Linearizing Chains at LF

Jairo Nunes (Universidade de São Paulo)

1. Introduction[1]

Reconstruction effects clearly show that PF and LF may differ regarding the chain links they select for interpretation. The truism of this observation has led researchers to (tacitly) assume that the mechanisms that convert chains into their PF and LF outputs are themselves different. In this paper I explore a different possibility. I propose that although one may have different PF and LF outputs for a given chain, both are ultimately derived by a single linearization procedure that triggers applications of deletion within chains. I will show that differences between chain outputs at PF and LF are determined by distinct convergence requirements imposed by the interfaces, whereas their similarities (e.g. their general format) follow from economy considerations.

Before we proceed to the discussion proper, some remarks are in order. The standard assumption in the field is that linear order is relevant for PF, but not for LF. However, this seems to me to be an unwarranted extrapolation of the assumption that the A-P system is sensitive to precedence relations, but C-I is not. Precedence is just one of the infinite relations that share linear order properties (transitivity, asymmetry, irreflexivity, and totality). Thus, it is perfectly conceivable that C-I does not operate with precedence, but operates with some other type of linear order. In Kayne's (1994) influential work, for example, the LCA applies to all syntactic representations, including LF (see his section 5.2). Again, this does not entail that the LF interfaces are sensitive to precedence. All it means is that LF is structurally organized in such a way that it can be mapped into a linear order. Thus, the question is whether C-I requires a linearly ordered object and if the answer is affirmative, the next question is what

[1] This paper has received support from CNPq (227730/2015-8). Earlier versions have been presented at the workshop *Chains in Minimalism* (Yokohama National University) and the following universities: Maryland, Campinas, Alagoas, Basque Country, Connecticut, and São Paulo. I am thankful to these audiences for comments and suggestions. Special thanks to Marcelo Ferreira, Ángel Gallego, Norbert Hornstein, Roger Martin, Ellen Thompson, Gary Thoms, Juan Uriagereka, and an anonymous reviewer for comments on previous drafts. I would like to take this opportunity to also thank Juan Uriagereka for years of guidance and friendship. Um grande abraço, Juan!

kind of linear order relation C-I operates with (keeping the standard assumption that it is not precedence).

In this paper I argue that the answer to the first question is affirmative, while remaining uncommitted as to which specific type of linear order relation C-I is sensitive to.[2] The argument is essentially based on Nunes's (1995, 1999, 2004, 2011) proposal regarding the linearization of chains at PF and Chomsky's (1993) analysis of reconstruction at LF. For Nunes, nondistinct copies induce violations of the asymmetry and irreflexivity conditions on linear order, preventing structures containing nontrivial chains from being linearized at PF; these problems can however be remedied through deletion within chains. In Chomsky's proposal, in turn, reconstruction effects or lack thereof arise as a byproduct of applications of deletion within chains at LF. From an abstract point of view, the PF and LF outputs resulting from deleting material within chains turn out to be too similar for it to be an accident. Based on this similarity, I explore the hypothesis that deletion within chains at LF is also triggered by linearization in the sense that it obliterates the violations of the asymmetry and irreflexivity conditions on linear order induced by the "repeated" material within different chain links.

The paper is organized as follows. In sections 2 and 3, I review Nunes's (2011) linearization approach to deletion within chains on the PF side and Chomsky's (1993) approach on the LF side. In section 4, I then argue that Nunes's approach can be extended to LF by showing that reconstruction effects (or lack thereof) may follow from optimal applications of deletion within chains, triggered by linearization.

2. Deletion within chains at PF and linearization

Nunes's (1995, 1999, 2004) linearization approach to phonetic realization of chains combines convergence requirements and economy considerations. In his Agree-based implementation (Nunes 2011), the system works in the following manner. Take the abstract structure K in (1), where the syntactic object α with an unvalued feature *F:u* has been copied and merged in a higher position, valuing F and forming the chain CH = (α, α).

(1) K = [[**α β γ**]$_{\mathbf{F:\surd}}$ … [X … [**α β γ**]$_{\mathbf{F:u}}$ …]]

[2] See Hornstein (2001:85) for a suggestion that the relevant notion of linear order for LF is scope (hence his Scope Correspondence Axiom).

The idea is that when K is transferred to the phonological component, it cannot be linearized in accordance with Kayne's (1994) LCA as is, because the two copies of α induce problems of symmetry and reflexivity. If copies count as nondistinct in virtue of being associated with the same material of the numeration (see Chomsky 1995:227)[3], X in (1) is required to precede and be preceded by α, for it asymmetrically c-commands and is asymmetrically c-commanded by (a copy of) α. Similarly, α is required to precede itself, for its higher copy asymmetrically c-commands the lower one. In order to circumvent these contradictory requirements, the phonological component resorts to Chain Reduction, as described in (2).

(2) *Chain Reduction* (Nunes 1995:279)
Delete the minimal number of constituents of a nontrivial chain CH that suffices for CH to be mapped into a linear order in accordance with the LCA.

(2) does not instruct deletion to target particular links. The choice of the terms to be deleted is determined by optimal interactions of convergence requirements and economy considerations. For instance, if the two copies of α in (1) are deleted, the linearization problems disappear. However, this is not an optimal solution: it arguably violates recoverability and employs two applications of deletion (each targeting a different copy), whereas a single application targeting one of the copies suffices. Similarly, simple deletion of the upper link circumvents the linearization problems, but does not lead to a convergent result at PF, for the surviving lower link does not have its F-feature valued. All of these convergence and economy computations conspire to yield the result that all things being equal, the optimal chain output at PF is the one in which the chain head is pronounced.

Although this is the most common situation, it is not the only one (see e.g. the collection of papers in Corver and Nunes 2007 and references therein). One may also find cases where a lower copy is pronounced (cf. (3)), cases where different parts of different links are pronounced (cf. (4)), and even cases where more than one link is fully pronounced (cf. (5)):

(3) a. *Romanian* (Bošković 2002):
Ce precede **ce**?
what precedes what

[3] See Martin and Uriagereka 2014 for relevant discussion.

'What precedes what?'

b. *Chain output at PF* (Bošković 2002):

[ce$_{SU}$ ~~ceов~~ precede **ceов**]

(4) a. *Bulgarian* (Rudin, Kramer, Billings, and Baerman 1999):

Dal li **si** **mu** **(gi)** parite?

given Q are him-DAT them the-money

'Have you given him the money?'

b. *Chain output at PF* (Bošković 2001):

[[~~si-mu-gi~~**-dal]-**li ... **[si-mu-gi-**~~dal~~**]** ...]

(5) a. *German* (McDaniel 1986):

Wen glaubt Hans ***wen*** Jakob gesehen hat?

whom thinks Hans *whom* Jakob seen has

'Who does Hans think Jakob saw?'

b. *Chain output at PF* (Nunes 2004):

[$_{CP}$ **wen** glaubt Hans [$_{CP}$ #[$_{C0}$ **wen**-C^0]# Jakob ~~wen~~ gesehen hat]]

The noncanonical patterns illustrated in (3)-(5) precisely involve cases where things are not equal. Take (3a), for example, which appears to involve *wh-in situ* in a multiple *wh*-fronting language such as Romanian. Bošković (2002) shows that this exceptional pattern arises when both subject and object *wh*-elements are identical and argues that pronunciation of the lower link of the object *wh*-chain circumvents this identity avoidance restriction (cf. (3b)). As for the Bulgarian pattern in (4a), Bošković (2001) argues that in the syntactic component, the complex head involving the verb and the clitics (*[si-mu-gi-dal]*) moves and left-adjoins to the interrogative particle *li*, leaving a copy behind. Deletion of the lower copy yields a grammatical output in Macedonian but not in Bulgarian, because in the former, the clitics are proclitic and *li* is enclitic but in the latter, these elements are all enclitic. A convergent reduction of the complex head chain in Bulgarian must therefore delete the copies of the clitics in the higher link and the copy of the verb in the lower link (cf. (4b)). Finally, Nunes (2004) argues that in languages that allow *wh*-copying, an intermediate *wh*-copy can be fused with a declarative C^0 in the morphological component (cf. (5b)). The fused copy then becomes invisible to (syntactic) linearization and ends up occupying a slot in the PF output in virtue of being an integral part of the linearized C^0 head.

Independent convergence requirements of the phonological component can therefore alter the set of optimal chain outputs at PF and trigger more (cf. (4b)) or less (cf. (5b)) applications of deletion than what one usually finds. There is still a question pending, though. Recall that pronunciation of the head of the chain in (1) instead of the lower copy was taken to follow from convergence requirements, as the latter does not have its F-feature valued. What then happens in cases such as (3b) and (4b), where a lower link or its parts are phonetically realized? Nunes (2011) argues that again, economy is at play. He proposes that after the higher link of (1) has its F-feature valued, it could probe the structure and value the corresponding F-feature of its next lower copy in a cascade fashion and at the end of the day, all chain links could have all of its features appropriately valued. From the point of view of PF, this convergent result is not optimal, though. (Why should the system bother to fix objects that will not appear at PF?) Nunes's proposal is that such costly additional probing and valuation are licensed only when forced by convergence requirements at PF. That is, if pronunciation of the chain head violates a convergence requirement at PF, probing and valuation of the next copy is sanctioned; otherwise, no extra probing or valuation take place.[4]

To sum up, phonetic realization of anything other than the chain head is in principle not an economical solution, but it may be licensed in case independent convergence requirements of the phonological component so demand.

---

[4] A reviewer points out that this proposal seems to resort to look-ahead computations. The point is well taken and seems related to a more general point concerning the spell-out of material containing chains. For A-movement and head movement, the problem appears to be more tractable as the relevant copies are distributed within a single phase span. For A'-movement, things get more complicated thanks to the possibility of successive cyclic movement across phases. Take the (simplified) derivation of the sentence in (ia) sketched in (ib), for instance.

(i) a. Which book did John say that Mary bought?

b. [$_{CP}$ [which book]$_{\mathbf{F:\surd}}$ did-Q John [$_{vP}$ [which book]$_{\mathbf{F:u}}$ say-v [$_{CP}$ [which book]$_{\mathbf{F:u}}$ that Mary [$_{vP}$ [which book]$_{\mathbf{F:u}}$ bought-v [which book]$_{\mathbf{F:u}}$]]]]

Arguably, the *wh*-phrase is merged in (ib) with an unvalued feature, which gets valued later on, after the *wh*-phrase reaches the matrix [Spec,CP]. Let us consider the lowest phase. If Transfer applies to the complement of the phase head, the derivation should in principle crash, due to the unvalued feature of the *wh*-phrase. Similar considerations apply to the next two higher phases. This unwanted result seems to suggest that when chains are involved, either Transfer is delayed or it is employed, but lower copies are held in a buffer until the head of the chain is computed. It is likely that whatever is the solution for the problem posed by (ii) can also be extended to account for the look-ahead problem noted by the reviewer.

3. LF deletion within chains in Chomsky 1993

Chomsky (1993) resorts to constrained applications of deletion within chains at LF to account for contrasts such as the one between (6a), where reconstruction is obligatory, and (6b), where reconstruction is optional.

(6) a. Which picture of John$_i$ did he$_{k/*i}$ see?
b. John$_i$ wondered which picture of himself$_{i/k}$ Bill$_k$ saw

Chomsky (p. 35-36) proposes that a phrase like *[$_{wh-}$ which picture of John]* should be converted into *[[which picture of John] [$_{wh-}$ t]]* or *[[which] [$_{wh-}$ t picture of John]]* "by an operation akin to QR', coupled with the requirement that 'in the operator position [Spec, CP], everything but the operator phrase must delete' and '[i]n the trace position, the copy of what remains in the operator position deletes, leaving just the phrase *wh-*.' Applied to (6a), this algorithm yields one of the outputs in (7).

(7) a. **[[which picture of John]** ~~[~~*~~wh-~~* *~~t~~*~~]~~**]** did he see **[**~~[which picture of John]~~ **[**~~*wh-*~~ ***t***]]
b. **[[which]** ~~[~~*~~wh-~~* *~~t~~* ~~picture of John]~~**]** did he see **[**~~[which]~~ **[**~~*wh-*~~ ***t*** **picture of John]]**

If both outputs were to feed C-I, (7a) should incorrectly allow coreference between *he* and *John*. In order to prevent this unwanted result, Chomsky (p. 41) adds a preference principle that chooses the option illustrated in (7b) over the one in (7a) whenever possible. Thus, this preference principle enforces the option in (7b) for (6a), yielding a Principle C effect.

As for the different behavior of anaphors illustrated in (6b), Chomsky (p. 40) assumes that at LF 'the anaphor or part of it raises by an operation similar to cliticization.' That being so, the two readings in (6b) result from cliticization applying to the lower or the upper copy of the anaphor, as respectively shown in (8a) and (9a), which are then converted into (8b) and (9b).

(8) a. John wondered [which picture of himself] Bill **self**-saw [which picture of him**self**]
b. John wondered **[[which]** ~~[~~*~~wh-~~* *~~t~~* ~~picture of himself]~~**]** Bill self-saw **[**~~[which]~~ **[**~~*wh-*~~ ***t*** **picture of himself]]**

(9) a. John **self**-wondered [which picture of him**self**] Bill saw [which picture of himself]

b. John self-wondered **[[which picture of himself]** ~~[~~*wh*- *t*~~]~~**]** Bill saw **[**~~[which picture of himself]~~ **[***wh*- ***t***]**]**

Crucially, the preference principle is taken to be inapplicable in (9a), for deletion of *picture of himself* in the upper *wh*-copy 'would break the chain (*self*, t_{self}), leaving the reflexive element without a θ-role at LF' (p. 41).

Although Chomsky's system has some appeal and covers a broad range of data, there are some conceptual and empirical problems that call for a reanalysis of his selective chain deletion approach. The QR-like operation that creates the structures to which deletion applies, for instance, appears to rely on self-adjunction, which is arguably at odds with Last Resort, and employs both copies and traces. Crucially, this use of traces is not just a shorthand for copies, for they are to be interpreted as variables in the relevant logical forms. However, the resort to both copies and traces ends up undermining part of the rationale for assuming copies instead of traces within minimalism.

The preference principle also raises some questions of its own. It is formulated as an economy principle, choosing among convergent derivations. The problem is that there is no clear reason why the option illustrated in (7b) should be more economical than the one in (7a). At face value, the two options seem to have the same derivational cost, as they both employ two applications of deletion. Further questions arise in cases like (10), where the *wh*-chain has more than two copies (see Thoms 2010 for relevant discussion).

(10) a. Which picture of $John_i$ did Mary say $he_{k/i}$ saw?

b. [**[which picture of John]** did Mary say [**[which picture of John]** he saw **[which picture of John**]]]

If the *wh*-copies (10b) undergo raising and deletion along the lines of (7b) in consonance with the preference principle, we incorrectly predict a Principle C effect for (10a).

As Chomsky restricts the QR-like operation and the applications of deletion illustrated in (7) to A'-relations, there remains the issue of how copies in A-chains are to be interpreted.

The fact that *pictures* in (11a) below, for instance, can be interpreted as being part of the idiom *take pictures* indicates that the lower copy in object position is accessed for interpretation. In turn, the lack of a Principle C effect in (11b) indicates that lower copies of *that picture of John* in the embedded clause cannot be interpreted. Discussing these sorts of facts, Chomsky suggests (p. 42) that in the case of A-chains, the (lowest) trace of an A-chain is computed for purposes of θ-marking and idiom interpretation, whereas the head of the chain is interpreted for scope 'and other matters.'

(11) a. Several pictures were taken.
a'. [several pictures] were taken [several pictures]
b. That picture of John$_i$ seems to him$_{k/i}$ to be embarrassing
b'. [that picture of John] seems to him [that picture of John] to be embarrassing

This suggestion does describe the facts, but if we focus just on the lower copies, the picture gets a bit awkward. In order to account for the lack of Principle C in (11b), the tacit assumption is that the lexical content of the lower copy in (11b') is ignored/deleted. If this is to hold generally, the idiom interpretation in (11a) should then not be based on the content of the lower link of the A-chain in (11a'), which should be deleted as in (11b'), but rather on its position as sister of *taken*. However, to assume that a syntactic position may be interpreted if its lexical content is ignored/deleted does not accord well with bare phrase structure.

In addition, A-movement may sometimes allow reconstruction. Two well-known cases involve indefinites and bound pronouns, as illustrated in (12) below. In the most natural reading of (12a), the indefinite is interpreted under the scope of *likely* and in (12b) the matrix subject must be interpreted in an intermediate position lower than the quantifier (allowing the pronoun to be bound), but not lower than *her* (bleeding Principle C).

(12) a. Someone is likely to win the lottery.
b. [His$_i$ mother]$_k$'s bread seems to [every man]$_i$ to be known by her$_k$ to be the best there is. (Lebeaux 1991)

In the next section, I show how the relevant applications of deletion within A- and A'-chains of the type discussed by Chomsky can be better understood if they are analyzed as being

triggered by linearization, some plausible convergence requirements of C-I, and economy considerations.

4. Deletion within chains at LF as linearization

4.1. General proposal

In sections 2 and 3 we saw several similarities in the way PF and LF assign interpretation to nontrivial chains. In particular, the most common situation is the one in which 'repeated' material within a chain does not receive an interpretation, which suggests that such material is deleted before it reaches the interfaces. Extending Nunes's (1995, 1999, 2004, 2011) linearization approach regarding chain outputs at PF, I would like to propose that the type of deletion within chain links at LF proposed by Chomsky (1993) is also motivated by linearization considerations. More specifically, I propose that the C-I interface can only operate with syntactic objects that can be mapped into a linear order (not necessarily precedence; see section 1). In particular, a nontrivial chain induces linearization problems analogous to the ones we witnessed on the PF side and must undergo Chain Reduction in order for the relevant syntactic structure to be readable by C-I. I argue below that the different types of chain outputs at LF also follow from optimal applications of Chain Reduction, taking into consideration convergence at C-I and economy considerations.

4.2. Analysis

4.2.1. Interpreting the head of the chain or a lower copy

Let us start with the (simplified) structures in (13) (cf. (11b') and (6a)), which have been formed in the overt component after the boldfaced constituents moved to have an uninterpretable feature valued (and deleted for LF purposes).[5]

(13) a. [**[that picture of John]**$_{F:\surd}$ seems to him [**[that picture of John]**$_{F:u}$ to be embarrassing]]

b. [**[which picture of John]**$_{F:\surd}$ did-Q he see **[which picture of John]**]$_{F:u}$

As proposed above, the two copies in both (13a) and (13b) prevent the structure from being linearized and their chains must undergo Chain Reduction. In each case, deletion of either

[5] For expository purposes, I ignore copies at the edge of vP. For relevant discussion that is compatible with the approach to defended here, see especially Thoms 2010.

of the copies allows the structure to be linearized. However, there is a crucial difference between the two copies in each structure: the higher copy has its F-feature valued, but not the lower one. In more general terms: copies rather than occurrences constitute the relevant notion for movement and once a copy is created, it has a derivational life of its own. Hence, a derivation involving deletion of the upper link can only converge if prior to deletion, the upper copy probes and values the F-feature of the lower copy (see section 2). In turn, deletion of the lower link does not require this additional operation of probing and valuation. Thus, like what we saw on the PF side, deletion of the lower link is more economical than deletion of the higher link if everything else remains constant. Hence, the lack of a Principle C effect in (11b) results from an optimal application of Chain Reduction to (13a), as sketched in (14).

(14) [**[that picture of John]**$_{F:\surd}$ seems to him [~~[that picture of John]~~$_{F:u}$ to be embarrassing]]

That being so, one wonders what prevents this reasoning from applying to (13b), thereby incorrectly bleeding Principle C. It is worth observing in this regard that Chomsky (1995:291) adds unselective binding to the two interpretive mechanisms exemplified in (7) in order to account for *wh-in situ*. Suppose we then simplify the system by dropping the mechanisms illustrated in (7) and adopt unselective binding as the only convergent option, as stated in (15).

(15) *Generalized Q-Binding Condition* (GQBC)
If the lexical item LI has a *wh*-feature, LI must be within the scope of an interrogative complementizer Q at LF.

(15) basically says that Q is the element that marks scope and *wh*-elements are to be unselectively bound by Q. Applying to (13b), the GQBC prevents Chain Reduction from deleting the lower copy of the *wh*-phrase, for the surviving copy would not be in the scope of Q. In a convergent derivation, the higher copy must therefore probe and value the F-feature of the lower copy in the covert component, as represented in (16a) below, before the *wh*-chain undergoes Chain Reduction. Applying to (16a), Chain Reduction then yields the (simplified) structure in (16b), which is consonance with the GQBC, correctly yielding a Principle C effect.

(16) a. [**[which picture of John]**$_{F:\surd}$ did-Q he see **[which picture of John]**$_{F:\surd}$]
b. [~~[which picture of John]~~$_{F:\surd}$ did-Q he see **[which picture of John]**$_{F:\surd}$]

Let us now consider the lack of Principle C in cases with more than one copy like (10), whose updated (simplified) representation is given in (17).

(17) [**[which picture of John]**$_{F:\surd}$ did-Q Mary say [**[which picture of John]**$_{F:u}$ he saw **[which picture of John]**$_{F:u}$]]

In order for (17) to be properly linearized to be shipped to C-I, the *wh*-chain must undergo Chain Reduction. Deletion of any two copies suffices for linearization purposes. However, the derivation can only converge if the surviving copy has all of its features valued and is bound by Q, in accordance with the GQBC. Applying to (17) as is, no output of Chain Reduction can meet these two requirements. In this scenario, the higher copy is then allowed to probe and value the F-feature of the next lower copy, as illustrated in (18a) below. Applying to (18a), Chain Reduction can then delete the highest and the lowest copy, yielding the convergent output in (18b), which does not induce a Principle C effect, as desired.[6]

(18) a. [**[which picture of John]**$_{F:\surd}$ did-Q Mary say [**[which picture of John]**$_{F:\surd}$ he saw **[which picture of John]**$_{F:u}$]]

b. [~~[which picture of John]~~$_{F:\surd}$ did-Q Mary say [**[which picture of John]**$_{F:\surd}$ he saw ~~[which picture of John]~~$_{F:u}$]]

Notice that after the intermediate copy in (18a) had its F-feature valued, it could probe the lowest copy and value the F-feature of the latter, in which case Chain Reduction could also yield a convergent result deleting the two higher copies (and enforcing a Principle C effect). This additional probing and valuation is blocked by economy considerations, though. If the derivation can converge with a single additional probing by a copy with valued features, further valuations in a cascade fashion are preempted.

Like what finds on the PF side, the optimal output of an application of Chain Reduction at LF is, in principle, the one in which the head of the chain is interpreted, for its features have all been valued. However, if accessing this link does not meet independent convergence requirements of C-I, the computational system is forced to value the unvalued features of next

[6] In section 4.2.3 below I address the issue of how C-I interprets a representation such as (18b), for instance, as involving a variable in the most embedded object position.

lower link (in a cascade fashion) so that it can satisfy the relevant convergence requirements if it survives Chain Reduction.

Note that this approach only tangentially relies on A/A' properties, thus being also compatible with cases of reconstruction in A-chains. Take the sentence in (12a), for example, whose updated (simplified) structure is provided in (19).

(19) [**someone**$_{\mathbf{F:\surd}}$ is likely [**someone**$_{\mathbf{F:u}}$ to win the lottery]]

The pragmatically odd reading where *someone* scopes over *likely* can be obtained via deletion of the lower copy of *someone*. More relevant to our concerns is the derivation that conveys the pragmatically more natural interpretation of (19), with wide scope for *likely,* which could be generated via deletion of the upper copy of *someone*. The question is why the two readings are available, if the former looks more economical than the latter in that it does not require additional probing to value the F-feature of the lower copy of *someone*.

The fact that reconstruction is optional with the A-chain of (19), but blocked with the A-chain of (13a) indicates that what matters for reconstruction is the content of the chain links rather than the type of chain involved. In the case at hand, the difference seems to revolve around (in)definiteness. Following Heim (1982), among others, let us assume that indefinites do not have quantificational force on their own and that their existential interpretation results from the default insertion of existential quantifiers in the mapping to logical form. For concreteness, let us further assume that a default ∃ can in principle be associated with any link of an indefinite chain. That being so, there are in fact two derivations to consider in the case of (19), depending on whether the existential quantifier is associated with the upper or the lower copy of *someone*, as respectively illustrated in (20).

(20) a. [∃**someone**$_{\mathbf{F:\surd}}$ is likely [**someone**$_{\mathbf{F:u}}$ to win the lottery]]
b. [**someone**$_{\mathbf{F:\surd}}$ is likely [∃**someone**$_{\mathbf{F:u}}$ to win the lottery]]

If Chain Reduction deletes the lower copy of *someone* in (20a), the interpretation ∃ > *likely* is appropriately derived. In turn, deletion of the lower copy in (20b) yields a nonconvergent structure with an unbound variable in the matrix clause and vacuous quantification in the embedded clause. However, an alternative derivation can converge if the higher copy of *someone* values the F-feature of the lower copy, as shown in (21a) below; Chain Reduction

may then delete the upper copy, yielding the structure in (21b), which gives rise to the reading *likely* > ∃.

(21) a. [**someone**$_{\mathbf{F:\surd}}$ is likely [∃**someone**$_{\mathbf{F:\surd}}$ to win the lottery]]
b. [~~someone~~$_{\mathbf{F:\surd}}$ is likely [∃**someone**$_{\mathbf{F:\surd}}$ to win the lottery]]

I should emphasize that what is relevant for our present concerns is not the specific interpretive mechanism that licenses indefinites, such as ∃-insertion in (20). Whatever the most appropriate mechanism is, what matters is that it alters convergence computations in such a way that feature valuation *per se* cannot provide the basis for economy to determine which copies to delete. In other words, reconstruction in (12a) is not really optional; rather, the two different readings arise from two different noncomparable derivations and an apparently nonoptimal route may be taken once deletion of lower links does not result in a convergent output. [7]

4.2.2. Linearization of overlapping chains at LF

Let us now return to the ambiguity of cases such as (6b), repeated below in (22a). Given what was proposed in section 4.1, in order for the GQBC to be complied with, the upper copy of the *wh*-phrase in (22b) must probe and value the F-feature of the lower copy, as shown in (23a). An optimal application of Chain Reduction should then convert (23a) into (23b). However, (23b) incorrectly predicts that only the embedded subject reading for the anaphor should be available.

(22) a. John$_i$ wondered which picture of himself$_{i/k}$ Bill$_k$ saw.
b. [John wondered [**[which picture of himself]**$_{\mathbf{F:\surd}}$ Q [Bill saw **[which picture of himself]**$_{\mathbf{F:u}}$]]]

(23) a. [John wondered [**[which picture of himself]**$_{\mathbf{F:\surd}}$ Q [Bill saw **[which picture of himself]**$_{\mathbf{F:\surd}}$]]]

[7] I will keep the discussion to the simplest case. From the current perspective, complexities such as the blocking effect on A-reconstruction induced by negation (see e.g. Lasnik 1999 for relevant discussion) or the lack of ambiguity induced by the Parallelism Requirement (see e.g. Fox 2000), for example, should viewed as additional convergence conditions that may tip the balance toward one or another copy.

b. [John wondered [~~[which picture of himself]~~$_{F:\surd}$ Q [Bill saw **[which picture of himself]**$_{F:\surd}$]]]

Again, the logic of the system leads us to expect that a nonoptimal reduction of the *wh*-chain in (22b) to license the matrix subject reading can only be enforced if other convergence requirements prevent full deletion of the higher *wh*-copy. Suppose we follow Chomsky (1993) in assuming the anaphors involve movement, leaving aside the issue of whether this movement takes place covertly, as in Chomsky's proposal, or overtly, as in the proposals by Lidz and Idsardi (1997) and Hornstein (2001), among others. Applied to (23a), anaphor movement yields one of the structures in (24) depending on whether it targets the lower or the upper copy of the reflexive.

(24) a. [John wondered [**[which picture of himself]**$_{F:\surd}$ Q [Bill **himself-**saw **[which picture of himself]**$_{F:\surd}$]]]

b. [John **himself-**wondered [**[which picture of himself]**$_{F:\surd}$ Q [Bill saw **[which picture of himself]**$_{F:\surd}$]]]

The *wh*-chain and the reflexive chain in each of the structures in (24) overlap in the sense that they share a constituent. In (24a), for instance, the lowest copy of *himself* is at the same time the tail of the reflexive chain and part of the tail of the *wh*-chain. The question is how Chain Reduction should proceed in these circumstances. Interestingly, a similar situation has been discussed with respect to the PF side. Nunes (1995, 2001, 2004, 2012) has argued that a parasitic gap construction such as (25a) below involves sideward movement from the within the adjunct clause to the object of *file*, followed by standard movement to [Spec,CP], yielding the (simplified) representation in (25b) (with numbered copies for expository purposes). In (25b), two chains are formed: CH_1 = (copy^3, copy^1) and CH_2 = (copy^3, copy^2). These chains induce problems of linearization at PF and must undergo Chain Reduction. There is an additional complexity, though. As Nunes observes, copy^3 is the head of both chains. Thus, if Chain Reduction deletes it in its first application, the derivation is bound to crash. The two surviving copies (copy^1 and copy^2) prevent the structure from being linearized because they do not form a chain and accordingly, cannot undergo Chain Reduction. A convergent derivation must therefore delete the tail of the first chain being reduced, before handling the second chain, eventually yielding the sentence in (25a).

(25) a. Which paper did John file without reading?
b. [**[which paper]**^{3} did John [[file **[which paper]**^{2}] [without reading **[which paper]**^{1}]]]

If the interpretation of chains at LF is conditioned by linearization, as advocated here, we can now recast Chomsky's (1993) suggestion that deletion at LF cannot break a chain (see section 3) in more general linearization terms:

(26) *Generalization on the Reduction of Overlapping Chains (GROC)*:
Given a term τ such that τ is part of more than one chain, Chain Reduction must not delete τ if it is still part of another chain.

As stated in (26), the GROC is not a new principle, but just a convenient description of how the system works in an optimal way when overlapping chains are to be linearized, be it at PF or at LF. Suppose, for instance, that Chain Reduction targets the *wh*-chain in (24b) and deletes the upper copy in consonance with the GQBC, as sketched in (27) below. The two surviving copies of *himself* in (27) are now in the same situation of copy^{2} and copy^{1} in (25b) if copy^{3} is deleted: as they do not form a chain, they cannot undergo Chain Reduction and prevent the whole structure from being linearized.

(27) [John **himself-**wondered [~~[which picture of himself]~~$_{F:\sqrt{}}$ Q [Bill saw **[which picture of himself]**$_{F:\sqrt{}}$]]]

The optimal application of Chain Reduction of the *wh*-chain in (24b) must therefore be one that both satisfies the GQBC and is in consonance with the GROC. This is what we have in (28) below, where *which* is deleted in the higher copy of the *wh*-phrase (allowing the GQBC to be met) and *picture of himself* is deleted in the lower copy (in compliance with the GROC). The *himself*-chain in (28) can then undergo Chain Reduction and the whole structure can be properly linearized at LF, yielding the matrix reading interpretation for the anaphor of (22a).

(28) [John **himself-**wondered [**[**~~which~~ **picture of himself]**$_{F:\sqrt{}}$ Q [Bill saw **[which** ~~picture of himself~~**]**$_{F:\sqrt{}}$]]]

Notice that if scattered deletion of the sort seen in (28) could also be resorted to in the reduction of the *wh*-chain of (16a), repeated below in (29), the output should incorrectly bleed Principle C, as shown in (30).

(29) [**[which picture of John]**$_{F:\surd}$ did-Q he see **[which picture of John]**$_{F:\surd}$]

(30) [**[~~which~~ picture of John]**$_{F:\surd}$ did-Q he see **[which ~~picture of John~~]**$_{F:\surd}$]

Nothing special needs to be added in order to block (30), though. Like its PF counterpart (see section 2), scattered deletion is in principle not an optimal output of Chain Reduction at LF and is enforced only when deletion of the whole link does not lead to a convergent result, as was the case in (27). In particular, Chain Reduction of the *wh*-chain of (29) as in (30) leads to a convergent result (the structure can be linearized and the GQBC is satisfied), but is not economical: it involves (at least) two operations of deletion, whereas a single application targeting the upper copy should suffice for a convergent result to obtain (cf. (16b)). In sum, the proposal entertained here deduces Chomsky's (1993) preference principle from economy computations regulating applications of deletion under Chain Reduction.

This approach also provides a rationale for cases like (12b), repeated here in (31) (see Lebeaux 1991), where reconstruction in A-chains seems to be forced in order for the pronoun to be bound.

(31) [His$_i$ mother]$_k$'s bread seems to [every man]$_i$ to be known by her$_k$ to be the best there is.

If (31) simply involved A-movement of *his mother's bread*, as represented in (32), we should not expect reconstruction, for the head of the chain is the optimal link to survive Chain Reduction.

(32) [**[his mother's bread]**$_{F:\surd}$ seems to every man [**[his mother's bread]**$_{F:u}$ to be known by her [**[his mother's bread]**$_{F:u}$ to be the best there is]]]

Again, the logic of the system tells us to look for convergence reasons that would prevent the head of the chain in (32) from surviving Chain Reduction. Here is one possibility, based on Hornstein's (2001:177) proposal that bound pronouns are not part of the numeration and are inserted when movement fails. According to this proposal, the derivation of (33) below, for

instance, proceeds along the lines of (34a-b), where *everyone* moves to the external argument position in violation of the Left Branch Condition; this derivation can however be saved if a pronominalization operation targets the source of such movement and replaces it with a pronoun, as sketched in (34c).

(33) Everyone$_i$ loves his$_i$ mother.

(34) a. [loves [everyone's mother]]
b. [**everyone** loves [**everyone**'s mother]]
c. [**everyone** loves [**pronoun**'s mother]]

Leaving details of implementation aside, suppose the step in (34c) does not take place in the syntactic component, but is a matter of spell-out in the phonological component. For concreteness, suppose that an illicit instance of movement assigns * to the relevant copy, signaling to the phonological component that if pronounced, that copy should be realized as a pronoun. That being so, the structure that feeds C-I is (34b), rather than (34c). If something along these lines is on the right track, the derivation of (31) should then be as sketched below: in (35a) *[every man's mother's bread]* moves from the most embedded to the intermediate clause; in (35b) *seems* is merged and *[every man]* moves to its experiencer position in violation of the Left Branch Condition, triggering assignment of * to the lower copy; finally, after additional material is merged in (35c), *[*[every man]'s mother's bread]* moves to the matrix subject position, valuing its Case-feature.

(35) a. [$_{VP}$ seems [**[[every man]'s mother's bread]$_{F:u}$** to be known by her [**[[every man]'s mother's bread]$_{F:u}$** to be the best there is]]]
b. [$_{VP}$ **[every man]** seems [**[*[every man]'s mother's bread]$_{F:u}$** to be known by her [**[[every man]'s mother's bread]$_{F:u}$** to be the best there is]]]
c. [$_{TP}$ **[*[every man]'s mother's bread]$_{F:\surd}$** T [$_{vP}$ seems-v [$_{VP}$ **[every man]** seems [**[*[every man]'s mother's bread]$_{F:u}$** to be known by her [**[[every man]'s mother's bread]$_{F:u}$** to be the best there is]]]]]

In the phonological component, the chain involving *[[every man]'s mother's bread]* and the chain involving movement of *seem* to the light verb (see Chomsky 1995:305) are optimally reduced (i.e. the lower copies are deleted), as shown in (36a). Finally, the inherent Case

assigned by *seem* to *every man* is realized as *to* and the **[every man]* is realized as *his*, as shown in (36b). Crucially, *his* and *every man* in (36) are distinct on the PF side and do not prevent the structure from being linearized in this component.

(36) *PF*:

a. *Chain Reduction*:

[$_{TP}$ **[*[every man]'s mother's bread]$_{F:\surd}$** T [$_{vP}$ seems-v [$_{VP}$ **[every man]** ~~seems~~ [~~[~~*[every man]'s mother's bread]~~]$_{F:u}$~~ to be known by her [[~~[every man]'s mother's bread]~~]$_{F:u}$~~ to be the best there is]]]]]

b. *Phonological realization*:

[$_{TP}$ **[*his* mother's bread]$_{F:\surd}$** T [$_{vP}$ seems-v [$_{VP}$ *to*-**[every man]** ~~seems~~ [~~[~~*[every man]'s mother's bread]~~]$_{F:u}$~~ to be known by her [[~~[every man]'s mother's bread]~~]$_{F:u}$~~ to be the best there is]]]]]

Things are different on the LF side, though. By assumption, *-assignment only has effects on the PF side; hence, the structure that feeds the covert component is actually the one in (37) below (= (35c) without *s), with three chains to be reduced: the one involving *seem*, the one involving *[[every man]'s mother's bread]*, and the other involving *[every man]*.

(37) [$_{TP}$ **[[every man]'s mother's bread]$_{F:\surd}$** T [$_{vP}$ seems-v [$_{VP}$ **[every man]** seems [**[[every man]'s mother's bread]$_{F:u}$** to be known by her [**[[every man]'s mother's bread]$_{F:u}$** to be the best there is]]]]]

Since reduction of chain involving *seems* is immaterial to the present discussion, let us simply assume that it proceeds in an optimal way, deleting the lower copy. The case of interest is the reduction of the other two chains, for they overlap: the copy of *every man* within the intermediate clause is at the same time the tail of its own chain and part of the intermediate link of a bigger chain. In situations like this, the GROC states that Chain Reduction cannot delete a link that is shared by more than one chain, for otherwise the structure cannot be later linearized (surviving copies that are not in a chain relation induce linearization problems but cannot undergo Chain Reduction). Reduction of the chain involving *every man's mother's bread* must therefore preserve the intermediate link and delete the other two, which in turn requires that the upper copy first value the intermediate copy. After valuation and reduction, we obtain the

structure in (38), which can be properly linearized after the chain involving *every man* is further reduced:

(38) [TP ~~[[every man]'s mother's bread]~~$_{F:\sqrt{}}$ T [vP seems-v [VP **[every man]** ~~seems~~ [**[[every man]'s mother's bread]**$_{F:\sqrt{}}$ to be known by her [~~[[every man]'s mother's bread]~~$_{F:u}$ to be the best there is]]]]]

Similar considerations apply to the well-known cases of reconstruction blocked by a bound pronoun. (39) below, for instance, does not have the reading *every* > *some*, which should be licensed if the matrix subject could be reconstructed. If bound pronouns are disguised copies are suggested above, (39) is to be derived as sketched in (40).

(39) [Some student]$_i$ seems to his$_i$ classmates to like every professor

(40) a. K = [**[some student]** to like every professor]
L = [**[some student]**'s classmates]
b. [VP [**[some student]**'s classmates] seems [**[some student]** to like every professor]]
c. [TP **[some student]** T [vP seems-v [VP [***[some student]**'s classmates] seems [**[some student]** to like every professor]]]]

In (40a) *[some student]* undergoes sideward movement from K to L, which then merges in the specifier of *seem* (cf. (40b)). After further computations, *[some student]* moves from within the specifier of *seem* to the matrix subject position (triggering *-assignment), yielding (40c), where the highest copy forms a distinct chain with each of the lower copies. On the PF side, the starred-constituent is realized as a pronoun and the other lower copy is deleted. On the LF side, on the other hand, both chains in (40c) must be reduced. In accordance with the GROC, Chain Reduction cannot delete the highest copy (otherwise, the other copies, which do not form a chain, prevent the structure from being linearized). Chain Reduction at LF then deletes the lower link of each chain in (40c), yielding lack of reconstruction in (39).

Following Chomsky (1993), I have assumed that it is not the case that all the material in each link of a given nontrivial chain is assigned an interpretation at C-I. Based on the fact that a comparable state of affairs is also found on the PF side, I have argued that deletion within chains is triggered by the same requirement (linearization) and executed by the same mechanism (Chain Reduction). Mismatches with respect to the outputs of Chain Reduction a

PF and LF are claimed to follow from the different convergence conditions each level is subject to, in combination with economy considerations on how these demands can be satisfied in an optimal way.

4.2.3. Uriagereka's address problem

Any derivation in which Chain Reduction at LF deletes the lowest copy of a chain raises the issue of how C-I recovers the relation that was established upon merger of this copy. This recoverability problem is reminiscent of Uriagereka's (1999:267) 'address' problem. If Spell-Out can apply multiple times in the course of the derivation, as proposed by Uriagereka, there arises the question of how spelled-out structures are appropriately 'plugged in' where they belong when the whole structure is computed. Nunes and Uriagereka (2000:23) propose that labels can provide the 'address' for the appropriate plugging in. Extending Nunes and Uriagereka's solution to the case under discussion, I propose that labels are preserved when chains are linearized at LF, providing the means for the relevant information to be recovered. To be specific, when Chain Reduction targets a given syntactic object K = {$\underline{\gamma}$, {α, β}}, with label $\underline{\gamma}$ and terms α and β (see Chomsky 1995), it may delete α, β, or the set {α, β}, but not its label $\underline{\gamma}$.[8] Under this view, (14), for example, should be seen as a shorthand for (41), where the head of the chain is to be interpreted as being thematically related to the embedded predicate because its label is nondistinct from the label of the reduced lower copy.

(41) [**{<u>that</u>, {that, [picture of John]}}** seems to him [**{<u>that</u>,** ~~{that, [picture of John]}~~**}** to be embarrassing]]

This proposal captures the intuition underlying Chomsky's (1993) system with respect to the interpretation of the tail of chains, but does so in a way compatible with bare phrase structure and without assuming traces in addition to copies. It is worth pointing out that the linearization problems induced by nontrivial chains relate to the nondistinct lexical items present in each link. Labels are constructed from lexical items (see Chomsky 1995), but are not lexical items themselves. Thus, the presence of nondistinct labels after a given chain undergoes

[8] For purposes of discussion, I am assuming with Chomsky (1995) that labels are produced as part of the inner workings of the operation Merge. As far as I can see, the proposal to be presented below may also be compatible with approaches under which labels are produced by a specific operation like Hornstein and Nunes's (2008) and Hornstein's (2009) Label operation, or by a labelling algorithm, as in Chomsky 2013.

Chain Reduction, as in (41), does not prevent the structure containing them from being linearized. Labels therefore play an important role in the mapping to C-I, as they preserve information that could be lost if Chain Reduction at LF radically erased the whole chain link.

Independent evidence for this approach comes from cases where labels are arguably computed with respect to binding theory. Consider the data below, for instance.

(42) a.*Which picture of who$_i$ did he$_i$ like?
b. *Who$_i$ did he$_i$ like?

(43) a. Which picture of who$_i$ did Mary say he$_i$ liked?
b. *Who$_i$ did Mary say he$_i$ liked?

Given the GQBC (see section 4.2.1), the optimal convergent reduction of the *wh*-chain associated with each sentence in (42) is the one which deletes the higher link so that the surviving link is c-commanded by the interrogative complementizer; hence the Principle C effect in both sentences. The lack of such an effect in (43a) is accordingly tied to the presence of a copy of the *wh*-phrase in the embedded [Spec, CP]; if this copy survives Chain Reduction at LF, it complies with the GQBC and is not c-commanded by the pronoun. (43b), which also involves a copy of *who* in the embedded [Spec, CP], should then behave in the same way. However, its ungrammaticality indicates that the availability of this additional copy does not suffice to mitigate the Principle C violation that is arguably involved in (42b). Neither can the mere difference in syntactic complexity between the *wh*-phrases of (43) be responsible for their grammaticality contrast, for the syntactically complex *wh*-phrase in (44a) below patterns like *who* in (43b). And to make matters more puzzling, if the *wh*-phrase of (44a) is embedded as in (44b), we again find a case of Principle C bleeding similar to the one in (43a).

(44) a. *Whose$_i$ mother did Mary say he$_i$ called?
b. Which picture of whose$_i$ mother did Mary say he$_i$ liked?

These data can be accounted for if traditional (*wh*-)pronouns may spell out phrasal objects (see e.g. Déchaine and Wiltschko 2002). For concreteness, let us assume that *who* involves a *wh*-layer on top of a ϕ-layer and *whose* involves a more complex structure with a *wh*-layer on top of a DP-layer, on a par with Dutch *wie zijn vrouw* (lit. who his wife) 'whose

wife' (see Barbiers, Koeneman, and Lekakou 2010:9). After Chain Reduction, the sentences in (42)-(44) should then be respectively associated with the (simplified) LF structures in (45)-(47).

(45) a. [**{<u>which</u>,** ~~{which, [picture of who]}~~**}** did-Q he like **{<u>which</u>, {which, [picture of who]}}**]

b. [**{<u>who</u>,** ~~{who, ϕP}~~**}** did-Q he like **{<u>who</u>, {who, ϕP}}**]

(46) a. [**{<u>which</u>,** ~~{which, [picture of who]}~~**}** did-Q Mary say [**{<u>which</u>, {which, [picture of who]}}** he liked **{<u>which</u>,** ~~{which, [picture of who]}~~**}**]]

b. [**{<u>who</u>,** ~~{who, ϕP}~~**}** did-Q Mary say [**{<u>who</u>, {who, ϕP}}** he liked **{<u>who</u>,** ~~{who, ϕP}~~**}**]]

(47) a. [**{<u>who</u>,** ~~{who, [his mother]}~~**}** did-Q Mary say [**{<u>who</u>, {who, [his mother]}}** he liked **{<u>who</u>,** ~~{who, [his mother]}~~**}**]]

b. [**{<u>which</u>,** ~~{which, [picture of whose mother]}~~**}** did-Q Mary say [**{<u>which</u>, {which, [picture of whose mother]}}** he liked **{<u>which</u>,** ~~{which, [picture of whose mother]}~~**}**]]

Satisfaction of the GQBC necessarily places *who(se)* in the c-command domain of *he* in (45), but not in (46a) or (47b), thanks to the copy in the embedded [Spec, CP]; hence, they must be disjoint in the former, but not necessarily in the latter. The interesting cases for our discussion are the ones in (46b) and (47a), which require a disjoint reading between *he* and *who(se)* despite the fact that they are not in a c-command relation. When labels are taken into account, the puzzle disappears. In (46a) and (47b), *who(se)* is buried within the deleted material of the lowest copy and is not available for binding theory computations. By contrast, in (46b) and (47a) *who(se)* is deleted in the lowest copy but its label survives deletion and is available for binding theory purposes, which forces it to be interpreted as disjoint from the c-commanding pronoun. The correlation between syntactic complexity and availability of additional copies above the pronoun is therefore spurious. What really matters is the label of the reduced copies that remain in the c-command domain of the pronoun.

The Principle C effect in (46b) and (47a) mimics the pronunciation of more than one copy (see section 2). In the case of PF, an additional pronounced copy does not cause problems of linearization because it has been morphologically fused with some head (Nunes 2004). Hence, it is not computed for linearization purposes on its own because it is not an independent

lexical item, but an integral part of the reanalyzed lexical item. In the case of (46b) and (47a), *who(se)* is in a sense being interpreted in more than one position al LF, also without creating problems of linearization. Crucially, we do not have two nondistinct copies of the lexical item *who(se)*, but an lexical item and a label determined by this lexical item.

4.2.4. Additional similarities: bottom-up linearization

4.2.4.1. ATB extraction

Consider the contrast in (48) (based on Munn 1993), which appears to show that reconstruction under ATB only affects the first conjunct.

(48) a. Which of John$_i$'s cousins does Mary like and he$_i$ hate?
b. *Which of John$_i$'s cousins does he$_i$ like and Mary hate?

Under a sideward movement approach to ATB extraction (see Nunes 1995, 2001, 2004, 2012), the *wh*-phrase of (48b) moves from the object of *hate* to the object of *like*, before moving to [Spec, CP], yielding the (simplified) structure in (49), with the chains CH_1 = (copy^3, copy^1) and CH_2 = (copy^3, copy^2). The linearization of these chains at LF yields the output in (50) or (51), depending on whether Chain Reduction targets CH_1 or CH_2 first.

(49) [**[which of John's cousins]**^{3} does [he like **[which of John's cousins]**^{2}] and [Mary hate **[which of John's cousins]**^{1}]]

(50) [~~[which of John's cousins]~~^{3} does [he like **[which of John's cousins]**^{2}] and [Mary hate ~~[which of John's cousins]~~^{1}]]

(51) [~~[which of John's cousins]~~^{3} does [he like ~~[which of John's cousins]~~^{2}] and [Mary hate **[which of John's cousins]**^{1}]]

If the order of applications did not matter, the sentence in (48b) should incorrectly allow co-reference between *he* and *John* thanks to the output in (51). The fact that we have a Principle C effect in (48b) but not in (48a) thus shows that Chain Reduction must apply to CH_1 prior to CH_2. This seems to indicate that Chain Reduction proceeds in a bottom-up fashion, perhaps as a consequence of phase-based computations. Tentatively, I propose that linearization of chains unfolds along the lines of (52).

(52) *Bottom-up Linearization:*

Given two chains CH_1 and CH_2, Chain Reduction applies to CH_1 first if links of CH_1 are lower than links of CH_2, where α is lower than β if α is asymmetrically c-commanded by β or α is asymmetrically c-commanded by γ and γ dominates β.

In the case of (49), CH_1 = ($copy^3$, $copy^1$) must be reduced first, because $copy^1$ is lower than $copy^2$ ($copy^1$ is asymmetrically c-commanded by the first conjunct, which dominates $copy^2$).

The above discussion leads us to expect that linearization at PF should also proceed in a bottom-up fashion, along the lines of (52). In the case of (49), we cannot see the effects of (52) at PF because reduction of each chain deletes the lower copy. But if pronunciation of the chain head causes the derivation to crash, lower copy pronunciation should involve the first conjunct, in a way parallel to what we saw with reconstruction in (49). An interesting case illustrating this possibility involves the contrast in (53) in (noncolloquial) Czech, as discussed by Dotlačil (2008).

(53) *Czech* (Dotlačil 2008):

a. *Zavolal **jsem** **Petra** a představil známým.
called *aux$_{1SG}$* *Petr$_{acc}$* *and* *introduced* *friends*
'I called Petr and introduced him to friends.'

b. Zavolal **jsem** **ho** a představil známým.
called *aux$_{1SG}$* *him$_{acc}$* *and* *introduced friends*
'I called him and introduced friends.'

(53a) shows that Czech does not allow auxiliary gapping and object drop under coordination. In this regard, the grammaticality of (53b), which differs minimally from (53a) in having a pronominal clitic in place of *Petra*, is rather unexpected. Dotlačil (2008) argues that the auxiliary and the pronominal clitics in (53b) in fact undergo ATB extraction (via sideward movement), yielding the simplified structure in (54a) below (with English words), 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 (54b). Similar reduction of the chain involving the first conjunct would leave the clitic cluster in clause initial position, which is not allowed in Czech. Chain Reduction then deletes the higher

copy, as shown in (54c). As the reader can see, the order of applications of Chain Reduction in (54) is the one required by (52).

(54) 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**~~]]

We thus have another case of parallelism between linearization of chains at PF and linearization of chains at LF.

4.2.4.2. Parasitic gaps

Chomsky (1986:60), citing Kearney (1983), observes that reconstruction in the parasitic gap position of constructions like (55a) is not possible. However, Williams (1989/1990:271) shows that in constructions such as (55b), the *wh*-phrase does reconstruct into the parasitic gap, inducing a Principle C effect. In addition, Munn (1994) shows that even in the case of Principle A, we may get the opposite pattern from the one discussed by Chomsky if the parasitic gap is within a subject, as illustrated in (55c).

(55) a. Which books about himself/*herself did John file *t* before Mary read *PG*?
b. *Whose$_i$ mother did we warn *t* before he$_i$ arrested *PG*?
c. Which picture of himself/*herself did every boy who saw *PG* say Mary liked *t*?

The complex pattern in (55) finds a simple account under the present analysis. Let us start with (55b). Recall that labels may give rise to Principle C effects even in absence of true reconstruction (see section 4.2.3). This appears to be the case of (55b). Notice that once *whose mother* is embedded in a *wh*-phrase, as in (56) below, the Principle C effect disappears. (57) further shows that reconstruction into the 'real' gap is indeed full reconstruction and not just an effect of labels. This leads to the conclusion that once the role of labels is filtered out, (55b) actually patterns with (55a).

(56) Which picture of whose$_i$ mother did Mary buy *t* after he$_i$ found *PG*?

(57) *Which picture of whose$_i$ mother did he$_i$ buy *t* after Mary found *PG*?

Let us then examine the contrast between (55a) and (55c), whose (simplified) representations under a sideward movement approach are given in (58).

(58) a. [**[which books about himself]**^{3} [did-Q John [file **[which books about himself]**^{2}] [before Mary read **[which books about himself]**^{1}]]]

b. [**[which picture of himself]**^{3} did [every boy who saw **[which picture of himself]**^{2}] say Mary liked **[which picture of himself]**^{1}]

Abstracting away from the details involving reflexive movement (see section 4.2.2), the determination of the site for reconstruction follows from the bottom-up approach to linearization outlined in section 4.2.4.1. If copy^1 counts as lower than copy^2 in both (58a) and (58b), Chain Reduction must apply to CH_1 = (copy^3, copy^1) before CH_2 = (copy^3, copy^2), yielding reconstruction into the site occupied by copy^2.

The assumption that copy^1 is lower than copy^2 in (58a) is admittedly less trivial, as it relates to the independent issue of how 'right-adjoined' adjuncts are to be linearized. Putting aside eventual amendments to (52) prompted by this issue, the directionality of Chain Reduction proposed above predicts that in adjunct PG constructions where pronunciation of the head of the chain causes the derivation to crash, the trace to be pronounced should not be the one within the adjunct, for its chain counts as lower and is reduced first. Evidence for this conclusion is provided by the Romanian sentence in (59) (see Bošković 2002 and Niinuma 2010).

(59) *Romanian* (Bošković 2002):

Ce precede **ce$_i$** fără să influenţeze **PG$_i$**?

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

'What precedes what$_i$ without influencing it$_i$?'

Under a sideward movement analysis, (59) involves movement of *ce* from the object of *influenţeze* to the object of *precede* before undergoing *wh*-fronting, as represented in (60) below, with the chains CH_1 = (copy^3, copy^1) and CH_2 = (copy^3, copy^1). Applying to CH_1, Chain Reduction deletes the lower copy of *ce*, yielding (61a). By contrast, if Chain Reduction deletes the lower copy of *ce* when applying to CH_2, the derivation will not converge due to the ban on adjacent homophonous *wh*-phrases in Romanian (see section 2). In order to circumvent this problem, Chain Reduction deletes the higher copy, as shown in (61b) (see Bošković 2002), and

the structure surfaces as (59), which superficially seems to involve a *wh-in situ* licensing a parasitic gap. Like what we saw with respect to the reconstruction in (55a), the copy within the adjunct counts as lower than copy in the matrix object position. Importantly, the correct output is only obtained if chains are reduced in a bottom-up fashion, as proposed in (52).[9]

(60) [ce$_{SU}$ **ce$_{OB}$**^{3} [[precede **ce$_{OB}$**^{2}] [fără să influenţeze **ce$_{OB}$**^{1}]]]

(61) a. [ce$_{SU}$ **ce$_{OB}$**^{3} [[precede **ce$_{OB}$**^{2}] [fără să influenţeze ~~ce$_{OB}$~~^{1}]]]
b. [ce$_{SU}$ ~~ce$_{OB}$~~^{3} [[precede **ce$_{OB}$**^{2}] [fără să influenţeze ~~ce$_{OB}$~~^{1}]]]

5. Concluding remarks

It goes without saying that the proposal outlined in the preceding sections is very programmatic and does not do proper justice to the large literature on the interpretation of chains under the copy theory that has been emerged under minimalism. But even being programmatic, it at the very least points out similarities in the way PF and LF handle chains, which do not seem to be coincidental and call for a unified approach. The specific unified approach I sketched extended to chain outputs at LF my earlier proposal that the different patterns of chain outputs one finds at PF are determined by linearization, combined with additional PF convergence requirements and general economy computations. Taking Chomsky's (1993) original proposal on how to interpret copies at LF as a testing ground, I have argued that the different patterns of chain outputs at LF are also determined by optimal linearization taking into account LF convergence conditions and economy considerations.

To say that LF is sensitive to linearization definitely causes some eyebrows to raise. However, I believe this is due to the fact that we informally equate linear order with precedence

[9] This analysis predicts that an analogous case involving ATB extraction should enforce pronunciation of a lower copy in the first and not the second conjunct. However, Niinuma (2010) reports that the reverse is true, as shown in (i) below. At the moment, I do not have an account of the unexpected pattern in (i).

(i) a. *Ce a precedat ce şi a influenţat?
what has preceded what and has influenced
b. Ce a precedat şi a influenţat ce?
what has preceded and has influenced what
'What preceded and influenced what?'

in our discussions on how overt syntax is to be mapped into a PF object and it is generally assumed that precedence is relevant for PF, but not for LF. As I mentioned in the introduction, I remain uncommitted as to which specific linear order relation C-I is sensitive to. What I hope to have shown is that some linear order relation seems to be ultimately responsible for the deletion of 'repeated' material within a chain so that asymmetry and irreflexivity are observed.

References

Barbiers, Sjef, Olaf Koeneman, and Marika Lekakou 2010. 'Syntactic doubling and the structure of *wh*-chains', *Journal of Linguistics* 46: 1-46.

Bošković, Željko 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena*. Amsterdam: Elsevier.

Bošković, Željko 2002. 'On multiple *wh*-fronting', *Linguistic Inquiry* 33: 351-383.

Chomsky, Noam 1986. *Barriers*. Cambridge, Mass.: MIT Press.

Chomsky, Noam 1993. 'A minimalist program for linguistic theory', in Kenneth Hale and Samuel Keyser (eds.). *The View from building 20: essays in linguistics in honor of Sylvain Bromberger*, pp. 1-52. Cambridge, Mass.: MIT Press.

Chomsky, Noam 1995. 'Categories and transformations'. In *The Minimalist Program*, pp. 219-394. Cambridge, Mass.: MIT Press.

Chomsky, Noam 2013 'Problems of projection'. *Lingua* 130: 33-49.

Corver, Norbert and Jairo Nunes (eds.) 2007. *The copy theory of movement*. Amsterdam/Philadelphia: John Benjamins.

Déchaine, Rose-Marie and Martina Wiltschko 2002. 'Decomposing pronouns', *Linguistic Inquiry* 33: 409-442.

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*, pp. 134-150. Ann Harbor: Michigan Slavic Publications.

Fox, Danny 2000. *Economy and Semantic Interpretation*. Cambridge, Mass.: MIT Press.

Heim, Irene 1982. 'The semantics of definite and indefinite noun phrases', doctoral dissertation, University of Massachusetts, Amherst.

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 2008. 'Adjunction, Labeling, and Bare Phrase Structure'. *Biolinguistics* 2: 57-86.

Kayne, Richard 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Kearney, Kevin 1983. 'Governing categories'. Ms., University of Connecticut, Storrs.

Lasnik, Howard 1999. 'Chains of arguments', in Samuel David Epstein and Norbert Hornstein (eds.). *Working minimalism*, pp. 189-215. Cambridge, Mass.: MIT Press.

Lebeaux, David 1991. 'Relative clauses, licensing, and the nature of the derivation', in Susan Rothstein (ed.). *Syntax and semantics 25: perspective on phrase structure: heads and licensing*, pp. 209-239. San Diego, Calif.: Academic Press.

Lidz, Jeffrey and William Idsardi 1997. 'Chains and Phono-logical Form'. *UPenn Working Papers in Linguistics* 8: 109-125.

McDaniel, Dana 1986. 'Conditions on *wh*-chains', doctoral dissertation, CUNY.

Martin, Roger and Juan Uriagereka 2014. 'Chains in minimalism', in Peter Kosta, Steven L. Franks, Teodora Radeva-Bork, and Lilia Schürcks (eds.), *Minimalism and beyond: Radicalizing the interfaces*, pp. 169-194. Amsterdam/Philadelphia: John Benjamins.

Munn, Alan 1993. 'Topics in the syntax and semantics of coordinate structures', doctoral dissertation, University of Maryland, College Park.

Munn, Alan 1994. 'A minimalist account of reconstruction asymmetries', in Mercé Gonzàlez (ed.). *NELS* 24, pp. 397-410. Amherst: University of Massachusetts, GLSA.

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 Samuel David Epstein and Norbert Hornstein (eds.), *Working minimalism*, pp. 217-249. Cambridge, Mass.: MIT Press.

Nunes, Jairo 2001. 'Sideward movement', *Linguistic Inquiry* 31.2: 303-344.

Nunes, Jairo 2004. *Linearization of chains and sideward movement*. Cambridge, Mass.: MIT Press.

Nunes, Jairo 2011. 'The Copy Theory', in Cedric Boeckx (ed.). *The Oxford handbook of linguistic minimalism*, pp. 143-172. Oxford: Oxford University Press.

Nunes, Jairo 2012. 'Sideward movement: Triggers, timing, and outputs', in Myriam Uribe-Etxebarria and Vidal Valmala (eds.). *Ways of structure building*, pp. 114-142. Oxford: Oxford University Press.

Nunes, Jairo and Juan Uriagereka 2000. 'Cyclicity and extraction domains', *Syntax* 3: 20-43.

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.

Thoms, Gary 2010. 'Syntactic reconstruction and scope economy', paper delivered at GLOW 33, University of Wrocław, 14/4/2010.

Uriagereka, Juan 1999. 'Multiple Spell-Out', in Samuel David Epstein and Norbert Hornstein (eds.). *Working minimalism*, pp. 251-282. Cambridge, Mass.: MIT Press.

Williams, Edwin 1989/1990. 'The ATB theory of parasitic gaps'. *The Linguistic Review* 6: 265-279.