized. Rather, morphological requirements of the intermediate C<sup>0</sup> may trigger fusion with the adjoined *wh*-copy, making it invisible for the LCA and, consequently, for deletion. Once the system only “sees” the highest and the lowest *wh*-copies in (47b), its linearization as in (48) is no different from the linearization of a standard *wh*-movement construction such as (49), where economy considerations regarding applications of operations before Spell-Out ultimately determine the deletion of the lower *wh*-copy (see section 3).

---

<sup>16</sup> Contrary to what may seem at first glance, movement of the verb from T to Foc in (42a) or movement of the *wh*-element from one head-adjoined position to another in (47a) is not incompatible with Baker’s (1988) account of the general ban on excorporation (if the ban indeed holds). According to Baker, given the head adjunction structure [Y<sup>0</sup> X<sup>0</sup> Y<sup>0</sup>], if X<sup>0</sup> moves, the morphological component will receive a head with an adjoined trace, which was taken to be an illicit morphological object. Under the copy theory, Baker’s proposal can be interpreted as saying that deletion of copies cannot take place under an X<sup>0</sup> element. Notice that it is a crucial feature of the analysis reviewed above that the V-copy adjoined to F in (42a) and *wh*-copy adjoined to the intermediate C<sup>0</sup> in (47a) do *not* delete.

- (49) a. What did John see?  
 b. [<sub>CP</sub> what did [<sub>IP</sub> John see ~~what~~]]

Finally, by having *wh*-copying be dependent on morphological fusion, we reach a natural explanation for why complex *wh*-phrases do not license *wh*-copying (cf. (46)). The more morphologically complex a given element is, the harder it is for it to be fused and be reanalyzed as part of a word. Thus, the unacceptability of sentences such as (46) is arguably due to the fact that the *wh*-phrases cannot undergo fusion with the intermediate  $C^0$  due to their morphological complexity. This in turn entails that all the copies of the moved *wh*-phrase are visible to the LCA and failure to delete all but one link prevents their structures from being linearized.<sup>17</sup>

Nunes (2003, 2004) shows that the reasoning presented above also accounts for phonetic realization of more than one link in remnant movement constructions. Consider duplication of emphatic focus in Brazilian Sign Language (*LSB*) in (50) below, for instance. Nunes and Quadros (2006, 2008) argue that in constructions such as (50), the focused element moves and adjoins to a Focus head, followed by remnant movement of TP and fusion between Foc and the adjoined element in the morphological

---

<sup>17</sup> It should be noted that one finds considerable dialectal and idiolectal variation among speakers who accept *wh*-copying constructions. From the perspective reviewed here, variation in this regard is not due to syntactic computations proper, but to the degree of morphological complexity a given dialect or idiolect tolerates under fusion. As a rule, the more complex a constituent, the less likely it is for it to undergo fusion and become invisible to the LCA.

component.<sup>18</sup> Under the implementation of the linearization approach to copy deletion proposed in section 3, the derivation of (50a), for instance, is as sketched in (51).

(50) *Brazilian Sign Language:*

a. I LOSE BOOK LOSE

‘I LOST the book.’

b. [JOHN BUY WHICH BOOK YESTERDAY]<sub>wh</sub> [WHICH]<sub>wh</sub>

‘Which book exactly did John buy yesterday?’

(51) a. [<sub>FocP</sub> Foc [<sub>TP</sub> I LOSE<sub>F:?</sub> BOOK]]

b. *Adjunction to Foc:*

[<sub>FocP</sub> [<sub>Foc</sub><sup>0</sup> LOSE<sub>F:√</sub> [<sub>Foc</sub><sup>0</sup> Foc<sup>0</sup>]] [<sub>TP</sub> I LOSE<sub>F:?</sub> BOOK]]

c. *Probing by the higher copy:*

[<sub>FocP</sub> [<sub>Foc</sub><sup>0</sup> LOSE<sub>F:√</sub> [<sub>Foc</sub><sup>0</sup> Foc<sup>0</sup>]] [<sub>TP</sub> I LOSE<sub>F:√</sub> BOOK]]

d. *Remnant movement of TP:*

[[<sub>TP</sub> I LOSE<sub>F:√</sub> BOOK] ... [<sub>FocP</sub> [<sub>Foc</sub><sup>0</sup> LOSE<sub>F:√</sub> [<sub>Foc</sub><sup>0</sup> Foc<sup>0</sup>]] [<sub>TP</sub> I LOSE<sub>F:√</sub> BOOK]]]

e. *Spell-Out + fusion:*

[[<sub>TP</sub> I LOSE<sub>F:√</sub> BOOK] ... [<sub>FocP</sub> # [<sub>Foc</sub><sup>0</sup> LOSE<sub>F:√</sub> [<sub>Foc</sub><sup>0</sup> Foc<sup>0</sup>]] # [<sub>TP</sub> I LOSE<sub>F:√</sub> BOOK]]]

---

<sup>18</sup> Independent differences aside, the analysis of duplication of focus in Brazilian Sign Language to be sketched below can also be extended to the constructions involving focus duplication in American Sign Language originally discussed by Petronio (1993) and Petronio and Lillo-Martin (1997) (see Nunes 2004 and Nunes and Quadros 2006, 2008 for further discussion).

f. *Chain Reduction of CH = (TP, TP):*

[[<sub>TP</sub> I LOSE<sub>F:√</sub> BOOK] ... [<sub>FocP</sub> # [<sub>Foc<sup>0</sup></sub> LOSE<sub>F:√</sub> [<sub>Foc<sup>0</sup></sub> Foc<sup>0</sup>]]] # [<sub>TP</sub> I LOSE<sub>F:√</sub> BOOK]]

As discussed earlier, after the verb adjoins to Foc in (51b), valuation of the lower copy by the higher one, as in (51c), is not economical and will be resorted to only if triggered by independent requirements. This is indeed the case here. After LOSE and Foc fuse in the morphological component, as in (51e), the fused copy becomes invisible to the LCA and Chain Reduction is not called upon to delete the lower link of the chain CH = (LOSE, LOSE) formed when the verb adjoined to Foc. Thus, in order for the derivation to converge at PF, valuation of the lower copy of LOSE in (51c) must occur before Spell-Out. That being so, the only chain subject to reduction is the TP chain and deletion of its lower link as in (51f) yields the sentence in (50a).

Despite being optional, focus duplication is a very pervasive phenomenon in Brazilian Sign Language, being able to affect several kinds of constituents. However, there is a major restriction on this construction: the duplicated material cannot be morphologically complex (see Nunes 2003, 2004, Nunes and Quadros 2006, 2008 for discussion), as illustrated in (52a) below, with a verb that requires agreement morphology (annotated by the indices), and (52b), with a *wh*-phrase. Once the phonetic realization of multiple copies is dependent on morphological fusion and fusion is sensitive to morphological complexity, the ungrammaticality of the sentences in (52) can be attributed to the impossibility of fusion involving the moved elements. The presence of multiple copies that are visible to the LCA in the phonological component then prevents the structures underlying these constructions from being linearized (see section 3).

(52) *Brazilian Sign Language:*

a. \*JOHN <sub>a</sub>LOOK<sub>b</sub> MARY <sub>a</sub>LOOK<sub>b</sub>

‘John LOOKED at Mary’

b. \*[JOHN BUY WHICH BOOK YESTERDAY]<sub>wh</sub> [WHICH BOOK]<sub>wh</sub>

‘Which book exactly did John buy yesterday?’

Martins’s (2007) analysis of verb duplication in emphatic affirmation constructions in European Portuguese provides further evidence for this approach. In European Portuguese, the denial of a previous statement in a given context may involve verbal duplication at the right edge of the sentence, as illustrated in (53a) and (53b) below. Martins argues that both types of sentences are derived by movement of the verb to the polarity head  $\Sigma$ , followed by verb movement to C and fusion between C and the verb in the morphological component. The difference between them is that (53a) involves ellipsis, whereas (53b) involves remnant movement, as sketched in (54a) and (54b), respectively.

(53) *European Portuguese:*

a. A: – O João não comprou um carro, pois não?

the João not bought a car POIS NEG

‘John didn’t buy a car, did he?’

B: – *Comprou, comprou.*

bought bought

‘Yes, he DID.’

b. A: – Ele não comprou o carro.

he not bought the car

‘He didn’t buy the car.’

B: – Ele comprou o carro, comprou.

he bought the car, bought

‘He did buy the car.’

(54) a. [CP [C' [C comprou] [<sub>SP</sub> [<sub>Σ</sub>' comprou] [<sub>TP</sub> [<sub>T</sub>' ~~comprou~~

bought bought bought

[<sub>VP</sub> ~~o João comprou o carro~~]]]]]]]]

the João bought the car

b. [CP [<sub>SP</sub> ele comprou o carro] [C' [C comprou] [<sub>SP</sub> ~~ele~~ [<sub>Σ</sub>' ~~comprou~~

he bought the car bought he bought

[<sub>TP</sub> [<sub>T</sub>' ~~comprou~~ [<sub>VP</sub> ~~ele comprou o carro~~]]]]]]]]

bought he bought the car

Relevant for our current purposes is Martins’s documentation of a series of contexts that block verbal duplication, as illustrated in (55), with compound verbs and verbs with stressed prefixes. As Martins argues, the ungrammaticality of sentences such as (55) is to be attributed to the morphological complexity of their verbs, which should block fusion; in turn, once more than one copy of the verb is visible to the phonological component, the whole structure cannot be linearized.

(55) *European Portuguese*:

a. A: — Ele não fotocopiou o livro sem autorização, pois não?

he not photocopied the book without permission, POIS NEG

‘He didn’t copy the book without your permission, did he?’

B: — ??*Fotocopiou, fotocopiou.*

photocopied photocopied

‘Yes, he DID.’

b. A: — O candidato não contra-atacou, pois não?

the candidate not counter-attacked, POIS NEG

‘The candidate didn’t counter-attack, did he?’

B: — ??*Contra-atacou, contra-atacou.*

counter-attacked counter-attacked

‘Yes, he DID.’

Remnant movement constructions thus provide further empirical support for the copy theory in that they can also allow more than one chain link to be phonetically realized, provided that linearization and morphological requirements are satisfied.

## 5. The Copy Theory and the Debate on Obligatory Control

In the same way the mapping from Spell-Out to PF can provide compelling evidence for the copy theory over the trace theory, it can also set up independent grounds for choosing between the two major minimalist approaches to obligatory control which are currently under debate. As far as the mapping from Spell-Out to PF is concerned, the PRO-based approach, be it in terms of null Case (see e.g. Chomsky and Lasnik 1993, Martin 2001) or in terms of Agree (see e.g. Landau 2004), is no different than the GB

approach. In other words, once PRO is by definition devoid of phonetic content, its chain will receive no interpretation at PF. By contrast, under the movement approach to obligatory control proposed by Hornstein (2001), obligatorily control is actually a trace (i.e. a copy) of the controller. Under this view, the fact that the controller is the element that surfaces at PF follows from the fact that in general heads of chains are the ones that are realized at PF. But recall that pronunciation of chain heads is just the optimal output when there are no additional requirements on specific chain links. Thus, given the several possibilities discussed in section 4 for a chain to be realized at PF, there arises the possibility that control constructions may also display similar unorthodox realizations at PF.

Two such realizations have gained prominence recently. The first involves backward control constructions (see e.g. Polinsky and Potsdam 2002 and Boeckx, Hornstein, and Nunes 2010 for references and relevant discussion). As convincingly argued by Polinsky and Potsdam (2002), Tsez, for instance, allows control constructions where the thematic matrix subject is obligatorily null and obligatorily bound by the embedded overt subject, as illustrated in illustrated in (56).

(56) *Tsez*:

[ $\Delta_{1/*2}$  [kidbā<sub>1</sub>    ziya        bišra]    yoqsi]

*girl.ERG cow.ABS feed.INF began*

‘The girl began to feed the cow.’

Polinsky and Potsdam present several kinds of evidence all pointing to the conclusion that the phonetically realized subject in sentences such as (56) does indeed sit in the embedded clause. For instance, (57) below shows that the case marking on the overt

subject is determined by the dative assigning verb in the embedded clause, whereas (58) shows that the overt subject cannot precede a matrix adverb. Polinsky and Potsdam then propose that backward control constructions involve movement to a thematic position, as in standard instances of control under Hornstein's (2001) approach, with the difference that a lower copy is pronounced instead, as sketched in (59).

(57) *Tsez*:

kid-ber    babiw-s    xabar    teq-a    7y-oq-si  
*girl.II-DAT father-GEN story.III-INF hear-INF begin-PAST.EVID*  
 'The girl began to hear the father's story.'

(58) *Tsez*:

a. ħuʃ            [kidbā    ziya    bišra] yoqsi  
*yesterday girl.ERG cow feed began*  
 b. \*kidbā    ħuʃ            [ziya bišra] yoqsi  
*girl.ERG yesterday cow feed began*  
 'Yesterday the girl began to feed the cow.'

(59) a. [DP<sub>1</sub> V [DP<sub>1</sub> ....]]

b. *Deletion in the phonological component (forward control):*

[DP<sub>1</sub> V [~~DP<sub>1</sub>~~ ....]]

b. *Deletion in the phonological component (backward control):*

[~~DP<sub>1</sub>~~ V [DP<sub>1</sub> ....]]

The second type of unorthodox control constructions discussed recently regards copy-control. Consider the data in (60) from San Lucas Quiavini Zapotec, discussed by Lee (2003).

(60) *San Lucas Quiavini Zapotec* (Lee 2003):

a. R-cààa'z Gye'eihlly g-auh Gye'eihlly bxaady.

HAB-wantMike IRR-eat Mike grasshopper

'Mike wants to eat grasshopper.'

b. B-quii'lly bxuuhahz Gye'eihlly ch-iiia Gye'eihlly scweel.

PERF-persuade priest Mike IRR-go Mike school

'The priest persuaded Mike to go to school.'

c. B-ii'lly-ga' Gye'eihlly zi'cygàa' nih cay-uhny Gye'eihlly zèèiny.

PERF-sing-also Mike while that PROG-do Mike work

'Mike sang while he worked.'

Each of the sentences (60) shows a bound copy in the embedded subject position.

Interestingly, the similarities of these constructions with standard control constructions go beyond translation. They also trigger a sloppy reading under ellipsis, as shown in (61), and the bound copy displays complementarity with a co-referential pronoun, as shown in (62).

(61) *San Lucas Quiavini Zapotec* (Lee 2003):

- a. R-cààa'z Gye'eihlly g-ahcnèe Gye'eihlly Lia Paamm zè'cy cahgza' Li'eb.  
 HAB-want Mike IRR-help Mike FEM Pam likewise Felipe  
 'Mike wants to help Pam, and so does Felipe (want to help Pam/\*want Mike to help Pam)'
- b. Zi'cygàa' nih cay-uhny Gye'eihlly zèèiny b-ii'lly-ga' Gye'eihlly  
 while that PROG-do Mike work PERF-sing-also Mike  
 zè'cy cahgza' Li'eb.  
 likewise Felipe  
 'While Mike<sub>i</sub> was working, he<sub>i</sub> sang, and so did Felipe<sub>k</sub> (sing while he<sub>k</sub> worked)'

(62) *San Lucas Quiavini Zapotec* (Felicia Lee p.c., 2003):

- a. R-caaa'z Gye'eihlly g-ahcnèe-ëng Lia Paamm.  
 HAB-want Mike IRR-help-3SG.PROX FEM Pam  
 'Mike<sub>i</sub> wants him<sub>k/\*i</sub> to help Pam'
- b. Zi'cygàa' nih cay-uhny-ëng zèèiny b-ii'lly-ga' Gye'eihlly.  
 while that PROG-do-3SG.PROX work PERF-sing-also Mike  
 'While he<sub>i/\*k</sub> worked, Mike<sub>k</sub> sang'

Boeckx, Hornstein, and Nunes (2007, 2008) argue that the data in (60)-(62) are indeed cases of control, i.e. movement to thematic positions, with both the controller and the controlee copies being phonetically realized. More specifically, they propose that these constructions involve morphological fusion of the controlee copy with a null 'self' morpheme available in this language.<sup>19</sup> As we should expect given the discussion

<sup>19</sup> Boeckx, Hornstein, and Nunes (2007, 2008) argue that fusion with this null 'self'

above, if a control chain involves morphologically encumbered copies, fusion will be blocked and phonetic realization of more than one copy leads to an ungrammatical result. That this prediction is correct is illustrated by the copy control constructions in (63a), which involves a quantifier phrase, and in (63b), whose links contain an anaphoric possessor.

(63) *San Lucas Quiavini Zapotec* (Lee 2003):

- a. \*Yra'ta' zhyàa'p r-càaa'z g-ahcnèe' yra'ta' zhyàa'p Lia Paamm.  
 every girl HAB-want IRR-help every girl FEM Pam  
 'Every girl wants to help Pam.'
- b. \*R-e'ihpy Gye'eihlly behts-ni' g-a'uh behts-ni' bx:àady.  
 HAB-tell Mike brother-REFL.POSS IRR-eat brother-REFL.POS grasshopper  
 'Mike told his brother to eat grasshoppers.'

Let us reexamine the adjunct copy-control case in (60c). As argued by Hornstein (2001), adjunct control involves sideward movement (in the sense of Nunes 2001, 2004) of the embedded subject before the adjunct clause attaches to vP. The fact that sideward morpheme is also what underlies the existence of copy-reflexive constructions in San Luca Quiavini Zapotec such as the ones illustrated in (i).

(i) *San Luca Quiavini Zapotec* (Lee 2003):

- B-gwa Gye'eihlly Gye'eihlly.  
*PERF-shave Mike Mike*  
 'Mike shaved himself.'

movement may also lead to phonetic realization of multiple copies shows that sideward movement is nothing more than one of the instantiations of Copy plus Merge.

Interestingly, there are languages which only allow adjunct copy-control, which indicates that the relevant head that triggers fusion in these languages is within the adjunct clause. In his detailed study on control structures in Telugu and Assamese, Haddad (2007) shows that adjunct copy-control constructions such as (64) and (65) below (*CNP* stands for conjunctive participle particle) display all the traditional diagnostics of obligatory control and argues that they should also be analyzed in terms of sideward movement and phonetic realization of multiple copies.

(64) *Telugu* (Haddad 2007):

[[Kumar sinima cuus-tuu] [Kumar popkorn tinnaa-Du]]

*Kumar.NOM movie watch-CNP Kumar.NOM popcorn ate-3-M.S*

‘While watching a movie, Kumar ate popcorn.’

(65) *Assamese* (Haddad 2007):

[[Ram-Or khong uth-i] [Ram-e mor ghorto bhangil-e]]

*Ram-GEN anger raise-CNP Ram-NOM my house destroyed-3*

‘Having got angry, Ram destroyed my house.’

Given the role of morphological fusion in making the phonetic realization of multiple copies possible, it comes as no surprise that multiple copies are only possible if, in Haddad’s (2007:87) words, the subject “does not exceed one or two words”, as illustrated by the ungrammaticality of (66).

(66) *Telugu* (Haddad 2007):

\*[[Kumar        maryu Sarita        sinim cuu-tuu] [Kumar        maryu  
*Kumar.NOM and Sarita.NOM movie watch-CNP Kumar.NOM and*  
 Sarita        popkorn tinna-ru]]  
*Sarita.NOM popcorn ate*

‘While Kumar and Sarita were watching a movie, they ate popcorn.’

To the extent that backward control and copy-control constructions are roughly analyzed along the lines suggested above, they provide decisive grounds for choosing between PRO-based and movement-based approaches to control. More specifically, these constructions prove fatal to PRO-based approaches to control, as PRO is taken to be a phonetically null element by definition. In contrast, backward control and copy-control are in fact expected under a movement-based approach to control under the copy theory, given its potential different outputs at PF.

## 6. Concluding Remarks

Chomsky’s (1993) original arguments for incorporating the copy theory into the minimalist framework had to do with interpretation effects and the mapping from the numeration to LF. However, optimizing this mapping by assuming the copy theory seems to have an unwelcome consequence as it appears to require stipulating that in the mapping from Spell-Out to PF, lower copies (the old traces) must be deleted. After all, the null hypothesis regarding the copy theory is that if  $\alpha_1$  is a copy of  $\alpha_2$ , they should have the same status in the computational system. Thus, if the highest can be pronounced, lower copies should in principle be pronounceable, as well.

Upon close inspection, vice can be turned into virtue. As seen in the previous sections, lack of phonetic realization is not an intrinsic property that characterizes traces as grammatical primitives. Traces or parts of traces may be phonetically realized if the pronunciation of the head of the chain causes the derivation to crash at PF. The fact that traces in the general case are not phonetically realized results from the interaction among convergence and economy factors. On the one hand, linearization requirements trigger deletion of “repeated” material within a chain. On the other hand, economy considerations regarding the valuation of lower copies by higher ones before Spell-Out and the number of applications of deletion within a chain work in such a way that they render higher copies more PF-palatable than lower ones. Thus, if the phonological component imposes no additional convergence conditions that can affect these optimality computations, the head of a chain will always be the optimal option for phonetic realization. However, these economy considerations may be overruled by convergence requirements in the phonological component in which case we may have pronunciation of a lower copy, pronunciation of different parts of different copies, and even pronunciation of more than one copy.

In sum, we have seen that rethinking movement operations in terms of the copy theory, which was driven by the minimalist search for conceptual elegance, has led to a considerable enlargement of the empirical coverage previously handled.

## **References**

- Aboh, Enoch (2006). ‘When Verbal Predicates Go Fronting’. *ZAZ Papers in Linguistics* 46:21-48.
- Aoun, Joseph and Jairo Nunes (2007). ‘Vehicle Change Phenomena as an Argument for Move-F’. *Linguistic Inquiry* 38:525-538.

- Baker, Mark (1988). *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.
- Barbiers, Sjef, Olaf Koenenman, and Marika Lekakou (2010). 'Syntactic Doubling and the Nature of *Wh*-chains'. *Journal of Linguistics* 46:1-46.
- Bobaljik, Jonathan (1995). *Morphosyntax: The syntax of verbal inflection*. Doctoral dissertation, MIT, Cambridge, Mass.
- Bobaljik, Jonathan (2002). 'A-Chains at the PF-interface: Copies and 'covert' movement'. *Natural Language and Linguistic Theory* 20.2:197-267.
- Boeckx, Cedric, Norbert Hornstein, and Jairo Nunes (2007). 'Overt Copies in Reflexive and Control Structures: A Movement Analysis'. *University of Maryland Working Papers in Linguistics* 15:1-45.
- Boeckx, Cedric, Norbert Hornstein, and Jairo Nunes (2008). 'Copy-reflexive and Copy-control Constructions: A Movement Analysis'. *Linguistic Variation Yearbook* 8: 61-99.
- Boeckx, Cedric, Norbert Hornstein, and Jairo Nunes (2010). *Control as Movement*. Cambridge: Cambridge University Press.
- Bošković, Željko (2001). *On the Nature of the Syntax-Phonology Interface: Cliticization and Related Phenomena*. Amsterdam: Elsevier Science.
- Bošković, Željko (2002). 'On Multiple *Wh*-fronting'. *Linguistic Inquiry* 33: 351-383.
- Bošković, Željko (2007). 'On the Locality and Motivation of Move and Agree: An Even More Minimal Theory'. *Linguistic Inquiry* 38:589-644.
- Bošković, Željko and Jairo Nunes (2007). 'The Copy Theory of Movement: A View from PF'. In Norbert Corver and Jairo Nunes (eds.), *The Copy Theory of Movement*, 13-74. Amsterdam/Philadelphia: John Benjamins.

- Ćavar, Damir and Gisbert Fanselow (1997). 'Split constituents in Germanic and Slavic'. Paper presented at International Conference on Pied-piping. Friedrich-Schiller Universität Jena.
- Chomsky, Noam (1993). 'A Minimalist Program for Linguistic Theory'. In Kenneth Hale and Samuel J. Keyser (eds.), *The View from Building 20*, 1-52. Cambridge, Mass.: MIT Press.
- Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam and Howard Lasnik (1993). 'The Theory of Principles and Parameters'. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann (eds.), *Syntax: An International Handbook of Contemporary Research*, 506-569. Berlin/New York: Walter de Gruyter.
- Corver, Norbert and Jairo Nunes (2007). *The Copy Theory of Movement*. Amsterdam/Philadelphia: John Benjamins.
- Fanselow, Gisbert and Anoop Mahajan (2000). 'Towards a Minimalist Theory of *Wh*-expletives, *Wh*-copying, and Successive Cyclicity. In Uli Lutz, Gereon Müller, and Armin von Stechow (eds.), *Wh-scope Marking*, 195-230. Amsterdam/Philadelphia: John Benjamins.
- Franks, Steven (1998). 'Clitics in Slavic'. Position paper presented at the Comparative Slavic Morphosyntax Workshop, June 5-7, Bloomington, In. (downloadable from <http://www.indiana.edu/~slavconf/linguistics/index.html>)
- Fujii, Tomohiro (2007). 'Cyclic chain reduction'. In Norbert Corver and Jairo Nunes (eds.), *The Copy Theory of Movement*, 291-326. Amsterdam/Philadelphia: John Benjamins.

- Gärtner, Hans-Martin (1998). 'Review of *The Copy Theory of Movement and Linearization of Chains in the Minimalist Program*'. *GLOT International* 8.3:16-20.
- Golston, Chris (1995). 'Syntax Outranks Phonology: Evidence from Ancient Greek'. *Phonology* 12: 343-368.
- Grohmann, Kleanthes (2003). *Prolific Domains: On the Anti-Locality of Movement Dependencies*. Amsterdam/Philadelphia: John Benjamins.
- Haddad, Youssef (2007). *Adjunct Control in Telugu and Assamese*. Doctoral dissertation, University of Florida.
- Halle, Morris and Alec Marantz (1993). Distributed Morphology and the Pieces of Inflection. In Kenneth Hale and Samuel J. Keyser (eds.), *The View from Building 20*, 111-176. Cambridge, Mass.: MIT Press.
- Holmberg, Anders (1986). *Word Order and Syntactic Features in the Scandinavian Languages and English*. Doctoral dissertation, University of Stockholm.
- Hornstein, Norbert (2001). *Move! A Minimalist Theory of Construal*. Malden: Blackwell.
- Hornstein, Norbert (2007). 'Pronouns in Minimal Setting'. In Norbert Corver and Jairo Nunes (eds.), *The Copy Theory of Movement*, 351-385. Amsterdam/Philadelphia: John Benjamins.
- Kandybowicz, Jason (2008). *The Grammar of Repetition: Nupe Grammar at the Syntax-Phonology Interface*. Amsterdam/Philadelphia: John Benjamins.
- Kayne, Richard (1994). *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.
- Kobele, Gregory (2006). *Generating Copies: An Investigation into Structural Identity in Grammar and Language*. Doctoral dissertation, University of California at Los Angeles.

- Koopman, Hilda (1984). *The syntax of verbs*. Dordrecht: Foris.
- Landau, Idan (2004). 'The Scale of Finiteness and the Calculus of Control'. *Natural Language and Linguistic Theory* 22:811-877.
- Lee, Felicia (2003). 'Anaphoric R-expressions as Bound Variables'. *Syntax* 6, 84-114.
- Lidz, Jeffrey and William Idsardi (1997). 'Chains and Phono-logical Form'. *UPenn Working Papers in Linguistics* 8:109-125.
- Martin, Roger (2001). 'Null Case and the Distribution of PRO'. *Linguistic Inquiry* 32:141-166.
- Martins, Ana Maria (2007). 'Double Realization of Verbal Copies in European Portuguese Emphatic Affirmation'. In Norbert Corver and Jairo Nunes (eds.), *The Copy Theory of Movement*, 77-118. Amsterdam/Philadelphia: John Benjamins.
- McDaniel, Dana (1986). *Conditions on Wh-chains*. Doctoral dissertation, CUNY.
- Nunes, Jairo (1995). *The Copy Theory of Movement and Linearization of Chains in the Minimalist Program*. Doctoral dissertation, University of Maryland, College Park.
- Nunes, Jairo (1998). 'Bare X'-theory and Structures Formed by Movement'. *Linguistic Inquiry* 29:160-168.
- Nunes, Jairo (1999). 'Linearization of Chains and Phonetic Realization of Chain Links'. In Samuel D. Epstein and Norbert Hornstein (eds.), *Working Minimalism*, 217-249. Cambridge, Mass.: MIT Press.
- Nunes, Jairo (2001). 'Sideward Movement'. *Linguistic Inquiry* 31:303-344
- Nunes, Jairo (2003). 'Head Movement, Remnant Movement, and Phonetic Realization of Chains'. *Syntax at Sunset 3: Head Movement and Syntactic Theory*, 161-177, UCLA/Potsdam Working Papers in Linguistics.
- Nunes, Jairo (2004). *Linearization of Chains and Sideward Movement*. Cambridge, Mass.: MIT Press.

- Nunes, Jairo and Ronice Quadros (2006). 'Duplication of *Wh*-elements in Brazilian Sign Language'. In Leah Batenab and Cherlon Ussery (eds.), *Proceedings of the Thirty-fifth Annual Meeting of the North East Linguistic Society* vol. 2, 463-477. GLSA, University of Massachusetts, Amherst.
- Nunes, Jairo and Ronice Quadros (2008). 'Phonetically Realized Traces in American Sign Language and Brazilian Language'. In Josep Quer (ed.), *Signs of the Time: Selected Papers from TISLR 2004*, 177-190. Seedorf: Signum-Verlag.
- Pesetsky, David (1997). 'Optimality Theory and Syntax: Movement and Pronunciation'. In Diana Archangeli and D. Terence Langendoen (eds.), *Optimality Theory: An Overview*, 134-170. Malden: Blackwell.
- Pesetsky, David (1998). 'Some Optimality Principles of Sentence Pronunciation'. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky (eds.), *Is the Best Good Enough*, 337-383. Cambridge, Mass.: MIT Press.
- Petronio, Karen (1993). *Clause Structure in American Sign Language*. Doctoral dissertation, University of Washington, Seattle.
- Petronio, Karen and Diane Lillo-Martin (1997). 'Wh-movement and the Position of Spec-CP: Evidence from American Sign Language'. *Language* 73:18-57.
- Polinsky, Maria and Eric Potsdam (2002). 'Backward Control'. *Linguistic Inquiry* 33, 245-282.
- Reis, Marga (2000). 'On the Parenthetical Features of German *Was...W*-constructions and How to Account for Them. In Uli Lutz, Gereon Müller, and Armin von Stechow (eds.), *Wh-scope Marking*, 249-270. Amsterdam/Philadelphia: John Benjamins.

- Rice, Curt and Peter Svenonius (1998). 'Prosodic V2 in Northern Norwegian'. Paper presented at the Seventeenth West Coast Conference on Formal Linguistics. University of British Columbia, Vancouver, February 1998.
- Richards, Norvin (2006). 'A Distinctiveness Condition on Linearization'. Ms., Massachusetts Institute of Technology.
- Rudin, Catherine, Christina Kramer, Loren Billings, and Matthew Baerman (1999). 'Macedonian and Bulgarian *Li* Questions: Beyond Syntax'. *Natural Language and Linguistic Theory* 17, 541-586.
- Saab, Andrés (2008). *Hacia Una Teoría de la Identidad Parcial en la Elipsis*. Doctoral dissertation, Universidad de Buenos Aires.
- Stjepanović, Sandra (1999). *What Do Second Position Cliticization, Scrambling, and Multiple Wh-fronting Have in Common?* Doctoral dissertation, University of Connecticut, Storrs.
- Stjepanović, Sandra (2003). 'A Word-order Paradox Resolved by Copy Deletion at PF'. *Linguistic Variation Yearbook* 3: 139-177.
- Stjepanović, Sandra (2007). 'Free Word Order and the Copy Theory of Movement'. In Norbert Corver and Jairo Nunes (eds.), *The Copy Theory of Movement*, 219-248. Amsterdam/Philadelphia: John Benjamins.
- Wilder, Chris (1995). 'Rightward Movement as Leftward Deletion'. In Uli Lutz and Jürgen Pafel (eds.), *On Extraction and Extraposition in German*, 273-309. Amsterdam/Philadelphia: John Benjamins.

## **Abstract**

This chapter reviews the original motivation for the incorporation of the copy theory of movement into the Minimalist Program and discusses its consequences for the syntax-PF mapping. In particular, it discusses cases where traces (lower copies) are pronounced, as well as cases where more than one chain link is phonetically realized.

**Keywords:** copy theory, traces, syntax-PF mapping, lower copy pronunciation, pronunciation of multiple copies

## **Biographical note**

Jairo Nunes is currently Associate Professor at the Universidade de São Paulo. He is the author of *Linearization of Chains and Sideward Movement* (MIT Press, 2004), co-author of *Understanding Minimalism* (Cambridge University Press, 2005) and *Control as Movement* (Cambridge University Press, 2010), and co-editor of *The Copy Theory of Movement* (John Benjamins, 2007) and *Minimalist Essays on Brazilian Portuguese Syntax* (John Benjamins, 2009). He is also co-editor of *Probus: International Journal of Latin and Romance Linguistics* (Mouton de Gruyter).

## **Index**

Chain Reduction

copy theory

copy-control

control

economy conditions

Inclusiveness Condition

LCA (Linear Correspondence Axiom)

nondistinctiveness

phonetic realization of lower copies

phonetic realization of multiple copies

reconstruction

remnant movement

scattered deletion

trace theory