

Sideward Movement: Triggers, Timing, and Outputs*

Jairo Nunes

1. Introduction

Within the GB (Government and Binding) model, sideward movement – i.e. movement from a syntactic tree K to another syntactic tree L independent from K – is not a theoretical possibility. D-Structure provides the computational system with a single root tree and all the syntactic computations after D-Structure must operate within this root syntactic object. Thus, it is not at all surprising that sideward movement was not explored in GB. However, the theoretical framework that prevented sideward movement in GB becomes completely different within minimalism and it is worth discussing whether it should still be prevented and at which cost.

Let us focus on two major differences between GB and minimalism that bear on this issue. First, D-Structure is dispensed with for not being an interface level. Generalized transformations are then revived, allowing the computational system to operate with more than one root syntactic tree at a time. This is in fact trivially true for the first steps of any syntactic derivation. Take the first steps of derivation of (1) below, for instance. Given the (simplified) numeration in (2), the computational system independently selects *saw* and *her* and then merges them, as shown in (3). Crucially, *saw* and *her* in (3b) are both root syntactic objects.

(1) The boy saw her

- (2) $N = \{the_1, boy_1, saw_1, her_1\}$
- (3) a. $N' = \{the_1, boy_1, saw_0, her_1\}$
 $K = saw$
- b. $N'' = \{the_1, boy_1, saw_0, her_0\}$
 $K = saw$
 $L = her$
- c. $N''' = \{the_1, boy_1, saw_0, her_0\}$
 $M = [saw\ her]$

Another case where the computational system must deal with more than one root syntactic object at a time involves complex specifiers or complex adjuncts. Consider again the derivation of (1). Chomsky (1995) has argued that the computational complexity of syntactic derivations can be substantially minimized if we assume the Extension Condition, which requires that projecting operations work at the root node. Given the derivational step in (3c), for instance, the Extension Condition excludes the continuation in (4), where *boy* merges with M and then *the* merges with the nonroot syntactic object *boy* (cf. (4c-d)).

- (4) a. $N'''' = \{the_1, boy_0, saw_0, her_0\}$
 $M = [saw\ her]$
 $O = boy$
- b. $N'''' = \{the_1, boy_0, saw_0, her_0\}$
 $P = [boy\ [saw\ her]]$

c. $N'''' = \{the_0, boy_0, saw_0, her_0\}$

P = [boy [saw her]]

Q = the

d. $N'''' = \{the_0, boy_0, saw_0, her_0\}$

R = [[the boy] [saw her]]

Instead, the Extension Condition enforces the continuation in (5) below, where *boy* and *the* are selected and merged (cf. (5a-c)) before the resulting object merges with M.

Crucially, at the derivational step in (5b), there are three root syntactic objects in the derivational workspace.

(5) a. $N'''' = \{the_1, boy_0, saw_0, her_0\}$

M = [saw her]

O = boy

b. $N'''' = \{the_0, boy_0, saw_0, her_0\}$

M = [saw her]

O = boy

P = the

c. $N'''' = \{the_0, boy_0, saw_0, her_0\}$

M = [saw her]

Q = [the boy]

d. $N'''' = \{the_0, boy_0, saw_0, her_0\}$

R = [[the boy] [saw her]]

The second difference between GB and minimalism relevant for our purposes is the copy theory of movement, which reinterprets Move as the output of the interaction between the more basic operations Copy and Merge. The adoption of the copy theory is motivated by the attempt to eliminate noninterface levels (see Chomsky 1993), as well as the attempt to reduce the theoretical apparatus by only assuming syntactic primitives that can be understood in terms of (a rearrangement of features of) lexical items (Chomsky's (1995) Inclusiveness Condition). Under the copy theory, the derivation of a sentence such as (6) below, for instance, proceeds along the lines of (7), where the computational system creates a copy of *John*, merges it with the previously assembled TP, and deletes the lower copy in the phonological component.¹ Again, notice that in a system that has Copy as a basic operation, it must be the case that the computational system must be able to handle more than one root syntactic object, namely, the copy newly created and the root syntactic object containing the replicated material (cf. (7b)).

(6) John was arrested.

(7) a. K = [TP was arrested John]

b. *Copy:*

K = [TP was arrested Johnⁱ]

L = Johnⁱ

c. *Merge:*

M = [TP Johnⁱ was arrested Johnⁱ]

d. *Delete:*

P = [TP Johnⁱ was arrested ~~Johnⁱ~~]

What is relevant for our discussion is that if the computational system can operate with more than one root syntactic object at a time and if movement is understood as the interaction between the basic operations of Copy and Merge, sideward movement becomes a logical possibility within the system. That is, given two root syntactic objects K and L, the computational system may copy α from K and merge it with L, as illustrated in (8).

(8) a. $K = [\dots \alpha \dots]$

$L = [\dots]$

b. *Copy*:

$K = [\dots \alpha^i \dots]$

$L = [\dots]$

$M = \alpha^i$

c. *Merge*:

$K = [\dots \alpha^i \dots]$

$P = [\alpha^i [L \dots]]$

Terminological metaphors aside, note that there is no intrinsic difference between the “upward” movement seen in (7), for instance, and the “sideward” movement sketched in (8) with respect to the computational tools employed. In both cases, we have trivial applications of movement, viewed as Copy plus Merge. Sideward movement is therefore *not* a novel operation or a new species of movement. This point is worth emphasizing, as it has been consistently misunderstood. The fact that α in (8) does not merge with the structure that contains the “source” of the copy, as opposed to *John* in (7), may have independent explanations. First, (7) differs from (8) in an obvious

way: the copy of *John* in (7) has only one syntactic object to merge with, whereas the copy of α in (8) has two. But more importantly, it may be the case that Last Resort licenses merger of the copy of α in (8) with L but not with K. The derivation of V-to-T movement under the sideward movement analysis sketched in (9) illustrates this point.²

- (9) a. [VP ... V ...]
 T
- b. *Copy:*
 VP = [... Vⁱ ...]
 T
 Vⁱ
- c. *Merge (by adjunction):*
 VP = [... Vⁱ ...]
 K = [_{T0} Vⁱ [_{T0} T]]
- d. *Merge:*
- TP
- 3
- T⁰ [VP ... Vⁱ ...]
- 2
- Vⁱ T⁰

If T and VP had merged in (9a), yielding [_{TP} T VP], the Extension Condition should then prevent the verb from adjoining to T, as T would no longer be a root syntactic object (cf. (4c-d)). However, V-to-T adjunction can comply with the Extension Condition if it proceeds as in (9b-d), with copying of V preceding merger of the two-

segment T^0 with VP. Crucially, once the derivational step in (9b) is reached, the copied V must merge with T rather than VP, as V arguably has features to check with T, but not VP.³

In other words, sideward movement looks outlandish only if we examine it wearing GB lenses. If we wear minimalist lenses instead, we realize that it is a mere label for a specific sequence of Copy and Merge, which arises as a natural consequence of the interaction among core architectural features of the Minimalist Program, namely, the abandonment of D-Structure, the copy theory of movement, and the Extension Condition. It is worth noting that these architectural features are in turn conceptually grounded on the minimalist attempt to eliminate noninterface levels and reduce the number of primitives and the computational complexity of syntactic derivations. Thus, from a minimalist approach adopting these architectural features, sideward movement comes for free and does not increase the grammatical apparatus. In fact, one would need to complicate the system in order to exclude it.

In this paper, I discuss some extensions and refinements of the specific implementation of sideward movement proposed in Nunes 1995, 2001, 2004. I will discuss new empirical evidence for sideward movement and show that standard minimalist assumptions suffice to prevent overgeneration. The paper is organized as follows. In section 2 I briefly present Nunes's (1995, 2001, 2004) analysis of parasitic gaps and Hornstein's (1999, 2001) analysis of adjunct control to illustrate the general mechanics of derivations employing sideward movement. In section 3, I show how unwanted instances of sideward movement can be blocked. In section 4, I discuss empirical evidence for sideward movement bearing on the copy theory. In section 5, I discuss the interaction of economy and convergence requirements regulating sideward movement by examining adjunct control in Portuguese. In section 6, I show that

sideward movement may also apply in the morphological component. Finally, section 7 offers some concluding remarks.

2. Some examples of sideward movement

In this section I present two analyses employing sideward movement: Nunes's (1995, 2001, 2004) analysis of parasitic gaps and Hornstein's (1999, 2001) analysis of adjunct control. The purpose of this presentation is just to familiarize the reader with the general mechanics of sideward movement approaches. I leave the discussion of more technical details to sections 3 and 5 below.

Nunes (1995, 2001, 2004) argues that parasitic gap constructions constitute empirical instantiations of sideward movement. A parasitic gap construction such as (10) can be analyzed along the lines of (11).⁴

(10) [which paper]_i did you file t_i after John read PG_i?

(11) a. K = [John read [which paper]]

L = file

b. K = [John read [which paper]ⁱ]

L = file

M = [which paper]ⁱ

c. K = [John read [which paper]ⁱ]

P = [file [which paper]ⁱ]

d. K = [John read [which paper]ⁱ]

Q = [you [file [which paper]ⁱ]]

- e. [_{VP} [_{VP} you [_{file} **[which paper]ⁱ]]] [_{after} John read **[which paper]ⁱ]]****
- f. [_{did} [_{TP} you T [_{VP} [_{VP} you file **[which paper]ⁱ]]] [_{PP} after John read **[which paper]ⁱ]]]]****
- g. [**[which paper]ⁱ** did [_{TP} you T [_{VP} [_{VP} you file **[which paper]ⁱ]]] [_{PP} after John read **[which paper]ⁱ]]]]****
- h. [**[which paper]ⁱ** did [_{TP} you T [_{VP} [_{VP} ~~you~~ file ~~**[which paper]ⁱ**]]] [_{PP} after John read ~~**[which paper]ⁱ**]]]]~~~~

After the derivational step in (11a) is reached, the computational system makes a copy of *which paper* (cf. (11b)) and merges it with *file* (cf. (11c)) – an instance of sideward movement. Further computations involve the building of the matrix vP (cf. (11d-e)) and the matrix CP (cf. (11f)). After standard *wh*-movement takes place in (11g), the higher copy of *which paper* forms a distinct chain with each of the lower copies, capturing the fact that the two object positions are interpreted as bound by the *wh*-phrase in the matrix [Spec,CP]. Deletion of these lower copies in the phonological component (cf. (11h)) finally yields the parasitic gap construction in (10).

A similar derivation is employed by Hornstein (1999, 2001) to account for adjunct control. Based on the syntactic distribution and semantic interpretation of obligatorily controlled PRO, Hornstein argues that it is a trace/copy left by movement to a thematic position (see Boeckx, Hornstein and Nunes 2010 for detailed discussion). Crucially, the controlled PRO of adjunct clauses is no exception. More specifically, Hornstein proposes that adjunct control involves sideward movement. An adjunct control construction such as (12), for instance, can be derived as in (13).

(12) [_{TP} John_i [_{VP} [_{VP} t_i greeted everybody] [_{before} PRO_i leaving the room]]]

- (13) a. K = [**John** leaving the room]
 L = [greeted everybody]
- b. K = [**John**ⁱ leaving the room]
 L = [greeted everybody]
 M = **John**ⁱ
- c. K = [**John**ⁱ leaving the room]
 P = [**John**ⁱ greeted everybody]
- d. [_{vP} [_{vP} **John**ⁱ greeted everybody] [before **John**ⁱ leaving the room]]
- e. [_{TP} **John**ⁱ T [_{vP} [_{vP} **John**ⁱ greeted everybody] [before **John**ⁱ leaving the room]]]
- f. [_{TP} **John**ⁱ T [_{vP} [_{vP} ~~**John**~~ⁱ greeted everybody] [before ~~**John**~~ⁱ leaving the room]]]

Given the syntactic objects K and L in (13a), the computational system makes a copy of *John* from K (cf. (13b)) and merges it with L (cf. (13c)), an instance of sideward movement that allows the external θ -role of the matrix clause to be discharged. After *before* merges with K and the resulting PP adjoins to vP (cf. (13d)), the subject moves to [Spec, TP] (cf. (13e)), the lower copies of *John* are deleted in the phonological component (cf. (13f)), and the structure surfaces as (12).

In section 1 we saw that sideward movement makes it possible for head adjunction to comply with the Extension Condition. In this section we have seen that sideward movement also provides a straightforward analysis for multiple gap constructions where, descriptively speaking, an expression appears to be simultaneously moving from more than one position as more than one chain is formed. In the case of the parasitic gap and adjunct control constructions discussed above, for instance, it looks as if in (10), *which paper* has moved from the two object positions and in (12), *John* in the matrix [Spec,TP] has moved from the matrix [Spec,vP] and the embedded

subject position. From standard Move-based approaches, this is simply not a possibility. However, once Move is reinterpreted as Copy plus Merge, the derivation of multiple gap constructions such as parasitic gap and adjunct control constructions is not different from standard instances of movement. That is, like the derivation of standard “upward” movement, the derivation of parasitic gap and adjunct control constructions also involves applications of Copy and Merge. The only (irrelevant) difference is that in instances of sideward movement, the copy created merges not with the syntactic object that contains the source of the copying, but with another root syntactic object that is available to the computational system. It is therefore an empirical virtue of approaches that do not enrich the computational apparatus by (explicitly or implicitly) excluding sideward movement that they can provide a uniform treatment for movement operations that result in single or multiple chains.

Of course, one must also show that applications of Copy and Merge yielding sideward movement do not overgenerate. But this is no different a task than what must be done with respect to “upward” movement. I show below that the same conditions that block unwanted instances of upward movement can be used to prevent overgeneration in the case of sideward movement.

3. Preventing overgeneration

Once Move is reinterpreted in terms of Copy and Merge, all the conditions that were taken to regulate Move should accordingly be understood as holding of Copy, Merge, or the (chain) relation established among the copies. Note that this should be so regardless of whether we are dealing with upward or sideward movement. For instance, if applications of Move were required to satisfy Last Resort and Minimality, so is the

interaction between Copy and Merge. Below I show how conditions that were taken to constrain Move can be used to rule out unwanted instances of sideward movement.

3.1. Last Resort

Exploring general least effort guidelines, Chomsky (1995) proposes that every syntactic operation must be motivated, that is, every syntactic operation must be subject to Last Resort. Moreover, in consonance with the general attempt to reduce the computational complexity of derivations, Last Resort must be computed in a local manner (see e.g. Collins 1997). With this in mind, let us consider the contrast in (14), which illustrates the well known fact that parasitic gaps can be licensed by arguments but not by adjuncts (see e.g. Postal 1993).

- (14) a. [which paper]_i did you file t_i after John read PG_i?
 b. *how_i did Deborah cook the pork t_i after Jane cooked the chicken PG_i?

From a sideward movement approach, the derivation of (14a) involves sideward movement of *which paper*, as shown in (15) below, and (14b), sideward movement of *how*, as shown in (16). The question is why the latter is not licensed.

- (15) a. K = [John read [**which paper**]]
 L = file
 b. K = [John read [**which paper**]ⁱ]
 M = [file [**which paper**]ⁱ]

(16) a. K = [Jane cooked the chicken **how**]

L = [Deborah cook the chicken]

b. K = [Jane cooked the chicken **howⁱ**]

M = [Deborah cook the chicken **howⁱ**]

As discussed by Hornstein and Nunes (2002) and Nunes (2004), sideward movement satisfies Last Resort in (15), but not in (16). More specifically, the copying of *which paper* in (15) is triggered by θ -considerations: *file* must assign its θ -role and this convergence requirement licenses the copying of *which paper*. By contrast, there is no comparable requirement in L in (16a) that could trigger the copying of *how*.

Although it is quite reasonable to say that *file* in (15a) needs an argument, it makes no sense to say that L in (16a) needs an adjunct. Once copying of *how* in (16a) is not (locally) licensed, the parasitic gap construction in (14b) is correctly excluded.⁵

In sum, the Copy operation underlying sideward movement is not different from the one underlying upward movement: both must (locally) comply with Last Resort.

3.2. Derivational timing and the directionality of sideward movement

Let us now consider the contrast in (17) below. (17a) involves the sideward movement depicted in (15). In turn, the derivation of (17b) requires the instance of sideward movement shown in (18). The derivational step in (18) cannot be excluded by Last Resort, for the θ -requirements of *file* can license the copying of *which paper*, as we saw in (15). Thus, it must be the case that *which paper* in (18a) is not accessible to Copy at the derivational step where *file* could have its θ -requirements satisfied.

- (17) a. [which paper]_i did you file t_i after John read PG_i?
 b. *[which paper]_i did you file t_i after John left the room without reading PG_i

- (18) a. K = [John [_{VP} [_{VP} left the room] [_{PP} without reading [**which paper**]]]
 L = file
 b. K = [**John** [_{VP} [_{VP} left the room] [_{PP} without reading [**which paper**]]]
 M = [file [**which paper**]]ⁱ

The puzzling contrast in (17) finds a straightforward answer if the computation works in a bottom-up and phase-by-phase fashion, as currently assumed (see e.g. Chomsky 2000, 2001, 2004). Assuming phase-based computations, movement must proceed from more to less embedded domains. This is indeed the case in both (15) and (18). However, there is a crucial difference between these two (see Nunes and Uriagereka 2000, Hornstein 2001, Nunes 2001, 2004, and Hornstein and Nunes 2002). In (18), movement/copying targets an expression that is inside an adjunct. Regardless of how one implements adjunct islands in minimalist terms, such movement should induce an island effect; hence the unacceptability of (17b). By contrast, in (15) no element containing *which paper* is an adjunct. Crucially, *adjunct* is not an absolute, but relational notion: a given expression is an *adjunct of* another. In (15a) K is just a root syntactic object. The fact that later on K will become an adjunct is irrelevant at the derivational step where movement takes place. In other words, there is no island configuration in (15) that would prevent copying. In fact, the copying seen in (15) is no different from the copying found in licit instances of upward movement (cf. (7)): in both circumstances, copying proceeds from a configuration that is not an island.

One could ask why copying of *which paper* in the derivation of (17b) cannot take place *before* the relevant PP becomes an adjunct, as illustrated in (19) below. In (19a) K is not an adjunct and therefore *which paper* is accessible for copying. If it is indeed copied, it may later merge with *file* and there would be no island violation, which would incorrectly rule (17b) in.

- (19) a. K = [reading [**which paper**]]
 L = [John left the room]
- b. K = [reading [**which paper**]ⁱ]
 L = [John left the room]
 M = [**which paper**]ⁱ
- c. P = [_{VP} [_{VP} John left the room] [_{PP} without reading [**which paper**]ⁱ]]
 M = [**which paper**]ⁱ
- d. P = [_{VP} [_{VP} John left the room] [_{PP} without reading [**which paper**]ⁱ]]
 M = [**which paper**]ⁱ
 Q = file
- e. P = [_{VP} [_{VP} John left the room] [_{PP} without reading [**which paper**]ⁱ]]
 R = [file [**which paper**]ⁱ]

Assuming that Last Resort must be computed in a local fashion, the derivation depicted in (19) violates Last Resort. Notice that in (19a) there is no motivation for copying *which paper*; hence, movement in (19b) is ruled out by Last Resort at this step. To put it in more general terms, if the computational system cannot resort to look-ahead (across phases), it cannot idly create a copy and leave it hanging around until it can be used. Triggers for copying must be locally available. Again, this is no different from

upward movement: standard adjunct island violations could be incorrectly circumvented if copies could be freely created and kept in stock for later use.

Another question that arises regarding the derivation of (17b) is why *which paper* cannot move from the object position of *file* to the object position of *reading*, as sketched in (20) below. Notice that in (20a) *which paper* is not within an adjunct and its movement in (20b) would satisfy Last Resort as *reading* would assign its internal θ -role.

(20) a. K = [file [**which paper**]]

L = reading

b. K = [file [**which paper**]ⁱ]

M = [reading [**which paper**]ⁱ]

The derivation sketched in (20) is excluded if syntactic computations must proceed from more to less embedded domains, an assumption that is independently made in phase-based approaches to reduce computational complexity. In a well-behaved computation, the system first builds the (embedded) adjunct clause before activating the matrix derivational workspace (cf. (18a)). However, this is not the case in (20), for the matrix domain is activated before the most embedded domain is completed. To put it in general terms, the assumption that derivations unfold from more to less embedded domains has the effect that in the specific case of sideward movement, it must proceed from a “will-be” adjunct to the matrix derivational domain and not vice-versa.⁶ This assumption not only rules out the derivational step in (20), which would incorrectly allow (17b), but also makes interesting empirical predictions.

Suppose, for instance, that a given expression becomes inert for purposes of movement in some specific configuration. If derivations proceed from more to less embedded domains, the prediction is that a freezing configuration for an application of sideward movement may be found in the more embedded domain (the launching site), but not in the less embedded domain (the target of movement). Hornstein and Nunes (2002) and Nunes (2004) argue this is what is behind contrasts such as (21), originally noted by Postal (1993).

- (21) a. This is the book *which*_i I was [[given *t*_i by Ted] [after reading PG_i]]
 b. *This is the book *which*_i I [[read *t*_i] [before being given PG_i by Ted]]

In order for the parasitic gap constructions in (21) to be derived under a sideward movement analysis, the computational system must copy *which* from the more to the less embedded domain, as respectively illustrated in (22) and (23) (irrelevant details omitted).

- (22) a. K = [reading **which**]
 L = given
 b. K = [reading **which**ⁱ]
 L = [given **which**ⁱ]
- (23) a. K = [being given **which** by Ted]
 L = read
 b. K = [being given **which**ⁱ by Ted]
 M = [read **which**ⁱ]

In (22) and (23) we have sideward movement involving the object positions of *give* and *read*, the only difference being whether they sit in an embedding or embedded domain. When *read* sits in embedded domain, as in (22), sideward movement yields a licit result (cf. (21a)), but when *given* does, as in (23), the result is not well formed (cf. (21b)).

Assuming that the theme of double object constructions receives inherent Case and that inherent Case renders an element inert for purposes of A-movement, Hornstein and Nunes argue that the contrast in (21) independently follows from the directionality of sideward movement. In (23a), *which* is assigned inherent Case by *given* and is therefore frozen for purposes of (sideward) A-movement; hence the unacceptability of (21b). By contrast, in (22) *which* receives inherent Case only after sideward movement takes place. Hence, the movement in (22) may lead to a well formed result (cf. (21a)). Crucially, if sideward movement could proceed from embedding to embedded domains, (21b) would be incorrectly allowed in a derivation employing the steps in (24) below where the embedding domain is activated before the embedded domain is completed. Thus, the contrast in (21) provides independent support for the assumption that derivations proceed from more to less embedded domains and the corresponding directionality of sideward movement.

(24) a. K = [read **which**]

L = given

b. K = [read **which**ⁱ]

M = [given **which**ⁱ]

To summarize, if derivations are to proceed in a phase-by-phase fashion and if look-ahead must be minimized, sideward movement (like upward movement) becomes quite constrained and a good number of unwanted derivations involving sideward movement are excluded based on the way derivations unfold (from embedded to embedding domains) and the derivational timing of the Copy operation. An application of Copy is licensed by Last Resort only if its trigger is available to the computation at the derivational step where Copy takes place.

3.3. Deletion of copies and linearization of chains

One question that arises in any version of the copy theory of movement is why (in general) it is only one copy that surfaces at PF. Why must the structure in (25), for instance, surface as (26a) and not (26b)?

(25) [**John**ⁱ [was [arrested **John**ⁱ]]]

- (26) a. John was arrested.
 b. *John was arrested John.

Nunes (1995, 1999, 2004) argues that linearization considerations prevent (25) from surfacing as (26b). The gist of the proposal is the following. A chain is a discontinuous element, occupying different positions at a time. A PF object, on the other hand, is a linear string. Thus, if the system attempts to realize the whole chain at PF, no linear order will obtain. Consider the linearization of (25) without deletion, for instance. Assuming that linearization is guided by (some version of) Kayne's (1994)

LCA, *John* must precede *was* because the upper copy of *John* asymmetrically c-commands *was*; by the same token, *was* must precede *John* as it asymmetrically c-commands the lower copy of *John*. However, these two instances of *John* are nondistinct (they relate to the same material in the initial numeration); thus, the linearization of (25) without an application of deletion yields the contradictory requirement that *John* must precede and be preceded by *was*. Likewise, *John* would be required to precede itself as the higher copy asymmetrically c-commands the lower one. Nunes (1995) proposes that the deletion of chain links, which he refers to as *Chain Reduction*, allows the computational system to linearize structures containing chains. If the lower copy of *John* in (25) is deleted, for example, the structure can be trivially linearized as (26a) and no contradiction arises.

Assuming that deletion of copies is performed by Chain Reduction, Nunes (1995, 2001, 2004) further argues that linearization considerations also rule out unwanted instances of sideward movement. Consider, for instance, the well known contrast in (27), which in GB was taken to show that parasitic gaps must be licensed at S-Structure (see e.g. Chomsky 1982).

- (27) a. [which paper]_i did you file t_i without reading PG_i?
 b. *Who filed [which paper]_i without reading PG_i?

Under a sideward movement analysis, the derivation of either construction in (27) involves a licit application of sideward movement from the object position of *reading* to the object position of *filed*, as shown in (28).

- (28) a. K = [reading **which paper**]
 L = file/filed
- b. K = [reading **which paper**ⁱ]
 M = [file/filed **which paper**ⁱ]

However, the derivations underlying the sentences in (27) differ after their final structures are submitted to linearization. In the structure associated with (27a) given in (29a) below, two chains can be formed: CH₁ = (copy¹, copy²) and CH₂ = (copy¹, copy³). Crucially, no chain can be formed between copy² and copy³ due to lack of c-command between them. Applying to CH₂, Chain Reduction deletes copy³, yielding (29b). Applying to CH₁, Chain Reduction deletes copy² (cf. (29c)) and the structure surfaces as (27a) after it is linearized.

- (29) a. [[**which paper**]¹ [did you [[file [**which paper**]²] [after reading
 which paper]³]]]]
- b. [[**which paper**]¹ [did you [[file [**which paper**]²] [after reading
 ~~which paper~~³]]]]]
- c. [[**which paper**]¹ [did you [[file ~~which paper~~²] [after reading
 ~~which paper~~³]]]]]

By contrast, Chain Reduction cannot apply to (30) below (the structure underlying (27b)), because no chain can be formed between the two copies of *which paper*. Once Chain Reduction is inapplicable, the two nondistinct copies of *which paper* prevent the structure from being linearized for basically the same reason (25) cannot be linearized:

without, for instance, is subject to the contradictory requirement that it must precede and be preceded by *which paper* (Recall that the two copies are nondistinct).

(30) [who [[filed [**which paper**]ⁱ] [without reading [**which paper**]ⁱ]]]

To sum up, sideward movement is drastically constrained by linearization considerations. Its output yields an acceptable result only if further computations allow an additional copy to form an independent chain with each of the copies related by sideward movement. That is the case in (29a) (see also (11g), (13e)), but not in (30).

3.4. Summary

In this section we saw how sideward movement can be adequately constrained so that it does not overgenerate. The discussion was not meant to be exhaustive (see Nunes 2001, 2004 for further discussion). Rather, the point was to illustrate that the same conditions that regulate “upward” movement regulate sideward movement.

4. Sideward movement and noncanonical phonetic realization of copies

So far we have seen canonical instances of upward and sideward movement, where the highest link of the relevant chain is kept and the lower copies are deleted in the phonological component. However, an increasing body of literature has documented cases where it is the highest chain link that is deleted and even cases where more than one chain link is phonetically realized.⁷ Below I show that the general circumstances that allow these two types of exceptions also affect chains that result from sideward

movement, as should be expected if the output of both upward and sideward movement is subject to the same conditions on PF realization.

4.1. Pronunciation of lower copies

Recall from section 3.3 that Chain Reduction deletes lower chain links in order to allow linearization of structures containing chains. The question is why (in the general case) Chain Reduction does not delete the head of the chain and keep a lower copy for phonetic realization. Given the structure in (31), for instance, why must it surface as (32a) and not as (32b)?

(31) [**John**ⁱ [was [arrested **John**ⁱ]]]

- (32) a. John was arrested.
 b. *Was arrested John.

Nunes (1995, 1999, 2004) proposes that the general pattern illustrated in (32) follows from economy considerations. Roughly speaking, as movement allows feature checking/valuation, higher copies have more features checked/valued than lower copies. Thus, all things being equal, the system generally keeps the highest copy as it is the more optimal copy for PF realization for having the greatest number of features checked/valued. However, given that this is a choice based on economy considerations, in case independent convergence requirements are violated if the highest copy is phonetically realized, Chain Reduction deletes the highest copy and keeps the second highest copy.⁸

A clear case of lower copy pronunciation is presented by the contrasts in (33) and (34) below, as discussed by Bošković (2002). (33) shows that Romanian is a multiple *wh*-fronting language. However, the object *wh*-phrase does not appear to move if it is homophonous with the fronted subject *wh*-phrase, as shown in (34). Bošković proposes that Romanian has a low-level PF constraint against adjacent homophonous *wh*-phrases, which rules out (34b). As for the exceptional pattern in (34a), Bošković argues that it also involves multiple *wh*-fronting in the syntactic component, but in order to comply with the PF constraint on adjacent homophonous elements, the higher copy of the object *wh*-phrase is deleted and the lower one is pronounced instead, as sketched in (35) (irrelevant details omitted).

(33) *Romanian* (Bošković 2002):

a. Cine ce precede?

who what precedes

b. *Cine precede ce?

who precedes what

‘Who precedes what?’

(34) *Romanian* (Bošković 2002):

a. Ce precede ce?

what precedes what

b. *Ce ce precede?

what what precedes

‘What precedes what?’

(35) [_{CSUBJ} ~~ce~~OBJ [_{CSUBJ} precede ~~ce~~OBJ]]

Let us now consider instances of lower copy pronunciation in constructions involving sideward movement. We have seen in section 3.3 that Chain Reduction operates with chains and this is what accounts for why a parasitic gap construction such as (36a) below cannot be associated with the structure in (36b), which is derived by sideward movement of *which paper*. Given that the two copies in (36b) do not form a chain, Chain Reduction is inapplicable and the structure cannot be linearized. Combined with the approach on lower copy pronunciation presented above, this minimalist reanalysis of the S-Structure condition on parasitic gap licensing makes the prediction that a construction superficially similar to (36a) should be well formed if the object *wh*-phrase cannot surface in the fronted position. Bošković (2002) shows that this prediction is indeed borne out, as illustrated by (37).

- (36) a. *Who filed [which paper]_i without reading PG_i?
 b. [who [[filed [**which paper**]ⁱ] [without reading [**which paper**]ⁱ]]]

(37) *Romanian* (Bošković 2002):

Ce precede **ce**_i fara sa influenteze **PG**_i?

what precedes what without SUBF.PRT influence.3.SG

‘What precedes what_i without influencing it_i?’

Following Bošković, we assume that the object *wh*-phrase undergoes *wh*-fronting in the overt component, as is the standard case in Romanian, yielding the simplified

structure in (38), where an instance of *ce* moves from the object position of *influenzeze* to the object position of *precede* before undergoing *wh*-fronting.

(38) [_{ceSUBJ} **ce**³ [[precede **ce**²] [fara sa influenteze **ce**¹]]]

In (38) the fronted *wh*-object forms the chain CH₁ = (copy³, copy¹) and CH₂ = (copy³, copy¹).⁹ Applying to CH₁, Chain Reduction deletes the lower copy of *ce*, yielding (39a) below. By contrast, if Chain Reduction deletes the lower copy of *ce* when applying to CH₂, the derivation will not converge due to adjacency between the homophonous subject and object *wh*-phrases. In order to circumvent this problem, Chain Reduction deletes the higher copy, as shown in (39b), and the structure surfaces as (37), which superficially seems to involve a parasitic gap licensed by a *wh*-phrase *in situ*.¹⁰

(39) a. [_{ceSU} **ce**³ [[precede **ce**²] [fara sa influenteze **ce**¹]]]

b. [_{ceSU} **ce**³ [[precede **ce**²] [fara sa influenteze **ce**¹]]]

Another instance of lower copy pronunciation in constructions involving sideward movement is discussed by Dotlačil (2008) with respect to ATB movement in (noncolloquial) Czech. Nunes (1995, 2001, 2004) and Hornstein and Nunes (2002) argue that ATB extraction also involves sideward movement. Assuming this to be the case, Dotlačil discusses a curious case of ATB extraction of clitics in Czech. First, consider the contrast in (40).

(40) *Czech* (Dotlačil 2008):

a. *Zavolal **jsem** **Petra** a představil známým.

called aux_{1SG} Petr_{acc} and introduced friends

b. **Petra** **jsem** zavolala představil známým.

Petr_{acc} aux_{1SG} called and introduced friends

‘I called Petr and introduced him to friends.’

(40a) shows that Czech does not allow auxiliary gapping and object drop under coordination. The acceptability of (40b) in turn shows that missing auxiliaries and objects may be licensed if they can be analyzed in terms of ATB extraction; that is, in (40b) the object *Petra* and the auxiliary have undergone ATB extraction (via sideward movement) and bind a trace/copy in each of the conjuncts. That being so, let us now examine (41) below, which differs minimally from (40a) in that we have a pronominal clitic in place of *Petra*. Clearly, (41) cannot involve auxiliary gapping and object omission in the second conjunct; otherwise, (40a) should also be acceptable. At first sight, (41) could be derived by ATB extraction of the auxiliary and the pronominal clitics (like (40b)), followed by movement of *zavolal* to a higher position. However, the latter movement would violate the Coordination Structure Constraint, as *zavolal* would be moving from just one conjunct. So, how can (41) be derived?

(41) *Czech* (Dotlačil 2008):

Zavolal **jsem** **ho** a představil známým.

called aux_{1SG} him_{acc} and introduced friends

‘I called him and introduced friends.’

Dotlačil (2008) argues that (41) behaves exactly like (40b) in the syntactic component, that is, the auxiliary and the pronominal clitics undergo ATB extraction (via sideward movement), yielding the simplified structure in (42a) below (with English words for presentational purposes), where the highest copy of the clitic cluster forms a different chain with each of the lower copies. Reduction of the chain involving the second conjunct deletes the lower copy, as shown in (42b) (see footnote 9). Similar reduction of the chain involving the first conjunct would leave the clitic cluster in clause initial position, which is not allowed in (noncolloquial) Czech. Chain Reduction then deletes the higher copy, as shown in (42c). As Dotlačil shows, rather than being an apparent exceptional pattern of ATB extraction, (41) is another case of pronunciation of lower copies.

- (42) a. [**aux-him** [[called **aux-him**] and [introduced **aux-him**]]
 b. [**aux-him** [[called **aux-him**] and [introduced ~~aux-him~~]]
 c. [~~aux-him~~ [[called **aux-him**] and [introduced friends **aux-him**]]

To summarize, the above discussion shows that mismatches between syntactic structures and PF realization that can be solved by appealing to phonetic realization of lower copies do not distinguish upward from sideward movement. Again, this is good news for an approach that treats them in a uniform manner as different instantiations of Copy and Merge.

4.2. Pronunciation of multiple copies

We have seen in section 3.3 that linearization considerations in general prevent a chain from surfacing with all of its links phonetically realized. However, this account predicts that if for some reason a given link becomes invisible to the relevant linearization procedure, a chain may surface with more than one link realized at PF. Nunes (1999, 2004) argues that this may happen if a given copy is morphologically fused (in the sense of Halle and Marantz 1993) with another element. The reasoning goes as follows: if a given copy C fuses with a given element E in the morphological component, the blended result #C-E# (or #E-C#) behaves like an atomic vocabulary item with no internal structure accessible to further morphological or syntactic computations. In particular, the fused copy is no longer visible to linearization (the resulting atomic element is) and no contradictory requirement with respect to other copies will arise. Verb clefting constructions in Vata, as illustrated in (43), can exemplify this process.

(43) *Vata* (Koopman 1984):

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

Koopman (1984) shows that the two verbal occurrences in (43) cannot be separated by islands, which indicates that they should be related by movement. In terms of the copy theory, the verbal instances seen in (43) can then be analyzed as copies produced by the movement operation. The question that arises is why the presence of more than one copy does not create contradictory requirements for linearization. Nunes

(2004) proposes that 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 by the shaded material in (44a) below. Of the three verbal copies in (44a), the LCA only sees the lower two after the highest copy gets fused with Foc^0 . The lowest copy is then deleted (cf. (44b)) and the structure is linearized as in (43), with two copies of the verb phonetically realized. Evidence for morphological fusion in Vata verbal clefting is provided by the fact that the fronted verb cannot occur with tense or negative particles (see Koopman 1984), which makes sense if these particles render the verb morphologically too complex, thereby preventing the verb from undergoing fusion with the focus head.

(44) 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:*

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

With this overall picture in mind, one wonders if sideward movement can also yield outputs with more than one copy phonetically realized. In his detailed analysis of control in Telugu and Assamese, Haddad (2007, 2009) shows that constructions such as (45) and (46) 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.¹¹

(45) *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.’

(46) *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 (47) below.

(47) *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.’

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. 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. Of course, if a given copy is syntactically complex, i.e. it is phrasal, as in (47), it is also morphologically complex and not a good candidate to undergo morphological fusion. In turn, once fusion is prevented from applying, the presence of more than one copy induces a violation of the LCA.

To sum up, we again see that the exceptional circumstances and conditions that regulate acceptable instances of pronunciation of more than one copy apply to upward and sideward movement indistinctively.

5. *Wh*-movement and adjunct control

A general property of adjunct control constructions like (48) below, for instance, is that PRO must be controlled by the subject and not the object of the next higher clause.

Hornstein (2001) proposes that this subject-object asymmetry results from an economy choice at the derivational step sketched in (49).

(48) John_i saw Mary_k after PRO_{i/*k} eating lunch

(49) $N = \{\text{John}_0, \text{saw}_0, \text{Mary}_1, \text{after}_1, \text{eating}_0, \text{lunch}_0\}$

$K = [\text{John eating lunch}]$

$L = \text{saw}$

In (49), *saw* must assign its internal θ -role and there are two potential candidates to receive it: *Mary*, which is still in the numeration, and *John* in the subject position of the gerundive clause. If *Mary* is selected and merged with *saw*, as seen in (50a) below, the derivation converges as a subject control structure, after *John* undergoes sideward movement to [Spec,VP] (cf. (50b)).

(50) a. $N' = \{\text{John}_0, \text{saw}_0, \mathbf{Mary}_0, \text{after}_1, \text{eating}_0, \text{lunch}_0\}$

$K = [\text{John eating lunch}]$

$M = [\text{saw } \mathbf{Mary}]$

b. $N' = \{\text{John}_0, \text{saw}_0, \text{Mary}_0, \text{after}_1, \text{eating}_0, \text{lunch}_0\}$

$K = [\mathbf{John}^i \text{ eating lunch}]$

$P = [\mathbf{John}^i \text{ saw } \text{Mary}]$

c. $[_{TP} \text{John}^i [_{VP} [_{VP} \text{John}^i \text{ saw } \text{Mary}] [_{PP} \text{after } \text{John}^i \text{ eating lunch}]]]$

On the other hand, if *John* is copied and merged with *saw*, as shown in (51a) below, the derivation should in principle converge as well, this time yielding an object control structure after *Mary* is plugged in as the external argument, as shown in (51b). Under the assumption that Merge is more economical than Move (see Chomsky 1995), Hornstein observes that the derivation in (50) is more economical than (51); hence, the subject-object asymmetry in (48).¹²

- (51) a. $N = \{\text{John}_0, \text{saw}_0, \text{Mary}_1, \text{after}_1, \text{eating}_0, \text{lunch}_0\}$
 $K = [\text{after } \mathbf{John}^i \text{ eating lunch}]$
 $M = [\text{saw } \mathbf{John}^i]$
- b. $N' = \{\text{John}_0, \text{saw}_0, \mathbf{Mary}_0, \text{after}_0, \text{eating}_0, \text{lunch}_0\}$
 $K = [\text{after } \text{John}^i \text{ eating lunch}]$
 $P = [\mathbf{Mary} \text{ saw } \text{John}^i]$
- c. $[_{TP} \text{Mary}^k [_{VP} [_{VP} \text{Mary}^k \text{ saw } \mathbf{John}^i] [_{PP} \text{after } \mathbf{John}^i \text{ eating lunch}]]]$

Although this preference for subject over object control may be the general case, convergence requirements may lead to the opposite situation. This can be seen in the interaction between *wh*-movement and adjunct control in (Brazilian and European) Portuguese. In these languages, the subject of infinitival adjunct clauses may be controlled by the matrix subject or the matrix object, depending on whether or not the matrix object undergoes *wh*-movement, as illustrated in (52) below.¹³ (52b) has a *wh*-in situ in the matrix clause and the result is subject control, as in (52a), with no *wh*-element involved. By contrast, (52c) has *wh*-movement and now both subject and object control are possible.

(52) *Portuguese:*

- a. $[\text{Os alunos}]_i \text{ entrevistaram } [\text{os professores}]_k \text{ antes de } \text{PRO}_{i/*k} \text{ sair de férias.}$
the students interviewed the professors before of leave of vacation
 ‘The students interviewed the professors before leaving on vacation.’

- b. [Os alunos]_i entrevistaram [que professores]_k antes de PRO_{i/*k} sair de
the students interviewed which professors before of leave of
 férias?

vacation

‘Which professors did the students interview before leaving on vacation?’

- c. [Que professores]_k é que [os alunos]_i entrevistaram t_k antes de PRO_{i/k}
which professors is that the students interviewed before of
 sair de férias?

leave of vacation

‘[Which professors]_k did [the students]_i interview before they_{i/k} left on
 vacation?’

Assuming with Bošković (2007) that the strong feature that triggers successive cyclic movement (*uF*) is hosted by the moving element, I have proposed (see Nunes 2010) that in languages like Brazilian and European Portuguese, with optional *wh*-movement, this feature is lexically optional on *wh*-elements.¹⁴ Moreover, the presence of this feature in the derivation has consequences for economy computations regarding Merge-over-Move. Recall that subject over object control is enforced in adjunct control constructions due to Merge being more economical than Move (see footnote 12). In the case of (52a), for instance, if *os alunos* ‘the students’ is in the subject position of the adjunct clause, it cannot undergo sideward movement to the complement of the matrix verb, for merger of *os professores* ‘the professors’ in this position is more economical. So, after *os professores* is merged, *os alunos* can only move to the matrix [Spec, vP], yielding subject control (cf. (49)-(50)).

Bearing this in mind, the object control reading of (52b) and (52c) should involve the derivations sketched in (53a) and (53b), respectively.

(53) *Portuguese:*

- a. *Os alunos [[entrevistaram [**que_{uF}** professores]_i] [antes de t_i sair de férias]]?
the students interviewed which professors before of leave of vacation
 ‘[Which professors]_i did the students interview before they_i left on vacation?’
- b. [**que_{vF}** professores]_i é que os alunos [[entrevistram t_i] [antes de t_i sair de férias]]?
which professors is that the students interviewed before of leave of vacation
 ‘[Which professors]_i did the students interview before they_i left on vacation?’

The *wh*-element of both derivations in (53) enters the numeration specified with a strong feature *uF*, which in turn requires that the *wh*-phrase must move if possible. This requirement of the strong feature now overrules Merge-over-Move, for things are not equal anymore. If the *wh*-element sits in the subject of the adjunct clause and sideward movement to the matrix object position is possible, such movement must take place. Now, if Merge-over-Move is circumvented in the presence of a strong feature *Uf*, this strong feature must be checked. Hence, (53a) is unacceptable not because movement of the *wh*-element from the adjunct clause to the matrix object position violates Merge-over-Move, but because the strong feature of the *wh*-phrase remained unchecked. When it is checked by moving to [Spec,CP], as in (53b), the derivation converges, yielding an object control reading.¹⁵

To sum up. Focusing on successively cyclic (upward) movement, Bošković (2007) has argued that the computational complexity of syntactic derivations gets substantially minimized if the feature that drives movement is borne by the moving element rather than the probe. I have shown above that this general proposal not only extends to sideward movement, but also makes the correct empirical cut regarding the interaction between *wh*-movement and adjunct control in languages with generalized optional *wh*-movement such as (Brazilian and European) Portuguese.

6. Morphological sideward movement

Let us finally consider syntax-phonology mismatches involving preposition duplication in colloquial Brazilian Portuguese (henceforth *BP*). Take the data in (54) and (55), for instance.

(54)a. *Eu pensei **em o** João. *[formal/colloquial BP]*

I thought in the João

‘I thought about João.’

b. Eu pensei **no** João. *[formal/colloquial BP]*

I thought in-the João

‘I thought about João.’

(55)a. Eu pensei **em o** João fazer esse trabalho. *[formal BP]*

I thought in the João do-INF this job

‘I think that João should do this job.’

b. Eu pensei **no** João fazer esse trabalho. [*colloquial BP*]

I thought in-the João do-INF this job

‘I think that João should do this job.’

(54) shows that in BP the preposition *em* ‘in’ and the definite article *o* ‘the’ must contract when they are adjacent. In turn, (55) shows that if the definite article belongs to the embedded subject, we have contraction in colloquial BP, but not in its formal registers. Nunes and Ximenes (2009) analyze the difference between (55a) and (55b) as arising from two different structures. In formal registers of BP, the Case-marking preposition *em* precedes the whole infinitival CP, as shown in (56) below, and in this circumstance it is not adjacent to the determiner due to the intervention of C; lack of adjacency then yields lack of contraction (cf. (55a)). As for colloquial BP, Nunes and Ximenes argue that the preposition is realized as C, which renders it adjacent to the determiner, as sketched in (57), and contraction is obligatory (cf. (55b)).

(56) *Formal BP:*

[... X [**P** [_{infinitival-CP} C [_{TP} [_{DP} **D** ...]]]]]

(57) *Colloquial BP:*

[... X [_{infinitival-CP} **P/C** [_{TP} [_{DP} **D** ...]]]]

A very puzzling paradigm arises in colloquial BP when the contraction patterns depicted in (54)-(55) are combined with coordination, as illustrated in (58) and (59) (see Ximenes 2002, 2004, Ximenes and Nunes 2004, and Nunes and Ximenes 2009).

(58)a. *Eu pensei **no** João e **a** Maria. [formal/colloquial BP]

I thought in-the João and the Maria

‘I thought about João and Maria.’

b. Eu pensei **no** João e **na** Maria. [formal/colloquial BP]

I thought in-the João and in-the Maria

‘I thought about João.’

(59)a. Eu pensei **em o** João e **a** Maria [formal BP]

I thought in the João and the Maria

fazerem esse trabalho.

do-INF.3PL this job

‘I think that João and Maria should do this job.’

b. *Eu pensei **em o** João e **em a** Maria

I thought in the João and in the Maria

fazerem esse trabalho.

[formal/colloquial BP]

do-INF.3PL this job

‘I think that João and Maria should do this job.’

c. Eu pensei **no** João e **na** Maria [colloquial BP]

I thought in-the João and in-the Maria

fazerem esse trabalho.

do-INF.3PL this job

‘I think that João and Maria should do this job.’

(58) shows that contracting prepositions must be repeated if one of the conjuncts has a determiner that triggers contraction. This suggests that the Parallelism Requirement on

coordinated structures (see e.g. Chomsky 1995, Fox 2000, and Hornstein and Nunes 2002) also applies to the morphological component. That is, once contraction appears in one conjunct, it must appear in every conjunct. Thus, at first sight, (58) can converge only if there are two prepositions in the underlying numeration and the PPs headed by these preposition are accordingly coordinated, as sketched in (60).

(60) [Eu pensei [[PP **no** João] e [PP **na** Maria]]

I thought in-the João and in-the Maria

However, this account cannot be extended to the full paradigm of (59). The lack of contraction between the subcategorizing preposition and the determiner of the subject in (59a) is not surprising. It just replicates the pattern seen in (55a), which in formal BP is associated with the structure in (56), where the null complementizer breaks the adjacency between the preposition and the determiner. The ungrammaticality of (59b) in either register is not mysterious either. Given that the coordinated subject must involve DP coordination as it is interpreted as the agent of the embedded verb and triggers plural agreement on the inflected infinitival, (59b) is out due to the presence of a spurious preposition in the second conjunct. If a coordination involving PP is not an option for (59b), the question then is why a sentence analogous to (59b) becomes acceptable in colloquial BP if the prepositions get contracted with the relevant determiners (cf. (59c)).

Nunes and Ximenes (2009) (see also Ximenes 2002, 2004 and Ximenes and Nunes 2004 for discussion) argue that (59c) results from the application of sideward movement in the morphological component, that is, an application of copying¹⁶ and morphological merger. They propose that (59c) indeed involves coordination of DPs, as expected, and

that the second preposition is inserted in the morphological component. More specifically, if we have morphological merger (see Halle and Marantz 1993) in the boundary of one conjunct, the Parallelism Requirement triggers morphological merger in all conjuncts. The derivation of (59c), for instance, proceeds along the lines of (61).

(61)a. *Spell-Out:*

[... pensei [CP **em** [TP [_{andP} [DP **o** João] [_{and'} e [DP **a** Maria]]] fazerem...]]]

b. *Morphological merger:*

[... pensei [CP [TP [_{andP} [DP **em+o** João] [_{and'} e [DP **a** Maria]]] fazerem...]]]

c. *Copy and morphological merger:*

[... pensei [CP [TP [_{andP} [DP **emⁱ+o** João] [_{and'} e [DP **emⁱ+a** Maria]]]

fazerem...]]]

d. *Fusion:*

[... pensei [CP [TP [_{andP} [DP **no** João] [_{and'} e [DP **na** Maria]]] fazerem...]]]

Given that in colloquial BP, Case-marking prepositions are realized in C when they take infinitival complements (cf. (57)), the preposition *em* in (61a) is adjacent to the first determiner of the coordinated embedded subject in the spelled-out structure and morphological merger is obligatory in these circumstances, as seen in (61b). Once morphological merger affects the boundary of the coordinated subject, the Parallelism Requirement on coordinated structures kicks in and demands that the second conjunct also undergo morphological merger. Given that there is no preposition adjacent to the determiner of the second conjunct (recall that the embedded subject involves DP- and not PP-coordination), the preposition morphologically merged with the first conjunct is then copied and the resulting copy merges with the determiner of the second conjunct,

as shown in (61c). Finally, the prepositions and the determiners fuse, as shown in (61d), yielding the PF output in (59c), which at first glance appears to involve a quite exotic case of PP-coordination in a canonical subject position. Notice furthermore that the two copies of the preposition do not induce linearization problems as they become invisible for purposes of linearization after they fuse with the relevant determiners (see section 4.2).

To conclude, if the morphological component independently has the operations of copying and merger in its inventory, one should not be surprised if these operations interact and yield what may be viewed as sideward movement in the morphological component. After all, sideward movement is just a description of the interaction of these basic operations.¹⁷

7. Concluding remarks

In this paper I have argued that sideward movement is not a novel operation that makes the grammar's theoretical apparatus heavier. Sideward movement in fact arises as a natural consequence of the interaction among solid pillars of the minimalist enterprise: the attempt to eliminate noninterface levels, the postulation of syntactic entities in terms of lexical items, and the attempt to minimize derivational complexity, to name a few. In particular, once Move is interpreted within minimalism as the output of the interaction between the more basic operations of Copy and Merge, sideward movement is not a theoretical primitive, but is simply a mnemonic label for one particular instantiation of applications of Copy and Merge.

We have also seen that the same conditions that restrict upward movement also constrain sideward movement, again stressing the fact that under the copy theory, it

does not make sense to attempt to distinguish upward from sideward movement as *operations*, for Move is not understood as a primitive operation in the system. In other words, it is actually excluding sideward movement that requires complications in the theoretical apparatus and one has to ponder what is gained and what is lost with the required amendments.

Excluding sideward movement by brute force does not seem to lead to profitable gain from a minimalist perspective. One would be introducing redundancies in the system, as sideward movement is already quite constrained by the conditions that exclude unwanted instances of “upward” movement. Such exclusion would in fact face a considerable loss in empirical coverage. As discussed in the previous sections, sideward movement may not only be seen as cost free within a core set of minimalist assumptions, but also paves the way for sound analyses of intricate empirical phenomena. The empirical bar therefore requires that the exclusion of sideward movement should be accompanied by a uniform account of empirical phenomena such as the ones discussed here.

References

- Bobaljik, Jonathan (1995). *Morphosyntax: The Syntax of Verbal Inflection*. Doctoral dissertation, Massachusetts Institute of Technology.
- Bobaljik, Jonathan and Samuel Brown (1997). ‘Inter-arboreal Operations: Head-Movement and the Extension Requirement’, *Linguistic Inquiry* 28: 345-356.
- 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 (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 N. Corver and J. Nunes (eds.), *The Copy Theory of Movement*.
Amsterdam: John Benjamins, 13-74.
- Chomsky, Noam (1982). *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.
- Chomsky, Noam (1993). 'A Minimalist Program for Linguistic Theory', in K. Hale and S. J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, MA: MIT Press, 1-52.
- Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam (2000). 'Minimalist Inquiries: The Framework', in R. Martin, D. Michaels, and J. Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: MIT Press, 89-155.
- Chomsky, Noam (2001). 'Derivation by Phase', in M. Kenstowicz (ed.), *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press, 1-52.
- Chomsky, Noam (2004). 'Beyond Explanatory Adequacy', in A. Belletti (ed.), *Structures and Beyond*. Oxford: Oxford University Press, 104-131.
- Collins, Chris (1997). *Local economy*. Cambridge, MA: MIT Press.
- Corver, Norbert and Jairo Nunes (eds.) (2007). *The Copy Theory of Movement*.
Amsterdam: John Benjamins.

- Dotlačil, Jakub (2008). 'Across-the-Board Extraction and Clitic Omission in Czech', in A. Antonenko, J. Bailyn, and C. Bethin (eds.), *Proceedings of Formal Approaches to Slavic Linguistics 16: The Stony Brook Meeting*. Ann Harbor: Michigan Slavic Publications, 134-150.
- Fox, Danny (2000). *Economy and Semantic Interpretation*. Cambridge, MA: MIT Press.
- Haddad, Youssef (2007). Adjunct Control in Telugu and Assamese. Doctoral dissertation, University of Florida.
- Haddad, Youssef (2009). 'Copy control in Telugu', *Journal of Linguistics* 45: 69-109.
- Halle, Morris and Alec Marantz (1993). 'Distributed Morphology and the Pieces of Inflection', in K. Hale and S. J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, MA: MIT Press, 111-176.
- Hornstein, Norbert (1999). 'Movement and Control', *Linguistic Inquiry* 30: 69-96.
- Hornstein, Norbert (2001). *Move! A Minimalist Theory of Construal*. Oxford: Blackwell.
- Hornstein, Norbert (2009). *A Theory of Syntax*. Cambridge: Cambridge University Press.
- Hornstein, Norbert and Jairo Nunes (2002). 'On Asymmetries between Parasitic Gap and Across-the-Board Constructions', *Syntax* 5: 26-54.
- Kayne, Richard (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Koopman, Hilda (1984). *The Syntax of Verbs*. Dordrecht: Foris.
- Modesto, Marcello (2000). On the Identification of Null Arguments. Doctoral dissertation, University of Southern California.
- Niinuma, Fumikazu (2010). 'Across-the-Board and Parasitic Gap Constructions in Romanian', *Linguistic Inquiry* 41: 161-169.

- Nunes, Jairo (1995). *The Copy Theory of Movement and Linearization of Chains in the Minimalist Program*. Doctoral dissertation, University of Maryland at College Park.
- Nunes, Jairo (1999). 'Linearization of Chains and Phonetic Realization of Chain Links', in S. D. Epstein and N. Hornstein (eds.), *Working Minimalism*. Cambridge, MA: MIT Press, 217-249.
- Nunes, Jairo (2001). 'Sideward Movement', *Linguistic Inquiry* 31: 303-344
- Nunes, Jairo (2004). *Linearization of Chains and Sideward Movement*. Cambridge, MA: MIT Press.
- Nunes, Jairo (2010). 'Edge Features on Moving Elements: Evidence from Sideward Movement', paper presented at GLOW 33 (Generative Linguistics in the Old World), Universidade de Wrocław, 13-16/04/2010.
- Nunes, Jairo (2011). 'The Copy Theory', in Cedric Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism*. Oxford: Oxford University Press, 143-172.
- Nunes, Jairo and Juan Uriagereka (2000). 'Cyclicality and Extraction Domains', *Syntax* 3: 20-43.
- Nunes, Jairo and Cristina Ximenes (2009). 'Preposition Contraction and Morphological Sideward Movement in Brazilian Portuguese', in J. Nunes (ed.), *Minimalist Essays on Brazilian Portuguese Syntax*. Amsterdam: John Benjamins, 191-214.
- Postal, Paul (1993). 'Parasitic Gaps and the Across-the-Board Phenomenon', *Linguistic Inquiry* 24: 734-754.
- Rodrigues, Cilene (2004). *Impoverished Morphology and A-movement out of Case Domains*. Doctoral dissertation, University of Maryland at College Park.
- Uriagereka, Juan (1998). *Rhyme and Reason: An Introduction to Minimalist Syntax*. Cambridge, MA: MIT Press.

Ximenes, Cristina (2002). *Contração de Preposição em Estruturas Coordenadas*.

Master's thesis, Universidade Estadual de Campinas.

Ximenes, Cristina (2004). 'Preposition Contraction in Coordinated Structures in

Brazilian Portuguese', *MIT Working Papers in Linguistics* 47: 179–194.

Ximenes, Cristina and Jairo Nunes (2004). 'Contraction and Duplication of Prepositions

in Coordinated Structures in Brazilian Portuguese', in V. Chand, A. Kelleher, A.

J. Rodriguez, and B. Schmeiser (eds.), *WCFFL 23: Proceedings of the 23rd West*

Coast Conference on Formal Linguistics. Somerville: Cascadilla Press, 815-828.

* For valuable comments and suggestions, I would like to thank the audience of *Ways of Structure Building* (University of the Basque Country, 13-14/11/2008), Myriam Uribe-Echevarria, Vidal Valmala, Anna Snarska, and two anonymous reviewers. I would also like to thank FAPESP (2006/00965-2), CNPq (302262/2008-3), and Ikerbasque (Research Group GIC07/144-IT-210-07, Research Nets in Humanities HM-2009-1-1) for their support.

¹ Henceforth, superscripted indices will annotate copies.

² For relevant discussion, see Bobaljik 1995, Nunes 1995, 2001, 2004, Bobaljik and Brown 1997, Uriagereka 1998, and Hornstein and Nunes 2002, 2008.

³ Given that in (9b) V is a copy of the head/label of VP, merger between them would require that selfchecking be allowed, yielding massive overgeneration.

⁴ Throughout the paper irrelevant details will be omitted for the sake of presentation.

⁵ This indicates that θ -relations can trigger both Merge and Copy, whereas modification relations can only trigger Merge. That we have different conditions applying to different operations is not in itself surprising. However, it remains to be explained why modification cannot license copying.

⁶ Technically, the directionality from more to less embedded dominated can be enforced if we assume with Chomsky (2000) that a numeration is actually composed of subarrays, each of which containing one instance of a (strong) phase head, and that the computational system activates one subarray at a time. Relevant to our current discussion is the following consequence of this phase-by-phase approach: If the maximal projection determined by a subarray must be a phase, then prepositions that select clausal complements must belong to the “subordinating” array in convergent derivations, and not to the array associated with the complement clause – otherwise the CP phase will not be the maximal projection determined by the active subarray, for it will be embedded under a PP (see Nunes and Uriagereka 2000, Nunes 2001, 2004). With this in mind, consider the potential underlying subarrays for the sentence in (i) given in (ii), for instance.

(i) John called Mary after visiting Sue.

(ii) a. $N_1 = \{ \{A C_1, -ed_1\}, \{B John_1, v_1, call_1, Mary_1\}, \{C after_1, C_1, -ing_1\}, \{D v_1, visit_1, Sue_1\} \} \rightarrow \textcircled{\otimes}$

b. $N_2 = \{ \{E C_1, -ed_1\}, \{F John_1, v_1, call_1, Mary_1, after_1\}, \{G C_1, -ing_1\}, \{H v_1, visit_1, Sue_1\} \} \rightarrow \textcircled{\otimes}$

c. $N_3 = \{ \{I C_1, -ed_1\}, \{J v_1, call_1, Mary_1, after_1\}, \{K C_1, -ing_1\}, \{L John_1, v_1, visit_1, Sue_1\} \} \rightarrow \textcircled{\otimes}$

Regardless of which sequence of subarrays is activated, derivations starting with the numeration N_1 in (iia) are all illegitimate, for the maximal projection determined by subarray C is not a phase. As for N_2 in (iib), if the computation starts with subarray E, the derivation is doomed because although the complementizer can merge with T, the Extension Condition prevents T from acquiring a complement later on in the derivation. The same considerations apply to subarrays G and F. In the case of F, the problematic element is the preposition *after*. If the computational system first accesses the subarray F and builds [[John [_v [call Mary]]] *after*], the Extension Condition will later block noncyclic merger between *after* and the clausal complement. If the computation starts with subarray H in (iib), building [_{vP} _v [visit Mary]], no continuation leads to convergence either. If subarray G is activated after subarray H, *-ing* will successfully merge with the already assembled vP, but the external argument will be prevented from being inserted noncyclically later on. If subarray F is activated after subarray H, the external argument can be merged to [_{vP} _v [visit Mary]] in a cyclic manner, but a problem will then arise with *after*. Assuming that *after* selects for a clausal complement and not for a vP, it will not have its selectional features satisfied at the phase determined by subarray F and the Extension Condition prevents noncyclic introduction of the clausal ingredients present in subarray G. By contrast, N_3 in (iic) can lead to a convergent derivation if the computational system activates the subarrays L, K, J, and I in this order, as sketched in (iii). As the reader can see, if derivations work in a phase-by-phase fashion and if Extension holds, (sideward) movement is bound to proceed from more to less embedded domains.

(iii) a. *Computations based on subarray L:*

[_{vP} John [_v [visit Sue]]]

b. *Subsequent computations based on subarray K:*

[_{CP} C [_{TP} Johnⁱ -ing [_{vP} Johnⁱ [_v [visit Sue]]]]]]

c. *Subsequent computations based on subarray J:*

c'. *Sideward movement:*

[_{CP} C [_{TP} Johnⁱ -ing [_{vP} Johnⁱ [_v [visit Sue]]]]]]

[_{vP} Johnⁱ [_v [call Mary]]]

c''. *Merger:*

[_{vP} [_{vP} Johnⁱ [_v [call Mary]]] [_{PP} after [_{CP} C [_{TP} Johnⁱ -ing [_{vP} Johnⁱ [_v [visit Sue]]]]]]]]]

d. *Subsequent computations based on subarray I:*

[_{CP} C [_{TP} Johnⁱ -ed [_{vP} [_{vP} Johnⁱ [_v [call Mary]]] [_{PP} after [_{CP} C [_{TP} Johnⁱ -ing [_{vP} Johnⁱ [_v [visit Sue]]]]]]]]]]]

⁷ See e.g. Nunes 1999, 2004, 2011, Bošković and Nunes 2007, the collection of papers in Corver and Nunes 2007, and references cited there.

⁸ For different technical implementations, see Nunes 1995, 1999, 2004 for an approach in terms of feature elimination in the phonological component and Nunes 2011 for an approach in terms of chain internal probing.

⁹ In consonance with the overall bottom-up nature of the computation (see section 3.2), I assume that Chain Formation and Chain Reduction also procede from more to less embedded domains.

¹⁰ For further discussion on how lower copy pronunciation works in these cases, see Niinuma 2010.

¹¹ For a discussion of other instances of adjunct control with more than one copy phonetically realized, see Boeckx, Hornstein, and Nunes 2008.

¹² Under Nunes's (1995, 2001, 2004) system, a structure such as (51c) is independently excluded because it cannot be linearized, as the two copies do not form a chain and, accordingly, are not subject to deletion under Chain Reduction (cf. (30)). For the purposes of presentation, I will however put this possibility aside and frame the following discussion in terms of Hornstein's (2001) original Merge-over-Move approach. The proposal to be suggested below is compatible with either analysis.

¹³ The matrix clause of (52) involves a plural subject and a plural object, whereas the infinitival clause is uninflected; hence, we are dealing here with regular adjunct control rather than a *pro* licensed by an inflected infinitive. For original discussion of the finite counterparts of (52) in Brazilian Portuguese, see Modesto 2000 and Rodrigues 2004.

¹⁴ In fact, Bošković (2007) proposes that *wh*-elements in English are lexically specified as optionally having a strong feature *uF* and that its interrogative complementizer can only be checked by a *wh*-element marked with *uF*; hence, an interrogative complementizer cannot be checked by a *wh*-phrase *in situ*. However, adjunct control in English differs from what we find in Portuguese in that it always involves subject control, regardless of whether or not the object undergoes *wh*-movement, as illustrated in (i) below. English also differs from Portuguese in that the latter allows *wh-in situ* (in embedded clauses) even when there is a single *wh*-element, as shown in (ii).

- (i) a. Who_i greeted who_k after PRO_{i/*k} entering the room?
 b. Who_k did John_i greet t_k after PRO_{i/*k} entering the room?

- (ii) a. *John said that Mary is going to travel when?

b. *Portuguese:*

O João disse que a Maria viaja quando?

the João said that the Maria travels when

‘When did João say that Maria is going to travel?’

In Nunes (2010) I took the general availability of optional *wh*-movement in Portuguese to indicate that its optional specification for a strong feature is truly lexical. As for English, my conjecture is that *uF* is optionally assigned not in the lexicon, but *in the course of the derivation*, when phases are completed (as in Chomsky 2001). Assignment of such feature is however subject to Last Resort: only when the *wh*-element is not accessible to the computational system (typically, when it is not in a phase edge) can it be assigned a strong feature to undergo successive cyclic movement. Thus, a *wh*-phrase in the subject position of an adjunct clause will not be assigned *uF* and Merge-over-Move will be enforced, always yielding subject control (cf. (ib)). Notice that this suggestion is still compatible with Bošković’s (2007) main proposal that edge features are borne by the relevant moving elements and not by the heads of phases. I leave further development of this suggestion for another occasion.

¹⁵ For the sake of completeness, it remains to show how the subject control reading of both (52b) and (52c) can be obtained. In both cases, the subject control reading results from a derivation in which *os alunos* ‘the students’ is generated in the adjunct clause and undergoes sideward movement after the *wh*-phrase is merged in the matrix object position (in compliance with Merge-over-Move). As respectively shown in (i) and (ii) below, the difference between the two derivations involves the lexical specification of the *wh*-element. In (52b), it is not associated with an *uF* feature; hence we have *wh-in*

situ (cf. (i)). By contrast, in (52c) the *wh*-element has an *uF* feature and therefore must undergo *wh*-movement to check it (cf. (ii)).

(i) [Os alunos]_i [t_i [entrevistaram [**que** professores]_k] [antes de t_i sair de férias]]
the students interviewed which professors before of leave of vacation
 ‘Which professors did the students interview before leaving on vacation?’

(ii) [**Que**_{uF} professores]_k é que [os alunos]_i [t_i [entrevistaram t_k] [antes de t_i sair de férias]]
which professors is that the students interviewed before of leave of vacation
 ‘Which professors did the students interview before leaving on vacation?’

¹⁶ Such copying can be seen as a subtype of the standard operation involved in morphological reduplication.

¹⁷ However, as a reviewer observes, sideward movement in the morphological component is different from sideward movement in the syntactic component as the former does not involve two independent objects and does not seem to display directionality effects. This distinct behavior may perhaps follow if syntactic merger is subject to the Extension Condition, but morphological merger isn't. This difference in turn may follow if syntactic merger is a more complex operation involving Concatenate (which is akin to morphological merger) and Label, as proposed by Hornstein (2009). If so, Label is the operation that must apply in consonance with the Extension Condition and the difference between sideward movement in the syntactic and morphological

components would result from whether merger is associated with labeling or not. I will leave the discussion of the consequences of this speculation for another occasion.