

Rightward Movement from A Different Perspective

Cristiano Chesi

University of Siena

Abstract

A left-to-right (Phillips 1996) top-down (Chesi 2004-07) Minimalist Grammar is a generative procedure that operates on lexical items using minimalist structure building operations in order to isolate the best possible approximation of the (infinite) set of well formed linguistic expressions. In such a framework, all long-distance dependencies are “rightward” dependencies given that the dependency trigger must be found first (e.g. an unselected argument in a pre-verbal position), then the dependent (possibly non-local) constituent must be univocally identified (e.g. the selecting verbal head). Using a memory buffer (a well-known computational device) to store and retrieve constituents in a principled order/way so as to deal with long distance dependencies, we can correctly characterize standard (successive cyclic) movement, islandhood (Chesi 2004, 2007), parasitic gap constructions (Bianchi Chesi 2005-08), quantifier raising (Bianchi Chesi 2007) and A-binding (Bianchi 2007). In this paper I will show that classic rightward movements such as Extraposition and Heavy NP-Shift can also be accommodated in such a framework thus maintaining discrimination among

crucial cross-linguistic differences (e.g. Extraposition from subject Vs. object) and among special properties (e.g. clause-boundedness) that make Extraposition and Heavy NP-Shift peculiar compared to standard leftward movement.

Keywords: PP / relative clause Extraposition, Heavy NP-Shift, QR, Top-Down Left-to-Right derivation

1. Introduction

In any (standard) theory that includes the notion of feature-driven movement, rightward movement (like PP and (restrictive) Relative Clause (rRC) Extraposition (EXT) in SVO languages) shows peculiar properties that are usually resistant to a non-stipulative unified account. These properties are “clause”-boundedness (1) ((1).a from Baltin 2006, (1).b-b' from Akmajian 1975), adjunct/argument asymmetries sensitive to displacement directionality (2) and the definiteness constraint (3) (from Fox and Nissenbaum 1999):

(1) a.*[John was believed [to be certain _] by everybody]

that the Mets would lose.

- b. *[A review of [a book _]] appeared *by three authors*
- b'. [A review [of a book] _] appeared *by three authors*
- (2) a. We saw [a painting _] yesterday *of John*
- a'. *Of whom* did you see [a painting _]?
- b. We saw [a painting _] yesterday *from the museum*
- b'. **??* *From where* did you see [a painting _]?
- (3) a. I saw the (best) picture yesterday *from the museum.*
- a'. *??* I saw the (best) picture yesterday *of the museum.*
- a". I saw a (very good) picture yesterday *of the museum.*

In this paper I will show that this cluster of properties can be handled in a natural way if we drift away from the standard conception of bottom-to-top derivation and we redefine structure-building operations to be top-down, assuming a (weaker) version of the Linear Correspondence Axiom (Kayne 1994) and a left-to-right head-driven phase-by-phase derivation.

In the first part of this paper (§2.0), I will describe rightward movement in English, Dutch and Italian, summarizing some fundamental findings on Extraposition: what moves (§2.1), from where (§2.2) it does and where the extraposed element seems to be attached (§2.3).

In the second part (§3.0), I will present data related to Heavy NP-Shift that present an interesting pattern of properties that are also consistent with Extraposition (e.g. Clause-Boundedness).

In the third part (§4.0), I will analyze some of the proposed standard solutions, highlighting the main problems in the given analyses.

The last section (§5.0) is dedicated to the proposed solution: the memory-buffered approach to long-distance dependencies (Movement, QR and Binding) seems to be able to accommodate most of the asymmetries and problems reported in these pages, in the end providing a principled account of Extraposition and Heavy NP-Shift constraints without losing the ability to discriminate between classic “rightward” and “leftward” movement.

2. A case of rightward movement: Extraposition (EXT)

Let us first explore a set of well known distributional properties that characterize Extraposition (henceforth EXT) on the dimensions of what moves (§2.1), from where it does (§2.2) and where the extraposed constituent attaches (§2.3).

2.1. What

As mentioned before ((2), repeated in (4)) we can observe that Movement and EXT show an asymmetric sensitivity to the argument/adjunct distinction: while EXT is insensitive to such opposition, sub-extraction from NPs, in English ((4), Fox and Nissenbaum 1999), Italian (5) and Dutch (6),

show a pretty neat degradation if the extracted element is not a complement but an adjunct.

- (4) a. We saw [a painting _] yesterday *of John*
a'. *Of whom* did you see [a painting _]?
b. We saw [a painting _] yesterday *from the museum*
b'. **From where* did you see [a painting _]?
- (5) a. *pro* Abbiamo visto [un quadro _] ieri *di Gianni*
a'. *Di chi pro* hai visto [un quadro _]?
b. *pro* Abbiamo visto [un quadro _] ieri *dal museo*
b'. **Da dove* abbiamo visto [un quadro _]?
- (6) a. We hebben gisteren [een schilderij _] gezien *van Jan*.
We have yesterday a painting seen of J.
a'. *Van wie* heb jij [een schilderij _] gezien?
Of whom have you a painting seen
b. We hebben gisteren [sinaasappels _] gekocht *uit Spanje*
We have yesterday oranges bought from Spain
b'. **Waaruit / ?*waar* vandaan heb jij gisteren [sinaasappels _]
gekocht?
Where / from where from have you yesterday oranges
bought?

Another class of constituents that uniformly (English (7), De Vries 2006, Italian (8), and Dutch (9), De Vries 2002) undergoes Extraposition are restrictive relative clauses:

(7) [Some men _] appeared at the door *that Mary had been insulting*.

(8) Ho visto [un uomo _] oggi *che aveva perso la sua valigia*.

I have seen [a man _] today *who have lost his bag*

(9) Ik heb [de man _] gezien *die zijn tas verloor*.

I have [the man _] seen *who his bag lost*

More controversially¹ (see appositive/pseudo-relatives examples in (7).a-c (Citko 2006, De Vries 2003:65), (7).d (Smits 1989) and (7).f,g,h (De Vries 2006)), “appositive relatives, unlike restrictive relatives, do not undergo what is generally thought of as ‘Extraposition from NP’” (Emonds 1979:234, Vergnaud 1974:181)

(7) a. *[These men _] appeared at the door, *who Mary had been insulting*.

b. I have seen [Ann _] yesterday, *who is our director*.

¹ But this is not an issue I can deal with in these pages (see de Vries 2002). For sake of compactness I will limit the discussion to cases of extraposed rRCs.

- c. I met [John _] yesterday, *who I like a lot*.
- d. [Marie _] est là, *qui pleure comme une Madeleine*. (French)
 Marie is there, who cries like a fountain
 ‘Marie is there, and she is crying her heart out.’
- e. Ik heb [Joop _] gezien, *die twee zusters heeft*. (Dutch)
 I have Joop seen who two sisters has
 ‘I saw Joop, who has two sisters.’

In brief, these seem to be the basic properties that characterize the elements that undergo EXT:

- i. EXT, against movement, is insensitive to the argument/adjunct distinction;
- ii. PPs and rRCs can be extraposed and no crucial cross-linguistic variation (among English, Italian and Dutch) seems to hinge on that.

2.2. From where

Following Baltin (2006) let us call the position “related” to the extraposed constituent (i.e. the position “from where” EXT takes place) the *host* position. An host can be both an argument (10).a,b,c and an adjunct (10).d:

- (10) a. [A book _] appeared *which was written by Chomsky*
 b. I called [somebody _] yesterday *who I couldn't stand*

- c. I talked [to somebody _] about that *who was quite knowledgeable*
- d. I saw it [in a magazine _] yesterday *which was lying on the table*

However some restriction seems to apply when the constituent is moved: fronted PPs (11).a do not permit EXT to take place (prepositional stranding doesn't improve significantly, (11).a') while moved wh- constituents ((11).b, Culicover and Rochemont 1990) do.

- (11) a. *[In which magazine _]_i did you see it *t_i which was lying on the table?*
- a'. ^{??}[Which magazine _]_i did you see it in *t_i which was lying on the table?*
- b. [Who _] did you visit who was unhappy about the visit?

We should observe that while in languages like English or Dutch Extraposition from any argument/adjunct position is generally accepted, this is not the case for languages like Italian where, for instance, it is not possible to extrapose a PP/rRC from the pre-verbal subject position (independently from the transitive (12).a, intransitive (12).b, unaccusative (12).c, passive (12).d nature of the verb):

- (12) a. *[Un amico _] ha raccontato questa storia *di Gianni/che ho visto ieri*
A friend has told this story *of G. / which I have seen yesterday*

- b. *[Un amico _] ha parlato *di Gianni / che ho visto ieri*
 A friend has spoken *about G. / which I have seen yesterday*
- c. *[Un libro _] è uscito *di Chomsky/che è stato scritto da Chomsky*
 A book appeared *of C. / which was written by C.*
- d. *[Un libro _] è stato pubblicato *di Chomsky*
/ che è stato scritto da Chomsky
 A book has been published *of C. / which was written by C.*

Better results can be obtained with a post-verbal subject in Italian:

- (13) a. È uscito [un libro _] ieri *di Chomsky / che è stato scritto da C.*
 (it) appeared a book yesterday *of C. / which has been written by C.*
- b. Ha salutato Gianni [un signore _] ieri *che nessuno conosceva*
 (he_i) has greeted G. [a man]_i yesterday *who nobody knew*
 ‘A man who nobody knew has greeted G. yesterday’

On the other hand, EXT is fairly acceptable in Italian from object and indirect object, but it yields grammatical results only when the interrupting constituent is “light” (prepositional modifiers seem to block EXT):

- (14) a. Gianni ha mangiato [un panino _] ieri *che era avariato /*
con il prosciutto
 G. has eaten a sandwich yesterday *which was rotten /*

with ham

b. *Gianni ha mangiato [un panino _] in fretta *che era avariato/*

con il prosciutto

G. has eaten a sandwich in a hurry *which was rotten /*

with ham

(15) a. Gianni ha mangiato un panino [con un'amica _] ieri

che era ammalata /

di suo fratello

G. has eaten a sandwich with a friend yesterday

who was sick /

of his brother

b. *Gianni ha mangiato un panino [con un'amica _] in fretta

che era ammalata /

di suo fratello

G. has eaten a sandwich with a friend in a hurry

who was sick /

of his brother

Notice that both PP adjuncts, and an extra complement degrade EXT from direct object in Italian:

(16) a. Gianni ha spedito [una lettera _] ieri *senza francobollo*

G. sent a letter yesterday without stamp

a'. ??Gianni ha spedito [una lettera _] a Maria ieri
senza francobollo

G. sent a letter to M. yesterday
 without stamp

Another important restriction that seems to be subject to some degree of cross-linguistic variation (English, Italian Vs. Dutch) is constituent-boundedness, i.e. embedded constituents cannot be hosts in English (17).a / Italian (17).b while they can in Dutch (17).c (De Vries 1999):

- (17) a. *[A review of [a book _]] appeared *by three authors*
- b. *È apparsa [una recensione [di un libro _]] ieri *di tre autori*
- b. Ik heb [de papieren [van de man _]] gecontroleerd *die een rode koffer droeg.*
- I have [the papers [of the man _]] checked who a red suitcase carried

In sum:

- i. both arguments and adjuncts can be host for EXT in English and Dutch;
- ii. in Italian only the last argument/adjunct can be an host for EXT, provided that no “heavy” (prepositional) modifiers intervene between the extraposed constituent and the related host;

- iii. EXT from embedded constituents is impossible in English/Italian but it seems to be possible in Dutch;

2.3. *Where*

Looking at standard C-command tests, an extraposed element is not C-commanded by its host (Culicover and Rochemont 1997), since, if we accept that the pronominal object C-commands the second object (Larson 1988), we get a principle C effect:

- (18) a. I sent her_i [many gifts _] last year *that Mary_i didn't like*.
b. *I sent her_i [many gifts *that Mary_i didn't like*] last year.

Things get more complicated if we look at Condition C bleeding in complements ((19).a) vs. adjuncts ((19).b) (van Riemsdijk and Williams 1981, Lebeaux 1988, Taraldsen 1981, Fox and Nissenbaum 1999):

- (19) a. ^{??/*}[Which book *about John_i's library*] did he_i read _?
a'. [Which book *from John_i's library*] did he_i read _?
b. I gave him_i [a picture _] yesterday *from John_i's collection*.
b'. ^{??/*}I gave him_i [a picture *from John_i's collection*] yesterday.
c. I gave him_i [an argument _] yesterday *that supports John_i's theory*.

- c'. ^{??/*}I gave him_i [an argument that supports John_i's theory]
yesterday.
- d. ^{??/*}I gave him_i [a picture _] yesterday *of John_i's mother*.
- e. ^{??/*}I gave him_i [an argument _] yesterday *that this sentence
supports John_i's theory*.

Looking at these data, the pre-theoretical generalization we could make is that an object “reconstructs” under the host while an adjunct does not.

Other scope diagnostics, like the ‘Free choice’ *any* (that should be licensed in the scope of *look for/would*, Fox and Nissenbaum 1999), seem to point out that, when an extraposed constituent is an adjunct, the scope of the host is at least as high as the attachment site of the extraposed adjunct:

- (20) a. I looked very intensely for [anything *that would help me with my thesis*].
- b. *I looked for [anything _] very intensely *that will/would help me with my thesis*.
- c. I looked for [something _] very intensely *that will (likely) help me with my thesis*. (wide scope forced)
- d. I would buy [anything _] without making a fuss *that will/would help me with my thesis*. (under the scope of *would*)

It has been proposed that constituents extraposed from subjects attach to IP (availability of stranding if the VP is elided (21).a-b), while constituents extraposed from objects attach to VP (no stranding option if the VP is elided (21).c-d), i.e. an extraposed phrase is adjoined to the first maximal projection that dominates the phrase in which it originates (Baltin 1981, 2006). This is sufficient also to account for the classical nested dependency ((22).a Vs. (22).b):

- (21) a. Although [IP not [many people _] would [VP ride with Fred] *who knew just him*], some [IP would [VP_] *who knew his brother*].
- b. Although [IP [no reviews _] [VP appeared] *of Chomsky's book*], one [IP did [VP ___] *of Jakobson's book*].
- c. *Although he didn't [VP call [people _] up *who are from Boston*], he did [VP ___ *who are from New York*].
- d. *Although he didn't [VP call [people _] up *from Boston*], he [VP _ *from New York*].
- (22) a. [IP [Someone _1] [VP picked [some books _2] up [which were lying on the table]_2 VP] [who really didn't want to]_1 IP]
- b. *[IP [Someone _1] [VP picked [some books _2] up [who really didn't want to]_1 VP] [which were lying on the table]_2 IP]

Such constraints on the attachment site have been captured by Williams' generalization (1974):

(23) *Williams' generalization*

When an adjunct β is extraposed from a “source NP” α , the scope of α is at least as high as the attachment site of β (the Extraposition site).

3. Another case of rightward movement: (Heavy) NP-Shift

It has been proposed (Ross 1967) that arguments reordering could also be an instance of rightward movement; this is the case, for instance, when the direct object crosses over the indirect one or an adjunct (Staub et al. 2006):

(24) Lucy ate t_i with a fork [the extremely delicious, bright green broccoli]_i

This analysis has been seriously challenged by Larson (1988), Kayne (1994), somehow reformulated in Jackendoff (1990) and integrated by Belletti and Shlonsky (1995). In the following pages I will try to highlight some interesting parallelisms/differences that relates Heavy NP-Shift (henceforth HNPS) and EXT.

3.1. What

Looking at HNPS in English, we can observe that both direct and prepositional objects can undergo the alleged shifting operation:

- (25) a. I gave [the books which my uncle left to me as part of his inheritance] to Bill
- b. I gave _ to Bill [the books which my uncle left to me as part of his inheritance]
- c. I talk _ all the time [to my uncle who left me an enormous inheritance]

Despite the fact that such operation seems to be optional, a fairly evident bias (Pinker 94:131) is reported for preferring shifted versions ((26).a) of “heavy”² constituents ordering among many logically possible non-shifted alternatives (e.g. (26).b):

- (26) a. In my laboratory we use it as an easily studied instance of mental grammar, allowing us to document in great detail

² Usually we refer to a “heavy” NP as a NP containing a sentential/relative clause (Ross 1967). This (insufficient) definition will be better discussed in §5.

the psychology of linguistic rules
from infancy to old age
in both normal and neurologically impaired people,
in much the same way that biologists focus on the fruit fly
Drosophila to study the machinery of the genes.

b.^{??}In my laboratory we use it as an easily studied instance of
mental grammar, allowing
us to document
in much the same way that biologists focus on the fruit fly
Drosophila to study the machinery of the genes
in both normal and neurologically impaired people,
in great detail
the psychology of linguistic rules
from infancy to old age

In general, “light” NP-shift is not possible in English ((27).a’):

- (27) a. I gave the books to Bill
a'. *I gave to Bill the books

While in Italian “shifting” seems to be more free³ in allowing arguments reordering also in sentences like the one in (27):

- (27) b. *pro* ho dato a Bill i libri
b'. *pro* ho dato i libri a Bill

In brief:

- i. shifting is an optional (though preferred) operation that targets only “heavy” elements in English;
- ii. in Italian, argument reordering seems to be able to target also “light” constituents.

3.2. *Where and why*

Evidences that the “shifting” operation is not (always) a uniform “rightward” movement comes from Italian (Belletti and Shlonsky 1995): since the clitic *ne* (literally “some of them”) can be extracted only from the base object position (Belletti and Rizzi 1981, Burzio 1986, (28).b Vs. (29).b) the *ne* cliticization test suggests that, in case of linear PP NP order,

³ Different orders have different informational structure implications (Belletti and Shlonsky 1995, Zubizarreta 1998) that will be discussed in §3.2.

only when the NP object is heavy (31).a or focalized (31).b it seems to occupy the relevant base object position that licenses *ne* extraction ((30).a Vs. (30).b):

(28) a. Ho letto molti libri.

I read many books.

b. *Ne* ho letti molti.

I of-them read many.

(29) a. Hanno lavorato molti operai.

have worked many workers.

b. *[?]*Ne* hanno lavorato molti.

of-them have worked many.

(30) a. *Ne* ho dato/dati uno/tre a Gianni.

of-them (I) have given one/three to Gianni.

b. *[?]*Ne* ho dato/dati a Gianni uno/tre.

of-them (I) have given to Gianni one/three.

(31) a. *Ne* ho dato/dati a Gianni uno/tre *che mi avevano consigliato la settimana scorsa*.

of-them (I) have given to Gianni one/three that to-me (they) have suggested last week

‘I gave to Gianni one/three which they suggested to me last week’

b. *Ne* ho dato a Gianni *uno solo*.

of-them (I) have given to Gianni one only

On the basis of these facts, Belletti and Shlonsky (1995) propose that when the light object is postponed, it does not occupy anymore the object position (*Light NP Postposing*). This strategy is specific to Italian (and not available in English) because of the availability in such language of subject/object *pro* and it results in a focalization of the shifted constituent (this reminds of the obligatory focalization of post-verbal subject, Calabrese 1982, (32).a). Assuming that the focalized position to the “right edge” of the VP is unique, they can capture the impossibility of “light”-shifting both of the subject and the object at the same time⁴ ((32).b):

- (32) a. ?Ha dato un libro a Maria Gianni.
has given a book to Maria Gianni.
b. *(?)Ha dato a Maria un libro Gianni.
has given to Maria a book Gianni.

This solution however seems to be adequate only for light NPs shifting but not for the (adverbially) focalized object in (31).b for which Belletti and Shlonsky (1995) propose that PP must scramble before the NP, which, in

⁴ Italian post-verbal subjects are more restricted than in other null subject languages such as Spanish (Belletti 2001: 70f.).

fact, stays in situ allowing *ne* extraction. This solution is supported by the behaviour of the weak dative pronominal element *loro* that is forced to move closer to the verb (Cardinaletti 1992, (33).a) as predicted in the PP scrambling hypothesis; in this case, *ne* extraction from the direct object is allowed as expected ((33).b):

- (33) a. ho dato *loro* un libro.
 (I) have given to-them one book.
 b. *ne* ho dato *loro* uno.
 of-them (I) have given to-them one.

Unfortunately the scrambling analysis of PP, does not explain how the “shifted” heavy-NP behaves as if it were in a A'-position (instead of an A-position according to the *ne* cliticization test) when it licenses parasitic gaps:

- (34) a. I crossed _ without recognizing _ *a classmate with whom I attended primary school.*
 b. Ho incrociato _ senza riconoscere _ *quel compagno con cui avevo frequentato le scuole elementari.*
 (I) crossed without to recognize that classmate with whom (I) attended primary school.
 c. *Ho incrociato _ senza riconoscere _ *quel compagno.*
 (I) crossed without to recognize that classmate.

The contrast between (34).b (heavy NP) and (34).c (light NP) shows that heaviness is in fact the crucial property that, in a cross-linguistic uniform way, triggers a displacement operation that removes the direct object from its base position. This is the classic understanding of HNPS.

Before concluding this paragraph, I wish to stress a symmetry between HNPS and movement depending on the source position in double objects constructions: extractions from the first object ((35).a,b) seem to be much more difficult than from the second one ((35).c,d), Culicover and Rochemont 1997:15):

- (35) a. Bill [gave [John] _] yesterday *the book that he was looking for*
b. What_i did Bill [gave [John] t_i] yesterday
c. *Bill [gave _] the book yesterday *anyone who wanted it*
d. *Who did Bill [give t_i] the book yesterday

Summarizing the main facts discussed in this section:

- i. argument re-ordering is cross-linguistically sensitive at least to two factors: heaviness (classic NP-shifting), focalization of light constituents and availability of empty pronominal elements (to satisfy substantive criterial and/or verbal theta-role requirements and to license PP scrambling);

- ii. movement and HNPS behave the same (both degraded) with respect to extraction from the first object in double object constructions.

4. Some of the solutions proposed and their problems

In this section I will briefly review the main analyses that have been proposed to account for the data described in §2 (and partly in §3, which have been already discussed in §3.2). As it will be clear, none of them results in a sufficient coverage of the empirical evidence reported, especially in accounting for the three properties discussed in §1.0 (clause-boundedness, definiteness constraint and adjunct/argument distinction insensitivities to EXT).

4.1. Analysis 1: classic rightward movement

The first solution that has been proposed to account for EXT (this is straightforwardly also suitable for HNPS) has been a *rewriting transformation* (Rosenbaum 1965:11), adapted by Ross (1967) as follows:

(36) Extraposition from NP (Ross 1967:4):

X	[NP	-	S]	Y	→
					OPTIONAL
	1		2	3	1, ∅, 3+2

According to this rule, the underlying structure of (10).a is (10).b, repeated below:

- (10) a. [A book _] appeared *which was written by Chomsky*
b. [A book *which was written by Chomsky*] appeared

It is easy to show that this rule is too powerful and it cannot block ungrammatical transformations like (37).b:

- (37) a. [S [NP [S That [NP a gun [S which I had cleaned]]] [VP went off]]]
[VP surprised no one]]
b.*[S [NP [S That [NP a gun _] [VP went off]]] [VP surprised no one]
[S which I had cleaned]]

For this reason Ross proposes an ad-hoc constraint, that, in fact, excludes (37).b forcing clause-boundedness:

- (38) *Right Roof Constraint (RRC, Ross 1967)*

an element cannot move rightward out of the clause in which it originates.

At least three things remain unexplained under this fairly simple analysis:

- a. why movement is much more constrained on the right than on the left?
- b. what triggers Extraposition?
- c. how to account for the “last-cyclic” nature of the rule?

4.2. *Analysis 2: base generation*

The movement analysis has been criticized because of an incomprehensible asymmetry with respect to standard leftward movement (Akmajian 1975): why should *wh*-movement escape subjacency by successive cyclic applications of the movement operation while EXT cannot? Also from a licensing perspective, the movement analysis poses some problems: for instance, when a rRC is extraposed from a subject position, the trace it leaves would violate the Empty Category Principle (Chomsky 1981). To solve these problems an alternative analysis has been proposed, among others, by Guéron and May (1984), Culicover and Rochemont (1990-97); this analysis relies on the alleged satisfaction of the *Complement Principle* at some point in the derivation:

(39) *Complement Principle* (Guéron and May 1984)

In a sequence of categories $\alpha_i, \beta_i^1 \dots \beta_i^n$ in a structure $\Sigma, \beta_i^1 \dots \beta_i^n$ are complements to α_i only if α_i governs β_i^n (where α governs β iff α and β

are dominated by all the same maximal projections, and there are no maximal projection boundaries between α and β).

This principle obviously cannot apply at the surface structure, where the relevant locality (government) requirements between the host and the extraposed constituent are not met. They propose, then, that the principle must be satisfied at LF, crucially after Quantifier Raising (QR). This intuition elegantly accounts for the “locality” (clause-bounded because of QR) and for the definiteness constraint (no quantification, no QR) of EXT, but it contains at least one important flaw: the complement principle does not make any distinction between arguments and adjuncts, but this seems to be a very productive distinction as shown in (19);

4.3. Analysis 3: modification based account

Kayne (1994) proposes that the host and the extraposed constituent are generated together, then the host moves (to a C-commanding position) while the extraposed element is stranded in its base position.

Such analysis is the only possible solution assuming the Linear Correspondence Axiom (and its implicit ban on rightward movements⁵). This proposal, easily accounts for the *Definiteness Constraint* (Diesing 1992, (3)) since non-constituents cannot be moved; given the relative clause structure and the DP structure proposed by Kayne, this analysis is pretty straightforward (“the book” in (40).a is not a constituent, then it cannot move, while “two/a/those book(s)” it is, so it can move stranding the RC behind as expected):

- (40) a. [DP[D the]][CP[NP book_i][C[C that]][TP I [VP[V read]][NP t_i]]
 b. [DP[D 0]][CP[NP two/a/those book(s)_i][C[C that]][TP I [VP[V read]][NP t_i]]

Under this perspective the Right-Roof Constraint (38), is replaced by theta-role requirements and/or by LCA:

- (41) *The fact that [somebody]_i walked into the room is irrelevant [t_i who I knew]. (Kayne 1994:118)

This account however presents some problems:

⁵ If A asymmetrically C-commands B, then terminals dominated by B could not precede terminals dominated by A: this is why rightward adjunction could not be linearized.

- a. extraposed constituents seem not to be C-commanded by their host (e.g. (18));
- b. Examples like (10).c remain unexplained (since [P NP] should not be considered a constituent using Kayne's analysis);
- c. a stranding analysis would not easily predict the nesting dependencies reported in (22).

4.4. Analysis 4: "mixed" account

The last proposal I would like to quickly review is the solution proposed by Fox and Nissenbaum (1999). The general idea is that "overt" and "covert" movement simply differ for the pronunciation of the head of the chain (overt) Vs. pronunciation of the tail of the chain (covert). The relevant covert operation, in this case, as in Guéron and May's (1984) analysis, is QR; what is crucial in this proposal is the assumption that covert operations, i.e. QR, can precede overt ones: EXT, then, is the result of QR plus late merge of an adjunct. As for EXT of complements, since they need to satisfy thematic requirements (e.g. Complement Principle), they need to be merged in situ and QRed afterward.

This is the schematic derivation they propose:

- (42) i. [_{CP} John [_{VP} [_{VP} picked [_{DP} a book] up]
- ii. [_{CP} John [_{VP} [_{VP} picked [_{DP} a book] up] [_{DP} a book]]] (QR)

iii. [CP John [VP [VP picked [DP a book] up] [DP [DP a book] [CP which he really enjoyed]]]] (late merge)

iv. [CP John [VP [VP picked [DP a book] up] [DP ~~[DP a book]~~ [CP which he really enjoyed]]]]

(PF deletion of the head of the QRed nominal chain)

It is somehow surprising that the rightward (not leftward, as usually assumed) QR operation is not discussed in any detail in their work; nonetheless, this analysis allows us to capture many interesting facts:

- i. the clause-boundedness of Extraposition is determined by the clause-boundedness of QR (as in Guéron and May (1984), Culicover and Rochemont (1990-97));
- ii. the definiteness constraint is readily captured: definite NPs cannot be QR-ed (as in Guéron and May (1984), Culicover and Rochemont (1990-97));
- iii. late merge of adjuncts Vs. complements captures the condition C bleeding effect (e.g. (19)).

Moreover the late merge hypothesis would correctly predict that extraposed RCs do not reconstruct (Wilder 1995):

(43) a. *We talked [about her_i claim _] yesterday *that Mary_i will hire Peter.*

b. I gave him_i [an argument _] yesterday *that supports John_i's theory*

Despite the coverage of this elegant and relatively simple analysis, one main question remains unanswered: why QR is on the right?

5. The proposal: changing the derivational perspective

To account for the set of properties discussed in these pages, that EXT and HNPS seem to show, I propose adopting a radically different minimalist derivational perspective: tree structures must be built not from bottom to top (and, partly, “from right to left”, at least in alleged right branching languages), as generally assumed, but from left to right, expanding lexical “expectations” in a principled, top-down, way.

5.1. Deriving phrase structures Top-down, from left-to-right

A left-to-right (Phillips 1996) top-down (Chesi 2004-07) derivation is a generative procedure that operates on lexical items using minimalist structure building operations so as to isolate the best possible approximation of the (infinite) set of well formed linguistic expressions.

The lexicon is the core of our grammar and it has to be fully specified⁶: a lexical item is an ordered feature structure⁷ composed by phonetic/semantic features (that we cannot discuss here), functional features (characters prefixed with a “+” sign, expressing functional properties/positions, e.g. +D(eterminer)), selection features (characters prefixed with a “=” sign, expressing thematic requirements in terms of necessary additional constituents that must be present in the structure in order for the expression to be grammatical) and categorial features (unprefixed characters expressing N(ominal), V(erb) and A(djectival/dverbial) heads). This is how a lexical item (“John *runs*”) should look like⁸:

(44) < +T_{pres} +Agr_{3-sing} =<+D N> V *runs* >
 features: *functional* *selectional* *categorial* *phon/sem*

⁶ The mainstream Minimalist Program has often been accused to be deficient in this respect (Johnson and Lappin 1997, but see also Holmberg 2000); I follow Stabler 1997 and Adger 2007 in this respect.

⁷ Angled brackets express ordered sets: e. g. <A, B, C> or simply <A B C>

⁸ Let us take this as a simplified example. Obviously we could decompose such lexical items in its morphological subparts (<=<+D N> V *run*> <+T_{pres} ∅> <+Agr_{3-sing} *s*>) but this would require extra machinery to be discussed. Similar considerations hold for features: in this paper we can assume a pure privative system (Adger 2007), but full feature structures could make some task much more elegant (Pollard and Sag 1994, Chesi 2004, Adger 2007).

From this perspective, structure building operations Merge, Move and the Phase idea are reformulated according to such directionality constraint in the following (informal) way⁹:

(45) **Merge**: it creates (temporary, in the sense of Phillips 1996) constituency as a result of an unification procedure (Shieber 1986) that produces, if possible, a single rooted feature structure (i.e. a subtree with lexical items at leaves); compatibility among features is determined by universal constraints (i.e. Functional Sequence and the Linearization Principle, (48)-(50)); items to be merged can be, in this preference order, moved constituents (see below) or new items picked out from the lexicon;

(46) **Move**: it creates Long Distance (e.g. thematic) Dependencies by means of a *Memory Stack*: this is a memory buffer with a Last In First Out structure (the last element moved in such buffer is the first element to be re-merged in the structure); constituents are “moved” in the memory buffer when they are not “selected” (in a technical sense, that is, the lexical head with an appropriate select feature (e.g. =<+D N>) is not yet merged within the processed constituent when the item at issue

⁹ See Chesi 2007 for the full formalism and for some arguments on its generative power. This formalization (Phase-based Minimalist Grammars, PMGs) is a modification of Stabler’s 1997 proposal (Minimalist Grammars, MGs).

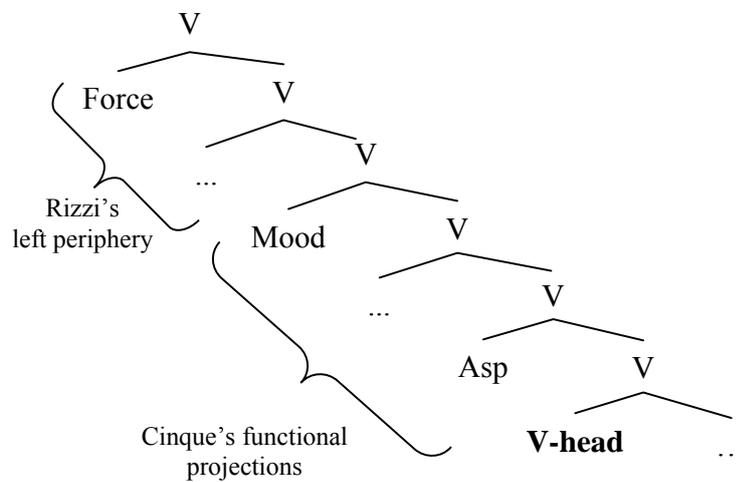
has been merged, e.g. [_v <+D N John>]) and retrieved/re-merged in the structure as soon as possible according to some Universal Principle that regulates the merge operation and the content of the memory buffer ((50)).

- (47) **Phase Projection:** depending on the (select) features present in the processed lexical item, this operation expands the feature structure introducing the minimal set of dominance relations complying with the selective requirements (the order of projections is determined by the select features order within the lexical item: <=<+D +Anim N> =<+D +Edible N> V *eat*> → [V *eat* [+D +Anim N] [_v [+D +Edible N]]]);

Some clarification is necessary in order to understand how the derivation unrolls: first of all, let me clarify the notion of *phase*. This is tightly related to the minimalist intuition (Chomsky 1999, 2005), namely a phase is a domain with the best degree of completeness with respect to interface requirements (i.e. thematic necessities, completeness in feature interpretability, full linearization), but it obviously requires some non trivial changes in order to be compatible with such unconventional directionality shift: assume that CP and D/PP are phases also in the new system; assume, moreover, that standard edge positions are in fact (part of the) expanded Functional Sequence (henceforth FS, e.g. for the CP layer we could assume

Rizzi's (1997) left periphery) and that the lexical items related to such FS are respectively Nominal and Verbal items¹⁰. The simplest assumption is considering FSs as Extended Projections (Grimshaw 1991) headed by an unique Nominal or Verbal head, accordingly, as exemplified for V in (48).

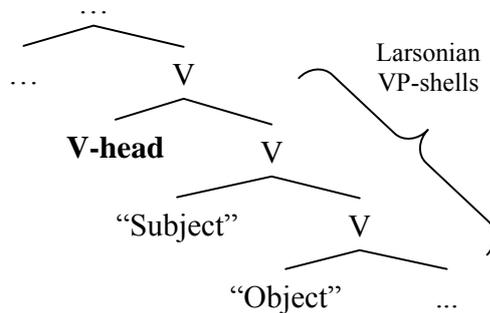
(48)



On the right side of this phase head, I will attempt to follow Larson's approach introducing VP-shells:

(49)

¹⁰ Somehow similar considerations should apply to Adjectival lexical items as well.



Any thematic requirement should be accommodated following a standard universal order (Uniform Theta-Role Assignment Hypothesis, Baker 1988) and an appropriate subcategorization of the phase head (according to this formalism, by means of select features).

The strongest assumption we want to pursue is that this skeleton is universal, that is, the FS always precedes the phase head while arguments always follow it. This can be guaranteed by a universal principle that uniformly maps precedence and dominance in a way that is reminiscent of (and compatible with, though weaker) Kayne’s Linear Correspondence Axiom (1994):

(50) *Linearization Principle*

- a. $\langle A, B \rangle$ if A (is a lexical head and) *selects* B as an argument
- b. $\langle B, A \rangle$ is B is a *functional specification* of A.

Obviously this order is (almost) never trivially evident in most languages, but this is why we need to implement movement/non local dependencies in our grammar: in this framework, we are able to distinguish *selected phases*, i.e. arguments that follows the phase head that do not need to be “moved”, from *unselected* ones, i.e. true adjuncts, preverbal arguments, relatives etc. Since arguments (e.g. <+D N> items) need to be selected (because of full interpretation any node has to be attached), when introduced in the phrase structure in a pre-head position¹¹, they are moved in the memory buffer. When a phase head is processed, phase projection will introduce new empty subtrees according to the select features in the selected lexical head and, depending of the element in the memory buffer, there will be an attempt to merge the most prominent item in the memory buffer within the first verbal shell. The structure of the memory buffer guarantees locality, therefore nested long distance dependencies can be readily captured. If the merge is successful, the re-merged item will be removed from the memory buffer and another element (if any) will be available for further merge. If the merge is unsuccessful or the buffer is empty a new element can be merged directly from the lexicon¹².

¹¹ This happens in order to satisfy a (criterial, in the sense of Rizzi 2006) functional feature, e.g. the subject/topic position in a preverbal location above the verbal head.

¹² Notice that in this system the standard minimalist notion of “Move” preempts the notion of “Merge”. This seems to be empirically tenable (Richards 1999).

5.2. Merge, Movement and Phase projection at work: nesting and the special status of the last selected argument

Summing up, there are three structure building operations that operate on lexical items; all of them access the features contained in these lexical items and delete/modify/complete them within a unified feature structure (i.e. a tree, that is the history of the derivation). There are some logical possibilities on the order of application of these three operations that crucially result in different derivations/structures¹³. I assume that in the default case, we first apply Phase Projection, then Merge and in the end Move.

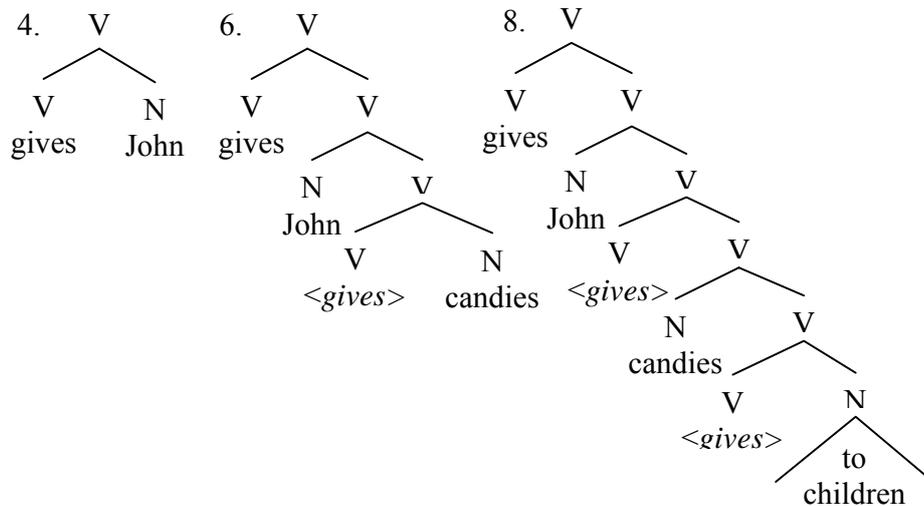
Below there is a sample derivation that shows how Phase Projection (PhP) and Merge operate in the case the Move buffer be empty¹⁴:

(51) Lexicon: {[=DP =DP =PP V gives], [+K to], [+D N John],
 [+D N children], [+D N candies]}

¹³ Interleaving in a way or another structure building operations results in grammars with different generative power; unfortunately, this is not an issue I can discuss here (Chesi 2007).

¹⁴ I will use squared brackets to mark feature structures/constituency, but remember that features within the feature structure are ordered; =DP/=PP are meant to be shortcuts for =[+D N] and =[+K +D N] respectively, where +K is a case feature.

1. *PhP*(=V) → [_v] (this is the default phase projection: it simply instantiates a verbal phase)
2. *Merge* ([_v],[=DP =DP =PP V gives]) →
[_v=DP =DP =PP V gives]
3. *PhP*([=**DP** =DP =PP V gives]) → [_v=DP =PP V gives [+D N]]
4. *Merge* ([=DP =PP V gives [+**D N**]], [+**D N** John]) →
[_v=DP =PP V gives [+D N John]]
5. *PhP*([_v=**DP** =PP V gives [+D N John]]) →
[_v V gives [+D N John] [_v =PP V [+D N]]]
6. *Merge*([_v V gives [+D N John] [_v =PP V [+**D N**]]],
[+**D N** candies]) →
[_v V gives [+D N John] [_v =PP V [+D N candies]]]
7. *PhP*([_v V gives [+D N John] [_v =**PP** V [+D N candies]]]) →
[_v V gives [+D N John] [_v V [+D N candies] [_v V [+K +D N]]]]
8. *Merge* ([_v V gives [+D N John] [_v V [+D N candies]
[_v V [+**K +D N**]]], [+**K +D N** to children]) →
[_v V gives [+D N John] [_v V [+D N candies]
[_v V [+K +D N to children]]]]



In this (simplified) derivation *Phase Projection* creates an empty (i.e. unlexicalized) constituent to the right edge of the structure depending on the first select feature of the processed lexical item; *Merge* fills this edge position either with the most prominent element in the memory buffer (if any, and this is not the case in the example) or with an element from the lexicon. The unification procedure, which is a necessary part of the Merge operation, will decide whether or not the features are compatible and, in case, which constituent will be created.

Let us now turn to *Movement*. This operation is tightly constrained by the notion of phase: the reason for using phases is that if we want to deal with computable derivations, we should keep our “working memory” as light as possible. In this sense, the minimalist idea that, at some point, the phase is “shipped-out”, that is, no further operations can tamper the set of dominance/precedence relations created up to this point, is exactly what we

want to implement. We decided (Chesi 2004, Bianchi and Chesi 2005-08) to fix this “memory limit” at the last operation triggered by a phase head, that is, the Phase Projection of the last selected complement. This produces some computational advantages (Chesi 2004-07) and allows us to make a crucial distinction between the right recursive branch of the tree, i.e. the last selected complement (the *sequential phase*, Bianchi and Chesi 2005-08), that once projected close the previous phase, and the *nested phases* (Bianchi and Chesi 2004-08), i.e. unselected phases and phases merged in the FS. Posing the constraint that every phase has its own memory buffer and that the inheritance of the content among memory buffers is sensitive to the sequential/nested distinction, we can constraint in an empirically tenable way the usage of the memory buffers and, then, the movement operation:

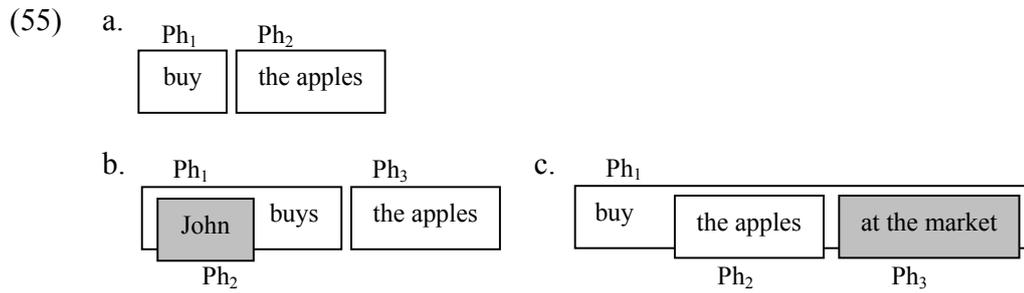
(52) *Constraint on the memory buffer inheritance*

the content of the memory buffer is either integrated within the phase, or discharged in (the memory buffer of) the LAST selected phase.

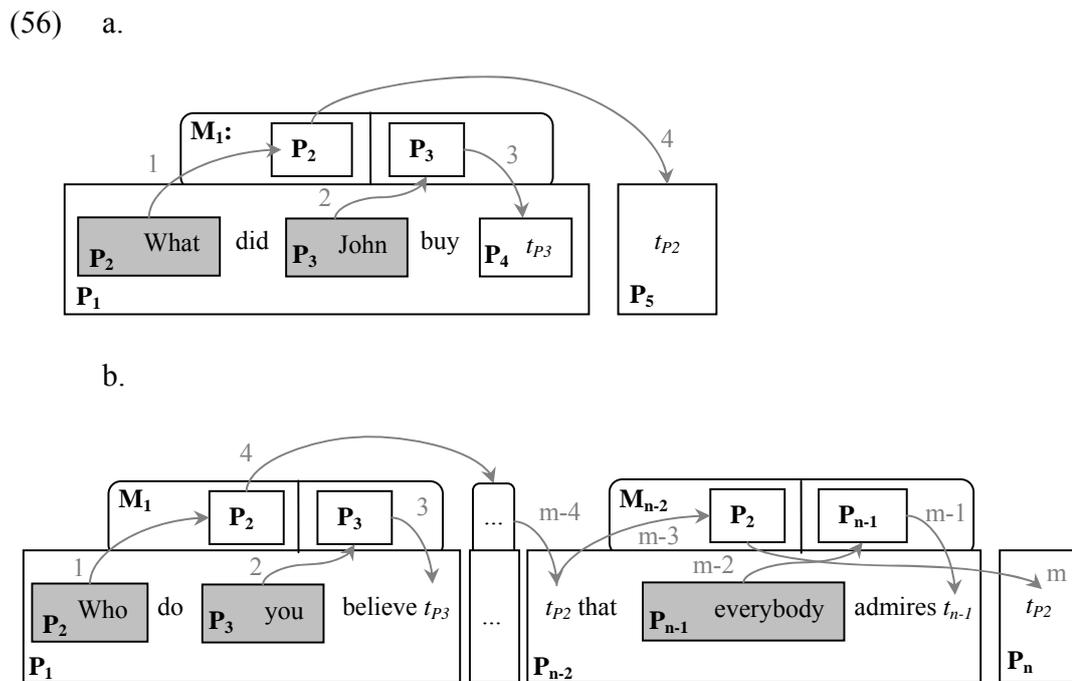
(53) *Success condition*

memory buffers must be empty at the end of the computation.

This is a graphical representation of the derivation of a simple SVO sentence in English (numbers on the arcs indicate the derivational sequence).



These assumptions are sufficient to capture (successive cyclic, (56).b) movement ((56).a) predicting the correct locality constraints¹⁶:



In (56).a, the steps of the derivations are the following ones:

¹⁶ In the original proposal, Chesi 2004, the memory buffer was structured in slots in order to capture *Relativized Minimality* effects (Rizzi 1990). The discussion of this point would unnecessarily makes this presentation more complex.

1. a default verbal phase is projected (P_1):

$$\mathbf{PhP}(Wh\text{-question}) = [+C_{wh} +T +K_{nom} V]^{17}$$

2. since the verbal phase is interrogative, this functional feature has to be explicitly marked; in English this can be done by merging the relevant *wh*-element within the specific “criterial” position. This is how $[+D_{wh} N \text{ what}]$, phase P_2 (computed as a nested phase), is introduced in the derivation:

$$\mathbf{Merge}([+C_{wh} +T +K_{nom} V], [+D_{wh} N \text{ what}]) = [+C_{wh}[+D_{wh} N \text{ what}] +T +K_{nom} V]$$

3. since it is unselected (by Linearization Principle) it is inserted (step 1) in the memory buffer (M_1) of the matrix V-phase (P_1):

$$\mathbf{Move}([+D_{wh} N \text{ what}]) = M_1 < [+D_{wh} N \text{ what}] >$$

4. *did* is compatible (unifiable) with a tense functional specification of the matrix V-phase, then licensed in this position:

$$\mathbf{Merge}([+C_{wh}[+D_{wh} N \text{ what}] +T +K_{nom} V], [+T \text{ did}]) = [+C_{wh}[+D_{wh} N \text{ what}] +T[\text{did}] +K_{nom} V]$$

5. $[+D N \text{ John}]$ (phase P_3 , again computed as a nested phase) is introduced to satisfy a subject-criterial (in the sense of Rizzi 2004a) requirement (functional specification of P_1 , i.e. $+K_{Nom}$) and moved in the

¹⁷ It is fair to assume that aux-subject inversion is decided (as parameterized option) at this level. *PhP* stands for the Phase Projection function.

memory buffer since it is unselected (step 2):

$$\mathbf{Merge}([+C_wh[+D_wh\ N\ what] \ +T[did] \ +K_nom\ V], [+D\ N\ John]) =$$

$$[+C_wh[+D_wh\ N\ what] \ +T[did] \ +K_nom[+D\ N\ John] \ v]$$

$$\mathbf{Move}([+D\ N\ John]) = M1 < [+D_wh\ N\ what], [+D\ N\ John] >$$

6. then $[= [+D\ N] \ = [+D\ N] \ v\ buy]$ is processed as the head of the matrix V-phase (P_1). Since it has two selection requirements to be satisfied (an agent and a patient, both N-phases), these select features will project two phases, P_4 and P_5 .

$$\mathbf{Merge}([+C_wh[+D_wh\ N\ what] \ +T[did] \ +K_nom[+D\ N\ John] \ v], [= [+D\ N] \ = [+D\ N] \ v\ buy]) = [+C_wh[+D_wh\ N\ what] \ +T[did] \ +K_nom[+D\ N\ John] \ v[= [+D\ N] \ = [+D\ N] \ v\ buy]]$$

$$\mathbf{PhP}([+D\ N]) \text{ (per two)} = [+C_wh[+D_wh\ N\ what] \ +T[did] \ +K_nom[+D\ N\ John] \ v[[buy\ [+D\ N]]\ [+D\ N]]]$$

7. P_4 is a nested phase and will be unified by (re-)merging P_3 (the first accessible element in $M1$);

$$\mathbf{Merge}([\dots \ v[[buy\ [+D\ N]]\ [+D\ N]]], M1[John]) = [\dots \ v[[buy\ [+D_nom\ N\ J.] \ [+D\ N]]]]$$

8. P_5 is the last Phase Projection: it is a selected/sequential phase where the last selectional requirement of the previous phase will be lexicalized by merging P_2 : $\mathbf{Merge}([+D\ N], M1[what]) = [+D_wh\ N\ what]$

Reiterating these steps recursively we can easily derive the successive cyclic movement in (56).b.

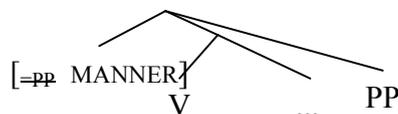
5.3. Right hand adjuncts

Following the Linearization Principle, we would not expect adjuncts, which are related to a functional position, to be placed on the right of the head¹⁸. To explain this fact, we need either to rely on an interleaving option between Phase Projection and Merge, or to the possibility of stacking the Phase Projections after the last selected complement. Before discussing which solution better fits with the data under analysis, let us understand how we can try to determine the head – adjunct order we expect: assume that this “shifting” operation happens when a functional feature is “complex”, that is, simply, when it bears a select feature. In this case the two possibilities should be available (delayed Phase Projection or Stacking), we would obtain something like (57).b from (57).a (notice the minimal difference between an adjunct selected by a functional feature (57).a and a selected (prepositional) argument (57).c).

(57) a.



b.



¹⁸ Notice that this is not a trivial issue for any anti-symmetric framework (e.g. Kayne 1994).

c.



There are reasons to believe that such solution is not arbitrary and it is probably justified by complexity reasons (Chesi 2007) that suggest to “minimize nesting” as long as the functional feature is not criterial and the dependency between the selecting feature/position and the “shifted” element is recoverable. On this line we can state a general principle that should allow us to predict shifting of “complex” constituents:

(58) *Minimize Nesting*

When a nested phase bears select features (i.e. it is *heavy/complex*), it would rather be (partially) processed in a phase-peripheral position (i.e. on the right), unless licensed functional/thematic features be unrecoverable.

(Partially) Processed in a phase-peripheral position means that the whole constituent could be shifted (unless functional/thematic features be unrecoverable) or else, only the selected phase(s) be projected in a phase peripheral position. The definition of Phase Projection however does not allow to freely project under non-local conditions, then the phase head has to be somehow moved in the new peripheral position. The next paragraph suggests a trigger for this operation.

As for the two available solutions, they are clearly distinct as shown below:

(59) a. Delayed Phase Projection

$$[[+A =A] [+B =B B] C \dots] \rightarrow [[+A [+B [C \dots [A] [B]]]]]$$

b. Stacking

$$[[+A =A] [+B =B B] C \dots] \rightarrow [[+A [+B [C \dots] B] A]]$$

The mirrored prediction made by stacking seems to be more promising (remember the nesting dependency discussed in (22)) and it will be first considered.

5.4. Rightward Quantifier Raising

In a left-to-right derivation there is no room for a leftward movement in the classical sense. Quantifier Raising is not an exception in this respect. What we proposed (Bianchi and Chesi 2007) is that the long distance dependency through memory buffers fits with QR as well¹⁹. There are obviously differences with respect to the Move operation that we need to capture:

- i. QR is covert (in the sense that the head of the chain is unpronounced);

¹⁹ As an anonymous reviewer noticed, this proposal somehow reminds of the lowering operation discussed in McCawley (1999).

- ii. it is not feature-driven²⁰;
- iii. it cannot be freely cyclic²¹.

All these properties can be accounted for, if we assume that QR is an operation that computes a selected position and removes from it the quantified element (QP) which cannot be interpreted there; this element should be remerged after a relevant constituent (the nuclear scope) has been created. It follows that QR is actually a *rightward* movement (as proposed, without much discussion, in Fox and Nissenbaum 1999).

More precisely, we proposed that QR:

- a. stores a QP in a dedicated memory buffer of the current phase (Schlenker 2005);
- b. integrates a coindexed variable in the corresponding argument position;
- c. when the top-down computation of the current phase is concluded, the QP function is retrieved from the Q-buffer and takes scope over the structure. The elements retrieved from memory buffers are (typically) not spelled out (Chesi 2004); hence QR is “covert”.

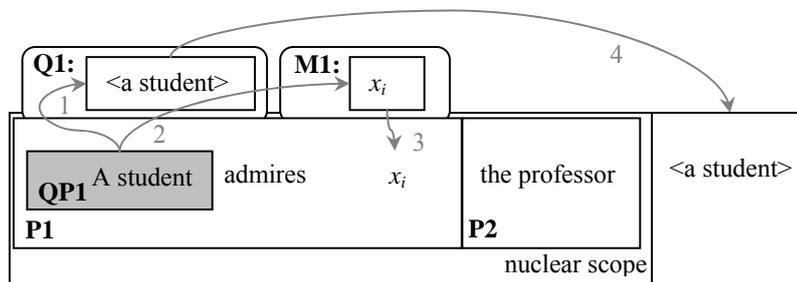
²⁰ But see Begelli and Stowell 1981 and discussion in Bianchi and Chesi 2007.

²¹ But see Cecchetto 2004 and discussion in Bianchi and Chesi 2007.

For example, in the computation of (60), P1 (a verbal phase) is the first phase computed and the subject QP1 (a quantified nominal phase) constitutes a nested phase. This is computed while the nuclear scope (P1+P2) is still incomplete. Then, the nested QP1 is stored in the Q-buffer (Q1) of the containing phase (P1) while the coindexed variable is stored in the already discussed Move Buffer (M1) and behaves as explained in §5.2.

Then we process the sequential phase (P2) which is still part of the nuclear scope and only at the end of it, we can remerge the QP.

(60)



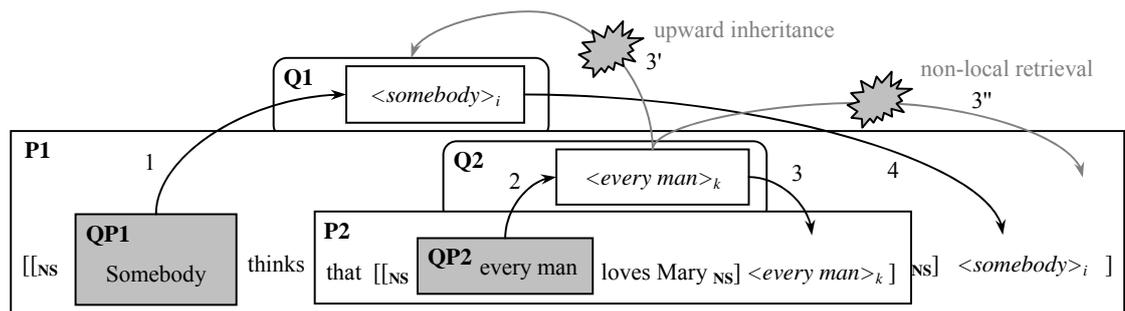
Notice that such device is sufficient to explain clause-boundedness (independently predicting the Right Roof Constraint in (38)): since buffers are phase-local, a QP cannot attach to a superordinate phase, this explains the impossibility of getting inverse scope in sentences like the following one:

(61) *Someone* expected [CP that *every Republican* would win].

$(\exists > \forall; * \forall > \exists)$

More explicitly, the Memory Buffer definition prevents any element from being remerged out of the originating phase (*non-local retrieval) or else copied/discharged in the memory buffer of a superordinate phase (*upward inheritance):

(62)



Then the clause-boundedness of QR follows from the computational sequencing of phases: the matrix subject QP1 is stored in the Q-buffer (Q1) of the containing phase P1 (the variable insertion in the argument position is ignored in the derivation). The embedded subject QP2 is stored in the Q-buffer of P2 (Q2), but it cannot “get into” the Q-buffer of the previously computed P1. As a result, QP2 will only have scope over the embedded P2, whereas QP1 will have scope over both P1 and P2. Thus, the phase boundaries determined by this top-down model, though not corresponding

to complete subtrees, derive the clause-boundedness of QR, subsuming one instance of the otherwise mysterious "right roof constraint".

5.5. *C-command and Pronominal Binding in a PMG*

As for QR, we can use the memory buffer device and the phase-based inheritance mechanism (no upward inheritance, no non-local retrieval) to implement any specific kind of relevant C-command relation. As we saw in (18)-(20), binding provides an important set of tests to guess the attachment point of an extraposed element.

Following Bianchi 2007 (phase-based extension of Schlenker's 2005 proposal) we can implement A-binding using a *Referential Buffer*²² that is a third type of memory buffer used to store the denotation of the referents (i.e. proper names, demonstrative pronouns and definite descriptions) that are processed during the derivation from left to right. The special properties of such buffer are:

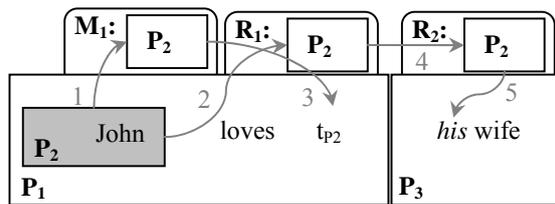
- i. elements within the memory buffer do not need to be discharged in the end for the sentence to be grammatical;

²² We need to introduce some simplification here to keep the discussion focused on a minimal, relevant, set of facts. For this reason I will ignore how these referents are indexed and how the sequence of referents is used to evaluate the truth conditions of the sentence (see Schlenker 2005 and Bianchi 2007 for discussion).

- ii. we admit some degree of freedom in retrieving the referents in the memory buffer (no rigid LIFO structure is assumed²³);

Below the schematic example on how referents are retrieved then evaluated:

(63)



When the referential expression (*John*, P_2) is processed, its referent is stored (step 2) in a phase-local *R(eferential)-buffer* (R_1) which is different from the *M(ove)-buffer* and from the *Q(uantifier)-buffer* since:

- a. it does not need to discharge/remerge the elements it contains at the end of the derivation;
- b. both nested and selected phases inherit the R-buffer of the containing open phase (step 4);

The bound pronoun retrieves the referent from within the R-buffer (step 5) and it is evaluated before the phase gets closed. Schlenker (2005) and Bianchi (2007) propose that whenever in the evaluation sequence (which roughly correspond to our memory buffer) an element which is already

²³ This is because locality does not seem to play a role as strong as for movement (but see Grosz et al. 1985).

present is reintroduced, we get a violation of the *Principle of Non-Redundancy* (i.e. Condition C violation).

5.6. *Extraposition from a Left-Right, Top-Down perspective*

Let us first distinguish EXT of arguments from EXT of adjuncts: in the case of EXT of arguments (“we saw a painting yesterday of John” schematically depicted in (64)), the “heavy” (i.e. containing select features) NP (“a painting of John”) competes with a right hand adjunct (“yesterday”, which is projected rightward because of a select feature in a specific functional position, as explained in §5.3, (57).b) for minimizing nesting (58). The four available solutions would be:

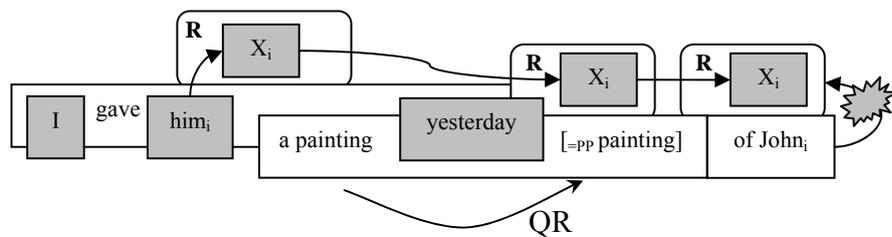
- i. keep the adjunct “in situ” (“*yesterday* we saw a painting of John”);
- ii. shift the whole heavy constituent (“we saw yesterday a *painting of John*”)
- iii. shift the phase projection of the heavy NP (“we saw a painting yesterday *of John*”)
- iv. leave the heavy constituent is situ (“we saw a *painting of John* yesterday”)

We are not interested now in ranking or excluding (for independent reasons) some of the proposed solutions²⁴. What is crucial, according to the discussion in §5.3, is that, in order to get the EXT version in iii., we need some movement trigger to keep Phase Projection local to the verbal head (guaranteeing a sort of Complement-Principle à la Guéron and May, 1984). Following Fox and Nissenbaum (1999) this trigger is the quantificational status of the DP (nominal phase) that, once QR-ed (rightward, as discussed in §5.4, without any stipulation) is remerged in a peripheral position after the nuclear scope has been computed; there it can satisfy the selection requirements of the phase head, first applying (the delayed) Phase-Projection, then Merging the relevant argument. Notice that the selection requirements of the nominal phase head is unsatisfied in situ and this prevents the relevant phase from being closed (the last selection requirement

²⁴ As an anonymous reviewer correctly pointed out, it should be useful to show how this framework treats the optionality of this operation. Unfortunately none of the main proposals currently discussed in the literature are suitable from this perspective: neither the probability distributions discussed in the Gradual Learning Algorithm (Boersma and Hayes 2001; this is because the algorithm needs to access the inner structure of the sentence and this is not trivial for EXT constructions), nor the underspecification of uninterpretable features (Adger 2006; this is because such operations are not feature-driven) seem to play a clear role. In this paper, my intent was just to show how different derivations can be produced in the current framework and which one should be preferred following general principles (e.g. minimize nesting). The intuition that some specific features present in the context, the variability in workload or in memory resources can account for optionality should be framed in more precise terms.

is not “phase projected”), then evaluated as discussed in §5.5. The content of the R-buffer is then available up to the point of the derivation in which the argument “of John” is sequentially merged. This produces a violation of *Non-Redundancy* as soon as the referent “John”, already present, since coreferent with “him”, is inserted in the local R-buffer yielding a condition C effect.

(64)

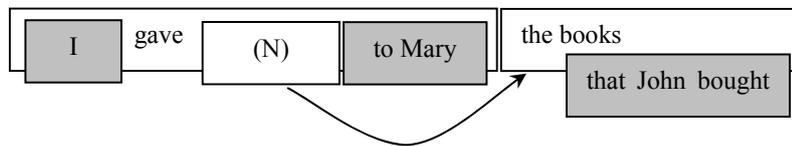


The minimal difference with respect to EXT of adjuncts resides on the fact that the nominal phase “a painting” is complete/interpreted when QR takes place; then when “from John’s collection” is late-attached, after the NP has been remerged after QR, the evaluation sequence (i.e. the memory buffer) is no more available.

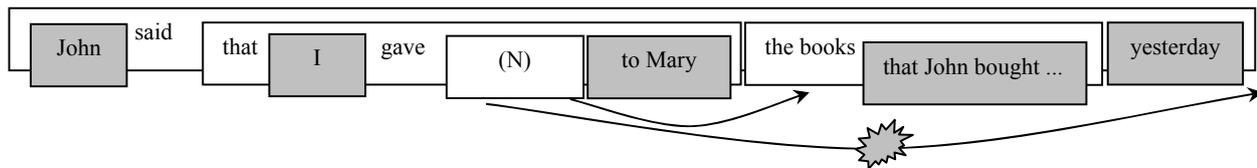
(65)

As briefly suggested in the paragraph before, Minimize Nesting (58) would predict also HNPS. Since selectional requirements have to be satisfied, by definition, in a local configuration with respect to the phase head, we simply predict “clause”-boundedness without any further assumption:

(67) a.



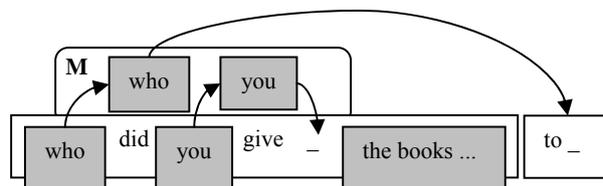
b.



Notice that when the shifted phase is selected, it becomes the “last selected phase”, that is, we can predict movement “from” this constituent ((68).a Vs. movement from the inner one (68).b):

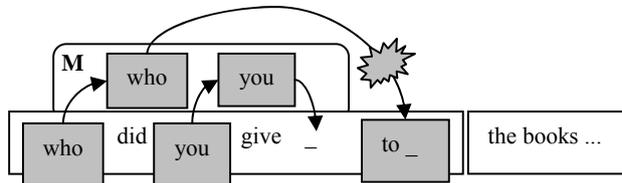
(68) a. Who did you [give *the books written by the venerable Prof. Plum*]

to?



b. *Who did you [give _] to *the books written by the venerable*

Prof. Plum?



Then we can correctly predict that shifted selected constituents behave as last selected (sequential) phases, hence the “ne” extraction from focalized and heavy NPs is straightforwardly captured ((30).b Vs. (31)).

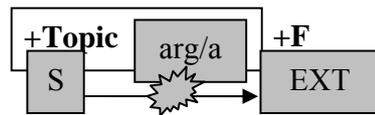
5.8. Remaining issues

In the remaining pages, I do not have space to deal with some interesting puzzles that still remain uncovered by the proposed analysis. However I would like to sketch the tentative solution of two main issues: the ban on Extraposition from the canonical subject position in Italian and from embedded constituent in Dutch.

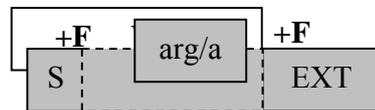
We noticed in §2.2 that EXT from the subject position is not cross-linguistically uniform: English ((10).a), for instance, allows for such operation but not in Italian (12). As we saw in §3.2, Italian allows for HNPS of subject over the direct object and this is possible because of the availability of *pro* in such language (this makes the criterial subject position “recoverable” under (58)). As noticed by Calabrese (1982) and Belletti and Shlonsky (1995) this shifting operation results in a focus on the shifted

subject. Assuming that EXT is also linked to focalization (Huck and Na 1990, among others), we could predict that the status of the Italian canonical subject position (topicalized and not focalized²⁵) makes it incompatible with the status of “focalized constituent” that is necessary in order to allow EXT of an inner constituent²⁶. This is indeed possible under focalized conditions, that seem to be the standard conditions for the subject in languages like English:

(69) a. Italian type languages (subject: +Topic)



b. English type languages (subject: +Focus)



As noticed by Culicover and Rochemont (1990) EXT from topicalized element is not allowed in English too:

²⁵ In this respect, I assume, following Cardinaletti (2003), that the “canonical” subject position is SubjP: a functional projection below CP (FinP) and above TP where the “subject of predication” is licensed. This licensing is a “criterion” in the sense of Rizzi (2006).

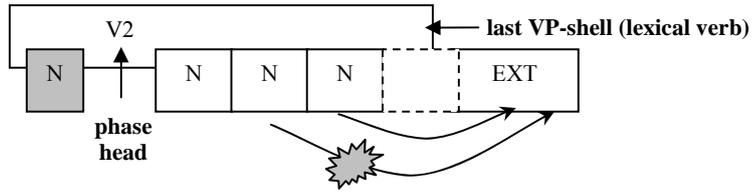
²⁶ Notice that HNPS of subject is not blocked in Italian; this means that the *pro* strategy is potentially available to satisfy a criterial subject/topic position.

- (70) a. John said he would meet a man at the party who was from Philadelphia, and meet *a man* at the party *who was from Philadelphia* he did.
- b. *John said he would meet a man at the party who was from Philadelphia, and meet *a man* at the party he did *who was from Philadelphia*.

Another open issue is the availability of Extraposition from embedded constituents in languages like Dutch (17). I am inclined to believe that such possibility is much more restricted than what has been reported in literature²⁷ and it is essentially possible only from the last selected constituent before the verb final position:

(71)

²⁷ “Müller's counter-examples [...] always involve extraposition along a chain of complements. Nonlocal complement extraposition from adjuncts still appears to be degraded, thus contrasting quite strongly with adjunct extraposition” (Crysmann 2008).



Despite the fact that these data deserve more scrutiny, this solution would be perfectly coherent with the nature of the last selected (sequential) argument.

6.0 Discussion

In this paper I propose that rightward movements, like Extraposition and Heavy NP-Shift, can be successfully accommodated in a Left-Right, Top-Down derivation (where every movement is to the right) without reducing their empirical peculiarities, that are: *clause boundedness*, insensitivity of Extraposition to *adjunct/argument asymmetry* and the *definiteness constraint*. All of these properties can be readily accommodated using the notion of Phase Projection (a phase-based version of theta role assignment) and a new conception of memory-buffered long distance dependencies. Within this radically different perspective we are however still able to characterize as correctly as possible such non feature-driven movements and to capture how they affect scope relations (e.g. selectively bleeding condition C effects, Fox and Nissenbaum 2001, following the derivational top-down implementation of binding principles proposed by Schlenker 2005). With regard to Extraposition I have stressed that the quantificational

status of the host is important for a relevant subset of phenomena (this can hardly be captured in a purely phonological way, e.g. Göbbel 2007), moreover, the directionality of QR does not need to be stipulated (Bianchi and Chesi 2007, Vs. Fox and Nissenbaum 2001).

In this system every long distance dependency (e.g. A'-movement) is regulated by specific triggers (e.g. criterial positions in a functional specification, unavailability of a relevant constituent to be computed, e.g. nuclear scope) using Last In First Out (LIFO) memory buffers locally connected to the notion of phase and regulated by a simple connecting mechanism (only the last selected, sequential phase can inherit the content of a previous memory buffer). In this way we can capture the productive distinction between nested and recursive constituents (the rightmost last selected complement) that allows for successive cyclic movements with no need of teleological features in order to be able to trigger intermediate steps. On the other hand, by relating heaviness to the presence of select features on a nested phase we can predict both shifting and Extraposition as a result of the tendency to reduce nesting (computational complexity). This leads to emarginate “heavy” phases by casting the required phase projections to the end (i.e. to the right) of the phase. Crucial properties that characterize Extraposition and Heavy NP-Shift simply follow from the notion of inheritance memory buffer: they are phase-bound because phase-projection always takes place within the superordinate phase and the memory buffers can neither be discharged upwards, nor on a non-selected phase.

References

- Adger, D. 2007. *A Minimalist Theory of Feature Structure*.
<http://ling.auf.net/lingBuzz/000583>
- Adger, D. 2006. "Combinatorial Variation." In: *Journal of Linguistics* 42, 503-530.
- Akmajian, A. 1975. More evidence for an NP-cycle. *Linguistic Inquiry* 6:114-130.
- Baltin, M. 2006. Extraposition. In *The Blackwell Companion to Syntax*, ed. Blackwell.
- Baltin, M. R. 1981. *Strict Bounding*: MIT Press.
- Belletti, A. 2001. "Inversion as Focalization." In: Hulk, A., Pollock, J. (eds.): *Subject Inversion and the Theory of Universal Grammar*. Oxford: Oxford University Press, 60-90.
- Belletti, A., and Shlonsky, U. 1995. The order of verbal complements: A comparative study. *Natural Language and Linguistic Theory*, 13(3), 489-526.
- Bianchi, v. 2007. *Non-redundancy and backward anaphora*. Ms., University of Siena [presented at GLOW 30, 2007, Tromsø]
- Bianchi, V. 2000. The Raising Analysis of Relative Clauses: A Reply to Borsley. *Linguistic Inquiry* 31:123–140.
- Bianchi, V., and Chesi, C. 2005-2008. *Phases, strong islands, and computational nesting*. Ms., University of Siena [presented at

- GLOW 28, 2005, Geneva] published in STiL - Studies in Linguistics, 2.
- Bianchi, V., and Chesi, C. 2007. *Quantifier Raising in a top-down grammar*, CGG XVII, Gerona
- Boersma, Paul and Bruce Hayes 2001. "Empirical Tests of the Gradual Learning Algorithm." In: *Linguistic Inquiry* 32, 45-86.
- Cardinaletti A. 2004. Toward a Cartography of Subject Positions. in Rizzi L. (Ed.) *The Structure of CP and IP*. OUP.
- Chesi, C. 2007. *An introduction to Phase-based Minimalist Grammars: why move is Top-Down from Left-to-Right*. STiL, 1.
- Chesi, C. 2004. Phases and Cartography in Linguistic Computation: toward a Cognitively Motivated Computational Model of Linguistic Competence. *Unpublished PhD Thesis*.
- Chomsky, N. 2005. *On Phases Manuscript*. MIT.
- Chomsky, N. 2001. *Beyond explanatory adequacy*. MIT Occasional Papers in Linguistics, no. 20. Cambridge, MA
- Chomsky, N. 2000. *Minimalist Inquiries: The Framework*. In *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*, edited by Robert Martin, David Michaels and Juan Uriagereka, 89-155. Cambridge, MA: The MIT Press, 2000.
- Chomsky, N. 1999. *Derivation by Phase*. MIT Occasional Papers in Linguistics, no. 18. Cambridge, MA.
- Chomsky, N. 1995. *The Minimalist Program*, MIT Press, Cambridge, MA.

- Chomsky, N. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Holland: Foris Publications
- Crysmann, B. 2010. *On the Locality of Complement Clause and Relative Clause Extraposition*. In this volume.
- Culicover, P., and Rochemont, M. 1990. Extraposition and the Complement Principle. *Linguistic Inquiry* 21:23-47.
- Culicover, P. and R. Jackendoff 2005. *Simpler Syntax*. MIT Press.
- de Vries, M. 1999. Extraposition of Relative Clauses as Specifying Coordination. Ms. University of Amsterdam (to appear in *proceedings of Console 7 (1998)*).
- de Vries, M. 2006. The Syntax of Appositive Relativization: On Specifying Coordination, False Free Relatives, and Promotion. *Linguistic Inquiry* 37:229-270.
- Fiengo, R. W., and Higginbotham, J. T. 1979. *Opacity in NP*.
- Fox, D., and Nissenbaum, J. 1999. Extraposition and scope: A case for overt QR. *Proceedings of the 18th West Coast Conference on Formal Linguistics* 18:132-144.
- Göbbel, E. 2007. Extraposition as PF movement. In *Proceedings of WECOL 2006*.
- Guéron, J., and May, R. 1984. Extraposition and Logical Form. *Linguistic Inquiry* 15:1–31.
- Holmberg, A. 2000. Am I Unscientific? A Reply To Lappin, Levine, And Johnson. *Natural Language and Linguistic Theory*, 18(4), 837-842.

- Huck, G. J. and Y. Na 1990. *Extrapolation and Focus*. *Language*, 66:51-77
- Husley S. and U. Sauerland 2006. *Sorting out relative clauses*. *NLP* 14:111-137
- Jackendoff, R. 1990. "On Larson's treatment of the double object construction". *Linguistic Inquiry* (21): 427-456.
- Johnson, D., and Lappin, S. 1997. A Critique of the Minimalist Program. *Linguistics and Philosophy*, 20(3), 273-333.
- Kayne, R. S. 1994. *The Antisymmetry of Syntax*: MIT Press.
- Larson, R. K. 1988. *On the double object construction*. *Linguistic Inquiry* (19): 335-392.
- Lasnik, H., and Saito, M. 1992. *Move Alpha: Conditions on Its Application and Output*: MIT Press.
- Lebeaux, D. 1988. *Language acquisition and the form of the grammar*, Doctoral dissertation, University of Massachusetts, Amherst.
- Müller, S., 2004. Continuous or discontinuous constituents? A comparison between syntactic analyses for constituent order and their processing systems. *Research on Language and Computation* 2(2): 209–257.
- Phillips, C. 1996. *Order and Structure*. Ph.D. Thesis. MIT.
- Pinker, S. 1994. *The Language Instinct*: HarperPerennial.
- Richards N. 1999. Dependency formation and directionality of tree construction. *MIT Working Papers in Linguistics* 34, 67-105.
- Riemsdijk, H., and Williams, E. 1981. NP-Structure. *The Linguistic Review* 1:171-217.

- Rizzi, L. 2006. On the Form of Chains: Criterial Positions and ECP Effects. in *WH-Movement Moving on*, S.Cheng & N.Corver eds., 97-134, MIT Press
- Rizzi, L. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of Grammar: Handbook of Generative Syntax*, Kluwer, Dordrecht, 281-337.
- Rizzi, L. 1990. *Relativized Minimality*. MIT Press
- Rochemont, M. S., and Culicover, P. W. 1997. Deriving dependent right adjuncts in English. *Rightward movement*:279–300.
- Rosenbaum, P. S. 1967. *The Grammar of English Complement Constructions*, MIT.
- Ross, J. R. 1967. *Constraints on Variables in Syntax*. MIT PhD Dissertation.
- Schlenker, P. 2005. Non-redundancy: towards a semantic reinterpretation of binding theory. *Natural Language Semantics* 13, 1-92
- Shieber S.M. 1986. *An Introduction to Unification-Based Approaches to Grammar*. Stanford, CA: CSLI Lecture Notes.
- Smits, R. 1989. *Eurogrammar: The Relative and Cleft Constructions of the Germanic and Romance Languages*: Foris.
- Stabler E. 1997. Derivational minimalism. in Retoré, ed. *Logical Aspects of Computational Linguistics*. Springer
- Staub, A., C. Clifton, L. Frazier 2006. Heavy NP shift is the parser's last resort: Evidence from eye movements. *Journal of Memory and Language* 54:389–406

- Taraldsen, K. T. 1981. The theoretical interpretation of a class of marked extractions. *Theory of markedness in generative grammar*:475–516.
- White, J. R. 1998. Syntax-LF mapping and the internal structure of comparatives. *UCL Working Papers in Linguistics* 10
- Williams, E. S. 1974. *Rule Ordering in Syntax*: MIT.
- Zubizarreta, Maria Luisa 1998. *Prosody, Focus, and Word Order*. Cambridge, Mass. The MIT Press.