Do the “right” move

Cristiano Chesi
NeTS - IUSS Center for Neurolinguistics and Theoretical Syntax, Pavia
cristiano.chesi@iusspavia.it

In this paper I argue that rightward dependencies such as Extraposition and Heavy NP-Shift can be accommodated, cross-linguistically, in a top-down (Chesi 2004-12) minimalist framework that results in a left-right phrase structure generation (cf. Phillips 1996, Phillips and Lewis, Kempson et al. and Kiaer in this volume). All the special properties that make Extraposition and Heavy NP-Shift peculiar compared to standard “leftward” movement will be considered, focussing on clause-boundedness, adjunct/ argument asymmetries with respect to dependency directionality, and the definiteness constraint.

1. Introduction

A Minimalist Grammar is top-down (Chesi 2004-12) and, as a consequence, left-right (cf. Phillips 1996, and contributions in this volume by Phillips and Lewis, Kempson et al. and Kiaer) if structure building procedures operate by expanding constituents from the root of the syntactic tree, rather than creating them by merging lexical items starting from inner constituents (Chomsky 1995). Reformulating the anti-symmetric intuition (Kayne 1994), I suggest that linear ordering (“left-right”, in the sense of the temporal order in which roughly we parse and produce sentences) derives from the hierarchical one (top-down). Surprisingly, at least from the mainstream perspective, in such a framework, all long-distance dependencies are “rightward” dependencies given that the dependency trigger must be found first (e.g. a *wh-* filler, Fodor 1978), then the dependent (possibly non-local) constituent will be identified (e.g. the selecting verbal head). Using memory buffers to store and retrieve constituents in a principled order/way we can correctly characterize (successive cyclic) movement and islandhood (Chesi 2004-07), parasitic gap constructions (Bianchi & Chesi 2008), quantifier raising (Bianchi & Chesi 2010) and A-binding (Bianchi 2010).
Here I argue that 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) ((1a) from Baltin 2006, (1b-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*

© Copyright 2013 Cristiano Chesi
STIL – Studies in Linguistics. 6:107-138
Special Issue on “Directionality of Phrase Structure Building” ISSN: 2281-3128

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

(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*.

Here I want to 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" (in the sense of Phillips 1996) derivation and we redefine structure-building operations top-down within a phase-based[1], head-driven[2] derivation.

In the first part of this paper (§2), I will compare 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), I will discuss data related to Heavy NP-Shift that present properties consistent with Extraposition (i.e. Clause-Boundedness).

In the third part (§4), I will analyse some of the proposed standard solutions, highlighting the main problems of the given analyses.

The last section (§5) is dedicated to the proposed solution: the memory-based 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, and it provides 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)

Extraposition (henceforth EXT) is a limited option that some language makes available to marginalize, at the end of the relevant clause, certain constituents (mainly PPs and rRCs). The well know distributional properties that characterize this phenomenon are summarized in this chapter, focussing on 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 _]?

[1] As in Chomsky 2008, here a phase will be a minimal computational domain bounding structure building operations, though this notion will be revisited according to the perspective shift.
[2] As in HPSG (Pollard and Sag 1994) I believe that the lexical verb within CP and the main noun in any DP/PP plays a crucial role that must be acknowledged

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

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*
pro have seen a painting yesterday by G.
a'. Di chi pro hai visto [un quadro _] ieri?
Of whom *pro* have seen a painting yesterday
b. pro Abbiamo visto [un quadro _] ieri con dei girasoli
pro have seen a painting yesterday with sunflowers
b'. *Preso da dove / * Da dove pro abbiamo visto [un quadro _]?
taken from where / from where *pro* have seen a painting

(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?

Also restrictive relative clauses undergo Extraposition (English (7), De Vries 2006, Italian (8), and Dutch (9), De Vries 2002)[3]:

(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 the his bag
(9) Ik heb [de man _] gezien *die zijn tas verloor*.
I have [the man _] seen who his bag lost.

Briefly:

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. A host can be both an argument (10a,b,c) and an adjunct (10d):

(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*

[3] The extraposition of appositive relative clauses is a more controversial issue in the linguistic literature (Vergnaud (1974:181), Emonds (1979:234), Citko (2008), de Vries (2006) and others). In this paper, I will restrict the discussion to restrictive relative clauses only.

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

Here some restrictions apply if the constituent is moved: Fronted PPs adjuncts (11a) do not permit EXT to take place (even though prepositional stranding improves EXT from these constituents, (11a')), while moved wh- arguments (e.g. (11b), Baltin 2006) 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 t_i *who was unhappy about the visit*?

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 it is not possible to extrapose a PP/rRC from the pre-verbal subject position (independently of whether the verb is transitive (12a), unergative (12b), unaccusative (12c), or in a passive voice (12d)):

(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 *of 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 (though slightly deviant for some speaker) are 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*
(pro$_i$) 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 a direct object (14a) or an indirect object (15a), but it yields grammatical results only when the interrupting constituent is "light" (prepositional modifiers seem to block EXT as shown in (14b) and (15b)):

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

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

(15) a. Gianni ha mangiato un panino [con un'amica _] ieri
di suo fratello / che era ammalata
G. has eaten a sandwich with a friend yesterday
of his brother / who was sick
b. *Gianni ha mangiato un panino [con un'amica _] in fretta
di suo fratello / che era ammalata
G. has eaten a sandwich with a friend in a hurry
of his brother / who was sick

Notice that not only PP adjuncts (14b, 15b), but also the presence of an extra argument (16b) results in a degradation of EXT from direct object in Italian:

(16) a. Gianni ha spedito [una lettera _] ieri *senza francobollo*
G. sent a letter yesterday *without stamp*
b. ??Gianni ha spedito [una lettera _] a Maria *senza francobollo*
G. sent a letter to M. *without stamp*

Last consideration is related to "constituent-boundedness": it seems that embedded constituents cannot be hosts of EXT in English (17a) / Italian (17b) while they can in Dutch (17c) (De Vries 1999)[4].

(17) a. *[A review of [a book _]] appeared *by three authors*
b. *È apparsa [una recensione [di un libro _]] ieri *di tre autori*
is appeared a review of a book yesterday *by three authors*
c. 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. Italian marginally allows the last argument/adjunct to be host for EXT, and no "heavy" (prepositional) modifiers/arguments should intervene between the extraposed constituent and the related host;
iii. EXT from embedded constituents is generally 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 we get a principle C effect in the following examples[5]:

(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.

[4] But see English counterexamples provided by Strunk & Snider (2013) and the discussion in §5.8.
[5] Following Larson (1988), the pronominal object C-commands the second object.

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

Looking at condition C bleeding under wh- movement and its sensitivity to the complements (19a) vs. adjuncts (19b) distinction (Van Riemsdijk and Williams 1981, Lebeaux 1988) so does EXT ((19c') vs. (19d'), Taraldsen 1981, Fox & Nissenbaum 1999)) we could conclude that the object "reconstructs" under the host while the adjunct does not:

(19) a. ??/*[Which book *about John$_i$'s library*] did he$_i$ read _?
b. [Which book *from John$_i$'s library*] did he$_i$ read _?
c. ??/* I gave him$_i$ [a picture *of John$_i$'s mother*] yesterday.
c' ??/* I gave him$_i$ [a picture _] yesterday *of John$_i$'s mother*.
d. ??/*I gave him$_i$ [a picture *from John$_i$'s collection*] yesterday.
d'. I gave him$_i$ [a picture _] yesterday *from John$_i$'s collection*.

As for the attachment point, it has been proposed that the constituents extraposed from subjects attach to IP (availability of stranding if the VP is elided (20a,b)), while the constituents extraposed from objects attach to VP (no stranding option if the VP is elided (20c-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, (21a) vs. (21b):

(20) 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 did [$_{VP}$ _ from New York].

(21) 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):

(22) *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

Following Ross' (1967) argument, reordering could also be considered as an instance of rightward movement; the direct object crossing over the indirect one or an adjunct is an example of this construction (Staub et al. 2006):

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

(23) Lucy ate _ with a fork [the extremely delicious, bright green broccoli]

This "rightward movement analysis" has been seriously challenged by Larson (1988) and 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 relate Heavy NP-Shift (henceforth HNPS) to EXT, suggesting that the rightward movement analysis might be on the right track in some specific case.

3.1. What

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

(24) 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]

Even though such operation seems to be optional, a fairly evident bias (Pinker 1994:131) is attested for preferring certain shifted versions, (25a), of "heavy"[6] constituents among many logically possible alternatives, (25b):

(25) a. In my laboratory we use it as an easily studied instance of mental grammar, allowing us [to document [in great detail] [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, (26a'), while, in Italian, "shifting" seems to be freely applicable in the very same context, (26b') [7]:

[6] A "heavy" NP is usually considered a NP containing a sentential/relative clause (Ross 1967). This (insufficient) definition will be better discussed in §5. In (25) the heavy constituent is in Italic.
[7] However, different orders have different informational structure implications (Belletti and Shlonsky 1995, Zubizarreta 1998). See discussion in §3.2.

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

(26) a. I gave the books to Bill
a'. *I gave to Bill the books
b. *pro* ho dato i libri a Bill
b'. *pro* ho dato a Bill i libri

In brief:

i. Shifting is an optional operation;
ii. Shifting only targets "heavy" elements in English, but it can also target light constituents in Italian;
iii. Shifting is the preferred option for "heavy" constituents.

3.2. Where and why

Evidences that the "shifting" operation is not (always) a uniform "rightward" movement come from Italian (Belletti and Shlonsky 1995): The clitic *ne* (literally "some of them") can be extracted only from the base object position (27b), but not, for example, from a post-verbal subject of unergatives (28b), Belletti and Rizzi 1981, Burzio 1986). *Ne* cliticization with light NPs is only possible if they are not shifted, (29a) vs. (29b). However, with heavy (30a) and focussed (30b) NPs that appear in a "shifted" position, *ne*-cliticization is also possible. This has been interpreted as the signature of the fact that the "shifted" object, e.g. (30), is, indeed, in his base position:

(27) a. Ho letto molti libri.
(I) read many books.
b. *Ne* ho letti molti.
(I) of-them read many.

(28) a. Hanno lavorato molti operai.
have worked many workers.
b. *?*Ne* hanno lavorato molti.
of-them have worked many.

(29) 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.

(30) 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

Because of this paradigm, Belletti and Shlonsky (1995) proposed 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

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

of the availability of subject/object *pro* in this language and it results in a focalization of the shifted constituent (this reminds of the obligatory focalization of post-verbal subject, Calabrese 1982, (31a)). Assuming that the focalized position to the "right edge" of the VP is unique, this solution can capture the impossibility of "light"-shifting both of the subject and the object at the same time[8] (31b):

(31) 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 the shifting of light NPs but not for the (adverbially) focalized object in (30b) for which Belletti and Shlonsky (1995) propose that the PP must scramble before the NP, which, in 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 1991, (32a)) as predicted in the PP scrambling hypothesis; in this case, *ne* extraction from the direct object is allowed as expected, (32b):

(32) 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 why the shifted heavy-NP behaves as if it were in an A'-position (instead of an A-position according to the *ne* cliticization test) when it licenses parasitic gaps:

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

The contrast between (33b) (heavy NP) and (33c) (light NP shifting) shows that heaviness is in fact the crucial property that triggers, in a cross-linguistic uniform way, 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 object constructions:

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

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

extractions from the first object (34a,b) seem to be much more difficult than from the second one ((34c,d), Culicover and Rochemont 1997:15):

(34) a. *Bill [gave _] the book yesterday *anyone who wanted it*
b. **Who* did Bill [give t_i] the book yesterday
c. Bill [gave [John] _] yesterday the book that he was looking for
d. What$_i$ did Bill [gave [John] t_i] 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 when the object is not in its canonical base position, i.e. next to the verb);
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 (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 (that is straightforwardly suitable also for HNPS) has been a *rewriting transformation* (Rosenbaum 1965:11), adapted by Ross (1967) as follows:

(35) 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 (10a) is (10b), 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 (36b):

(36) 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]]

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

For this reason Ross proposes an ad-hoc constraint, that excludes (36b) forcing clause-boundedness:

(37) *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 is movement much more constrained to the right than to the left?
b. What does trigger Extraposition?
c. How can we account for the clause-boundedness 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:

(38) *Complement Principle* (Guéron and May 1984)
In a sequence of categories $\alpha_i, \beta_i^1 \ldots \beta_i^n$ in a structure $\Sigma, \beta_i^1 \ldots \beta_i^n$ 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[9], but it contains at least one important flaw: the complement principle does not make any distinction between arguments and adjuncts (in both cases, no maximal projection boundaries, i.e. CP, intervene between the noun/verb and the DP/PP, then, according to (38), in both cases, arguments and adjuncts are governed), but this seems to be a very productive distinction as shown in (19).

[9] Walker (2013), running an experiment to test the speaker performance with respect to the Definiteness Constraint, verifies that this seems to be only a weak constraint, i.e. it can be violated in some significant way. I would prefer sticking on the strongest assumption and accept the performance issues raised in her study as the result of the availability, for some speaker, of a null existential Q operator on the otherwise simply definite DP (e.g. "(∃) the girl appeared..."). Notice that the verb of appearance facilitate this interpretation, thus also the contrast with respect to other verb classes seems less mysterious.

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

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 movement[10]). 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 (39a) is not a constituent, so it cannot move, while "two/a/those book(s)" is a constituent, so it can move stranding the RC behind as expected):

(39) 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 (37), follows from theta-role requirements and/or LCA:

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

Nevertheless, this account presents some problems:

a. Extraposed constituents seem not to be C-commanded by their host (e.g. (18));
b. Examples like (10c) remain unexplained (since [P NP] should not be considered a constituent using Kayne's analysis);
c. The stranding analysis would not easily predict the nesting dependencies reported in (21) (because of the order of complements, i.e. $[_{V}$ $[_{DP1}$ $[_{EXT1}]]$ $[_{DP2}$ $[_{EXT2}]]]$, unless we assume some sort of scrambling, standard movement operations would affect first the inner constituent, i.e. $[[_{DP2}$ $]$ $[_{V}$ $[_{DP1}$ $[_{EXT1}]]$ $[t_2$ $[_{EXT2}]]]]$, then the higher one, leading to a cross-serial dependency, i.e. .$[[_{DP1}$ $]$ $[[_{DP2}$ $]$ $[_{V}$ $[t_1$ $[_{EXT1}]]$ $[t_2$ $[_{EXT2}]]]]]$).

4.4. Analysis 4: "Mixed" account

The last proposal I would like to 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, can be the result of QR plus late merge of an adjunct. This is the schematic derivation Fox and Nissenbaum propose:

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

[10] If *A* asymmetrically C-commands *B*, then terminals dominated by *B* can not precede terminals dominated by *A* (Kayne 1994:33): this is why rightward adjunction could not be linearized.

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

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)

As for EXT of complements, if we assume that thematic requirements need to be satisfied everywhere (both at Deep Structure and at LF as predicted by the Complement Principle), the complements need to be merged before, then QRed afterward; so, no late merge of complements is available.

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 discussed in (19).

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

(42) 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 is QR assumed to be rightward?

5. The proposal: changing the orientation of the derivation

In order to account for the set of characterizing properties of EXT and HNPS, I propose to adopt a radically different minimalist derivational perspective: constituents must be built not from bottom to top and, consequently, from right to left (at least in right branching languages) as generally assumed, but strictly from left to right, expanding lexical "expectations" in a principled, top-down, way.

5.1. Deriving phrase Top-down

A top-down (Chesi 2004-12) derivation is a generative procedure that assembles lexical items in phrase structures, using structure building operations inspired by the Minimalist Program (Chomsky 1995-2008, Stabler 1997). The Phase-based Minimalist Grammar (PMG, Chesi 2007) that implements such (non-conventional) derivational direction requires a fully explicit lexicon, the formalization of the necessary structure building operations, and a clear specification of their domain of application.

As for the lexicon, I will adopt a simplified version of Stabler's (1997) formalism: a lexical item is a ordered feature structure[11] composed by functional features (prefixed

[11] Here to express ordered sets I will use squared brackets: e. g. [A, B, C] or simply [A B C].

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

by the "+" diacritic, expressing functional properties/positions, e.g. +D(eterminer)), selection features (characters prefixed by the "=" diacritic, expressing thematic requirements in terms of necessary additional constituents that must be present in the structure in order for the expression to be grammatical), categorial features (unprefixed characters expressing N(ominal), V(erbal) and A(djectival/dverbial) heads) and, in the end, phonetic/semantic features (that I do not discuss here and I indicate using placeholders). This is an example of lexical item (as in "John *runs*")[12]:

(43)		[+T_{pres} +$Agr_{3\text{-}sing}$	=[+D N]	V	*runs*]
	features:	*functional*	*selectional*	*categorial*	*phon/sem*

Structure building operations will target lexical items and phrases and access their feature structure in a constrained way. To understand such constraints we distinguish operations that establish local relations from the operations creating non-local ones. Looking at local relations, if we assume that:

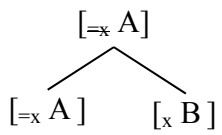
i. Merge is responsible for local phrase structure building,
ii. it is binary,
iii. it is successful between A and B (with A and B either lexical or phrasal) if and only if A selects B or B selects A[13],
iv. the item that selects also projects,

a natural Top-Down constraint follows:

(44) The selecting item is computed before the selected one.

This is so, because of the nature of the hierarchical structure: if *A* selects *B*, as in (45), *A* will be computed before *B* if the tree is explored Top-Down. I propose, then, that computing *A* will generate the expectation of *B*, because of the select feature on *A*.

(45)



Pushing forward this idea, the Top-Down Merge reduces to local lexical selection satisfaction (and, eventually, unification, Shieber 1985, Chesi 2012:51).

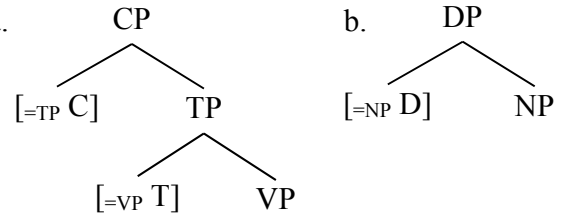
[12] Let us take this as a simplified example. Obviously we could decompose such lexical items in its morphological subparts ([=[+D N] V *run*] [+T_{pres} ∅] [+$Agr_{3\text{-}sing}$ *s*]) but this would require extra machinery to be discussed. Similar considerations hold for feature structures: in this paper, I assume a pure privative system, i.e. any feature is atomic and its presence/absence within a lexical item is completely arbitrary (Adger 2007). In this sense, every lexical item is an "exception" (Chomsky 1995:235). In many cases, using full feature structures could make many tasks much more elegant (Pollard and Sag 1994, Chesi 2004).

[13] Stabler (1997), for instance, uses this restriction to limit Chomsky's (1995) Merge. Similarly, Collins (2002) proposes that Merge itself is triggered by saturation considerations. The recent discussion on 'labeling' of tree nodes seems to me to go exactly in this direction (Cecchetto & Donati 2010): the 'label' (i.e. the interpretable result of Merge, accessible to other Merge operations) is the 'probe', i.e. the merged item that selects the other

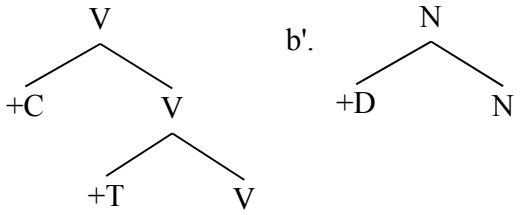
BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

As for non-local relations, some important constraint might follow from an answer to the question "How many features might be expected (hence selected) at once?". My tentative answer diverges from the minimalist one (that is, just one categorial feature, e.g. [$_{=N}$ D]), and it is crucially related to the minimalist notion of Phase. From a Top-Down perspective, the features that can be selected or expanded correspond to the minimal computational domain within which a non-local dependency should be evaluated. By assumption, this domain matches with the extended projection (Grimshaw 1991) of a lexical head (N, V or A) and its selected arguments. In this sense, the generally accepted structures in (46)a and (46)b will be replaced by (46)a' and (46)b' (i.e. extended projections of V and N, respectively):

(46) a. b.



a'. b'.



By using extended projection, selecting a DP (i.e. =DP) equals to say that a determined (+D) nominal (N) extended projection is expected (i.e. (46)b') while selecting a tensed declarative sentence would correspond to the expectation of the structure in (46)a'. In fact, because of the necessity to express the subject of predication, SVO languages should also include, a +S (criterial, in the sense of Rizzi 1997) feature in the verbal extended projection corresponding to a declarative sentence expectation, that is: [+C +S +T V].

Processing a "phase" means lexicalizing any expected feature using a compatible item taken from the lexicon. We predict that both phonologically null items ([+C Ø] to indicate a declarative force, and [+S_{nom} Ø] working as an implicit case marker) and features clusters ([+C_{wh} +D N *what*] or [+S +D N]) are present in our lexical knowledge. So it can happen that either lexicalized (e.g. [+C_{wh} +D N *what*]) or unlexicalized (e.g. [+S +D N]) feature clusters are used to expand just one expected features. By assumption, the "unexpected" features that are introduced must be "stored in memory" until they will be properly selected.

This is what triggers movement, i.e. a non-local dependency of the filler-gap, filler-first kind (Fodor 1978), both in the case of criterial wh- items and in the case of the pre-verbal subject; in an informal way this is how the derivation unrolls in the simple case of a finite declarative sentence, in a SVO language like English:

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

(47) a. a declarative sentence is expected: [+C +S +T V];
b. +C is lexicalized by a null [+C Ø];
c. [[+S_{nom} Ø] +D N] lexicalizes the +S expectation;
d. [+D N *John*] lexicalizes both +D and N, but since these features where unexpected, [+D N (*John*)] is stored in a memory buffer[14];
e. [+T V =DP *sings*] lexicalizes +T and V at once;
f. =DP selection of [=DP *sings*] creates the landing site for an item compatible with [+D N] expectation;
g. [+D N (*John*)] is discharged from the memory buffer;
h. The derivation ends since no more features must be processed.

In sum, the Phase idea, and the Merge and Move operations are reformulated according to the logical constraints here discussed and produce a left-right derivation as expected both in parsing and generation[15]:

(48) *Phase*: it is the minimal computational domains including an extended projection of an expected lexical head and its selection requirements;
(49) *Merge*: it lexicalizes/expands the expected features (by unification, Shieber 1986) introducing in the left-most unlexicalized position, a compatible item taken from the lexicon;
(50) *Move*: it creates Long Distance (Filler-Gap, Filler-First) Dependencies by means of a *Memory Stack*; constituents are "moved" in the memory buffer when they introduce unselected (i.e. unexpected) features and retrieved/re-merged in the structure as soon as possible when properly selected (Moved items preempts the insertion, i.e. Merge, of new items from the lexicon).

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

There are some logical possibilities on the order of application of the three operations just discussed that crucially result in different derivations/structures[16]. I assume that in the default case, we first apply Phase Projection, then Merge and in the end Move. Below, a sample derivation that shows how Phase Projection (PhP) and Merge operate in the case the Move buffer be empty[17] and a base verbal phase is expected:

[14] The null hypothesis is that features already used will not be re-merged twice; e.g. the first copy of a spelt out item is the one consuming its phonological features, this is why (*John*) is indicated under brackets. It is plausible to expect cross-linguistic parameterization on this, but I will not discuss these aspects here.
[15] See Chesi 2007, 2012 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 (and its revision in Collins and Stabler 2011) proposal (Minimalist Grammars, MGs).
[16] Interleaving in a way or another structure building operations results in grammars with different generative power (see Chesi 2007 for discussion).
[17] 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.

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

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

1. *PhP*(=V) → [$_{V}$] (this is a default phase projection that simply instantiates a verbal phase)

2. *Merge* ([$_{V}$],[$_{=DP\ =DP\ =PP\ V}$ gives]) → [$_{=DP\ =DP\ =PP\ \underline{V}}$ gives]

3. *PhP*([$_{=DP\ =DP\ =PP\ \underline{V}}$ gives]) → [$_{=DP\ =PP\ \underline{V}}$ gives [$_{+D\ N}$]]

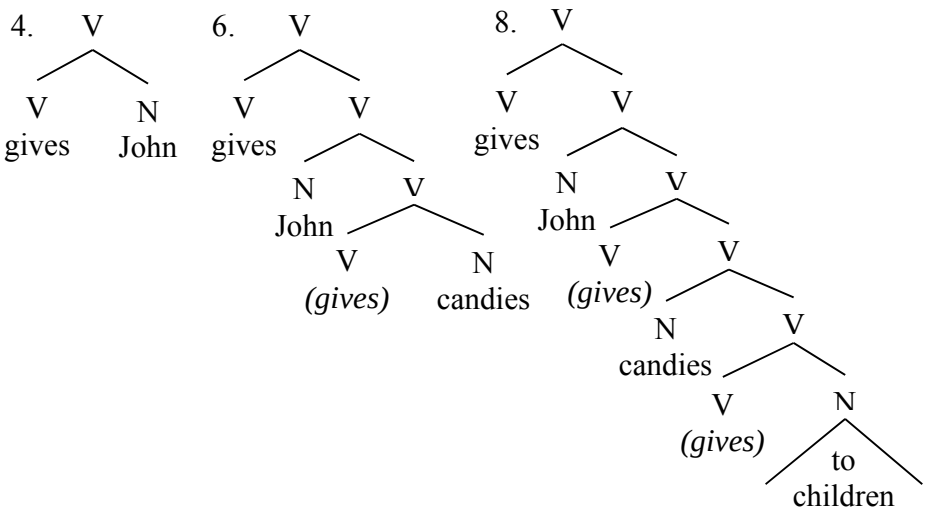
4. *Merge* ([$_{=DP\ =PP\ \underline{V}}$ gives [$_{+D\ N}$]], [$_{+D\ N}$ John]) → [$_{=DP\ =PP\ \underline{V}}$ gives [$_{\underline{+D\ N}}$ John]]

5. *PhP*([$_{=DP\ =PP\ \underline{V}}$ gives [$_{\underline{+D\ N}}$ John]]) → [$_{\underline{V}}$ gives [$_{\underline{+D\ N}}$ John] [$_{=PP\ (gives)}$ [$_{+D\ N}$]]]

6. *Merge*([$_{\underline{V}}$ gives [$_{\underline{+D\ N}}$ John] [$_{=PP\ (gives)}$ [$_{+D\ N}$]]], [$_{+D\ N}$ candies]) → [$_{\underline{V}}$ gives [$_{\underline{+D\ N}}$ John] [$_{=PP\ (gives)}$ [$_{\underline{+D\ N}}$ candies]]]

7. *PhP*([$_{\underline{V}}$ gives [$_{\underline{+D\ N}}$ John] [$_{=PP\ (gives)}$ [$_{\underline{+D\ N}}$ candies]]]) → [$_{\underline{V}}$ gives [$_{\underline{+D\ N}}$ John] [$_{=PP\ (gives)}$ [$_{\underline{+D\ N}}$ candies] [$_{(gives)}$ [$_{+K\ +D\ N}$]]]]

8. *Merge* ([$_{\underline{V}}$ gives [$_{\underline{+D\ N}}$ John] [$_{(gives)}$ [$_{\underline{+D\ N}}$ candies] [$_{(gives)}$ [$_{+K\ +D\ N}$]]]], [$_{+K\ +D\ N}$ to children]) → [$_{\underline{V}}$ gives [$_{\underline{+D\ N}}$ John] [$_{(gives)}$ [$_{\underline{+D\ N}}$ candies] [$_{(gives)}$ [$_{\underline{+K\ +D\ N}}$to children]]]]

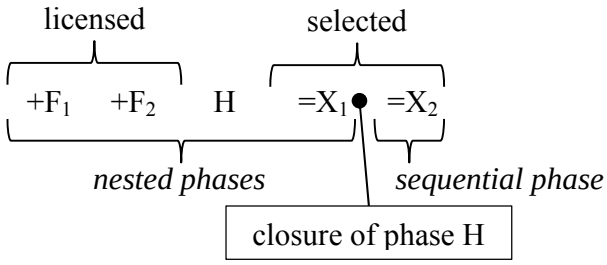


In this derivation, *Phase Projection* creates an empty (i.e. unlexicalized) constituent to the right edge of the structure depending on the first select feature of the processed lexical item; *Merge* fills this position either with the most prominent element in the memory buffer (if any, and this is not the case in the example) or with an element from the lexicon (as in (51) where the memory buffer is empty).

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

In the case the memory buffer is not empty, storage and retrieval is constrained by *phase*. The reason for using phases follows from cognitive and computational considerations, that is, we want to keep "working memory" as small as possible (coherently with processing evidence, Baddeley et al. 2009). In this sense, I will retain the minimalist intuition that, at some point, the phase is "shipped-out", and no further operations can tamper the set of dominance/precedence relations created up to this point. In this vein, we decided (Chesi 2004, Bianchi and Chesi 2008) to fix this "memory limit" at the last operation triggered by the phase head, that is, the Phase Projection of the last selected argument. This produces important computational advantages (recasting the *true* vs. *tail* recursion distinction, Abelson and Sussman 1996, in phasal terms, Chesi 2004-12) 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 2008), that once projected close the previous phase (since this allows for unlimited recursion, i.e. $=X_2$ in (52)), and the *nested phases* (Bianchi and Chesi 2008), i.e. unselected phases resulting from expansion of functional features or phases that are not the last selected one (that, in case of recursion, would create centre-embedding, Bever 1970):

(52)



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, as a consequence, the movement operation:

(53) *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.

(54) *Success condition*
Memory buffers must be empty at the end of the computation.

A simple example showing how movement operates is subject movement in SVO languages. According to what I said, we expect the basic word order of transitive verbal ph(r)ases to be VSO, as in (51). In order to derive the relevant SVO language we would have two options:

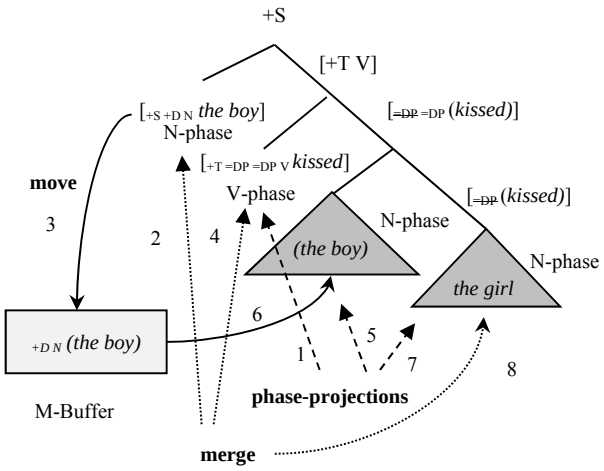
i. the verbal head could be spelled out at the lower VP-shell;
ii. the subject is generated first in preverbal position, then moved.

While the first solution could be cross-linguistically promising, here I propose that SVO order is obtained by exploiting the second strategy. This is a graphical

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

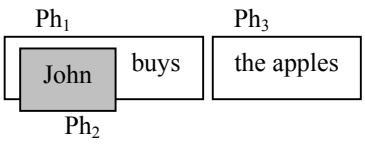
representation of the sketched derivation of a simple SVO sentence in English (the numbers labelling the arrows indicate the sequential order in which the operations apply).

(55)

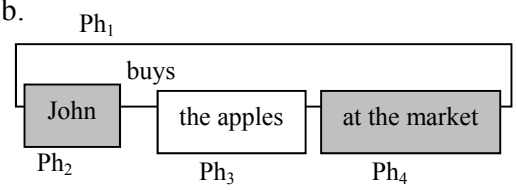


In the following pages, I will use a box-notation[18] to visualize derivations/structures as the one in (55). Nested phases are expressed by forefront boxes; when phases are unselected (e.g. adjuncts, like "at the market" in (56b) or the preverbal "John" in (56b)) they are indicated with grey boxes; in case of selected arguments (white boxes) I expect only the last selected argument that terminates the phase to qualify as sequential (e.g. "the apples" in (56a)), then in other cases, the arguments should be nested (as in the case of "the apples" in (56b), since Ph_3 is followed by a PP modifier, Ph_4, related to the matrix phase) though we might expect some degree of cross-linguistic variation in this respect. In the sequential case the boxes are white and aligned as in (56a):

(56) a.



b.



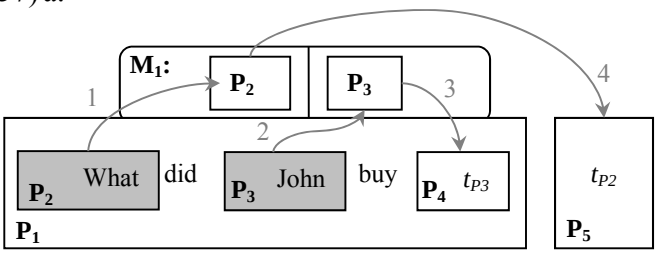
These assumptions are sufficient to capture (successive cyclic, (57b)) movement, (57a), predicting the correct locality constraints[19]:

[18] The "box-notation" is due to Valentina Bianchi. This notation allows us to keep track of the derivation in a compact and meaningful way. In a nutshell: squared boxes are phases, rounded boxes are memory buffers (the elements are ordered from left to right, i.e. the rightmost one is the first one to be remerged), indexes mark the univocal phase-memory buffer dependency.

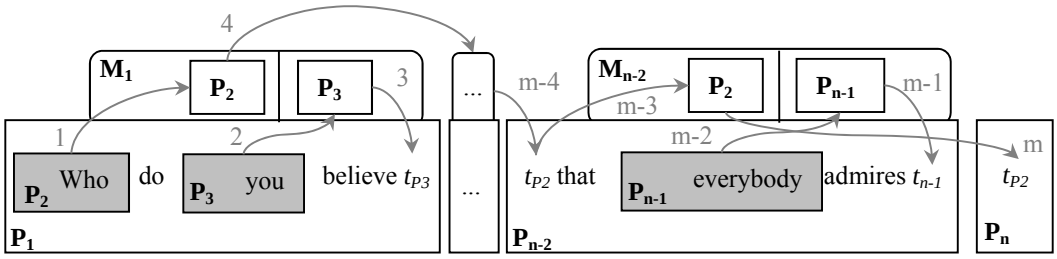
[19] In the original proposal, Chesi 2004, the memory buffer was structured in slots in order to capture *Relativized Minimality* effects (Rizzi 1990). Under the cue-base retrieval assumption (Van Dyke & Mc Elree 2006) this hypothesis seems unnecessary.

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

(57) a.



b.



In (57a), the steps of the derivations are the following ones:

1. A default verbal phase is projected (P_1):
   PhP(*Wh- question*) = [$_{+wh\ +T\ +S\ V}$][20]
2. Since the verbal phase is interrogative, this functional feature has to be explicitly marked; in English this is done by merging the relevant *wh*-element within the specific "criterial" position. This is how [$_{+wh\ +D\ N}$ what], phase P_2 (computed as a nested phase), is introduced in the derivation:
   Merge([$_{+wh\ +T\ +S\ V}$], [$_{+wh\ +D\ N}$ *what*]) = [$_{+wh}$[$_{\sout{+wh}\ +D\ N}$ *what*] $_{+T\ +S\ V}$]
3. Since [$_{+D\ N}$] are unselected [$_{+D\ N}$ (what)] is inserted (step 1) in the memory buffer (M_1) of the matrix V-phase (P_1):
   Move([$_{\sout{+wh}\ +D\ N}$ ~~*what*~~]) = M1< [$_{(+wh)\ +D\ N}$ *(what)*] >
4. *did* is compatible (unificable) with a tense functional specification of the matrix V-phase, then licensed in this position:
   Merge([$_{+wh}$[$_{+D\ N}$ *what*] $_{+T\ +S\ V}$], [$_{+T}$ *did*]) =
   [$_{+wh}$[$_{+D\ N}$ what] $_{+T}$[$_{\sout{+T}}$ did] $_{+S\ V}$]
5. [$_{+D\ N}$ *John*] (phase P_3, again computed as a nested phase) is introduced to satisfy a subject-criterial (in the sense of Rizzi 2006) requirement (functional specification of P_1, i.e. +S, triggering [$_{+S}$[$_{\sout{+S}}$ ∅] $_{+D\ N}$] phase expansion) and moved in the memory buffer since [$_{+D\ N}$] are unselected (step 2):

   Merge([$_{+wh}$[$_{+D\ N}$ *what*] $_{+T}$[*did*] $_{+S\ V}$], [$_{+S}$[∅] [$_{+D\ N}$ *John*]]) =
   [$_{+wh}$[$_{+D\ N}$ *what*] $_{+T}$[*did*] $_{+S}$[$_{\sout{+S}}$[∅] $_{+D\ N}$ *John*] $_{V}$]
   Move($_{+S}$[$_{\sout{+S}}$ [∅] [$_{+D\ N}$ ~~*John*~~]]) = *M1*< [$_{(+wh)\ +D\ N}$ *(what)*], [$_{(+S)\ +D\ N}$ *(John)*] >

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

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

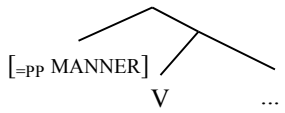
6. [$_{=[+D\ N]\ =[+D\ N]\ V}$ *buy*] is merged as the head of the matrix V-phase (P_1). Since it has two selection requirements to be satisfied (an agent and a patient, both N-phases), these select features will project two phases, P_4 and P_5, in sequence:
   Merge([$_{+wh}$[$_{+D\ N}$ *what*] $_{+T}$[*did*] $_{+S}$[[∅] [$_{+D\ N}$ *John*]] $_{V}$], [$_{=[+D\ N]\ =[+D\ N]\ V}$ *buy*]) =
   [$_{+wh}$[$_{+D\ N}$ *what*] $_{+T}$[*did*] $_{+S}$[[∅] [$_{+D\ N}$ *John*]] $_{V}$[$_{=[+D\ N]\ =[+D\ N]\ V}$ *buy*]]
   PhP(=[+D N]**)** = [$_{+wh}$[$_{+D\ N}$ *what*] $_{+T}$[*did*] $_{+S}$[$_{+D\ N}$ *John*] $_{V}$[$_{=[+D\ N]}$ buy [$_{+D\ N}$]]
7. P_4 is a nested phase and will be unified by (re-)merging P_3 (the first accessible element in M1);
   Merge(... $_{V}$[$_{=[+D\ N]}$ *buy* [$_{+D\ N}$]], *M1*[$_{(+S)\ +D\ N}$ (*John*)]) =
   $_{V}$[$_{=[+D\ N]}$ buy [$_{(+S)}$ ~~$_{+D\ N}$~~ (John)]]]
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:
   PhP(=[+D N]**)** = … $_{V}$[buy [$_{(+S)}$ ~~$_{+D\ N}$~~ *(John)*] [*(buy)* [$_{+D\ N}$]]]
   Merge(… $_{V}$[buy [$_{(+S)\ +D\ N}$ *(John)*] [*(buy)* [$_{+D\ N}$]]], *M1*[$_{(+wh)\ +D\ N}$ *(what)*]) =
   … $_{V}$[buy [$_{(+S)\ +D\ N}$ *(John)*] [*(buy)* [$_{(+wh)\ +D\ N}$ *(what)*]]]
9. The derivation here terminates since no more selection features must be computed and the memory buffer is empty as required by (54).

Reiterating these steps recursively we can easily derive the successive cyclic movement in (57b).

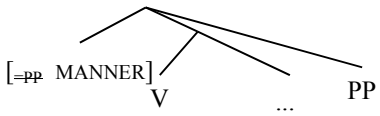
5.3. Right hand adjuncts

As in any antisymmetric framework, which with the present approach share the necessity of deriving linear order from hierarchical one, we would not expect adjuncts, which are functional-related positions (i.e. functional feature in the extended projection of a lexical head in the current approach, specifiers in an antisymmetric approach), to be placed to the right of the head. To explain the fact that certain adjuncts sometimes are rightward, we need either to rely on an interleaving option between Phase Projection and Merge, or on the possibility of stacking the Phase Projections after the last selected complement. Before discussing which solution better fits with the relevant data under analysis, let us understand how we can determine different head-adjunct orders. I assume that a rightward "shifting" operation becomes an option when a functional feature is "complex", namely when it bears a select feature. In this case, two possibilities should be available (delayed Phase Projection or Stacking) and we would obtain (58b) from (58a) (notice the minimal difference between an adjunct selected by a functional feature (58a) and a selected (prepositional) argument (58c)).

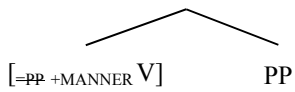
(58) a.



b.



c.



BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

There are reasons to believe that such solution is not arbitrary and it is justified by complexity reduction strategies (Chesi 2007) aiming at "minimizing nesting" as long as the relevant 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:

(59) *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 (e.g. to the right[21]), 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 us to project freely under non-local conditions, therefore the phase head has to be somehow moved in the new peripheral position. The next paragraph suggests a trigger for this movement operation.
As for the two available solutions, they are clearly distinct as shown below:

(60) a. Delayed Phase Projection
$[[_{+A\ =A}\ (X)]\ [_{+B\ =B}\ (Y)]\ _{C}\ ...] \rightarrow [[_{+A}\ (X)]\ [_{+B}\ (Y)]\ [_{C}\ ...\ X\ [_{A}]\ Y\ [_{B}]\ ]]$
b. Stacking
$[[_{+A\ =A}\ (X)]\ [_{+B\ =B}\ (Y)]\ _{C}\ ...] \rightarrow [[_{+A}\ (X)\ [_{+B}\ (Y)\ [_{C}\ ...]\ Y\ [_{B}]]\ X\ [_{A}]\ ]]$

The mirrored prediction made by stacking seems to be more promising (remember the nesting dependency discussed in (21)) and it will be considered first.

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 2010) is that the long distance dependency through memory buffers fits with QR as well[22]. 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);
ii. It is not feature-driven[23];
iii. It cannot be freely cyclic[24].

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 is not interpreted there; this element should be remerged (then interpreted) after the

[21] This might be another possible parameter: head initial languages marginalize at the right edge (shifting), while head-final languages marginalize at the left edge of the matrix phase (scrambling).
[22] Cf. lowering discussed in McCawley (1999).
[23] But see Begelli and Stowell 1981 and discussion in Bianchi and Chesi 2010.
[24] But see Cecchetto 2004 and discussion in Bianchi and Chesi 2010.

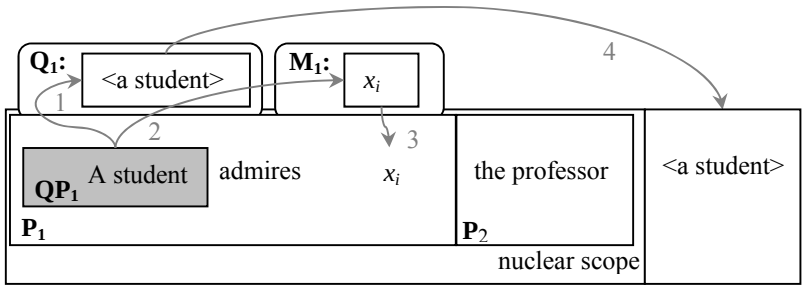
BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

relevant nuclear scope has been created. It follows that QR is actually a *rightward* movement (as proposed, without discussion, in Fox and Nissenbaum 1999).
More precisely, we proposed that QR:

i. stores a QP in a dedicated memory buffer of the current phase (Schlenker 2005);
ii. integrates a coindexed variable in the corresponding argument position or memory buffer (depending on where the QP is processed: selected vs. nested position);
iii. 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 built so far. The elements retrieved from memory buffers are (typically) not spelled out (footnote 14, Chesi 2004); hence QR is "covert".

For example, in the computation of (61), P_1 (a verbal phase) is the first phase computed and the subject QP_1 (a quantified nominal phase) constitutes a nested phase. This is computed while the nuclear scope (P_1+P_2) is still incomplete. Then, the nested QP_1 is stored in the Q-buffer (Q_1) of the containing phase (P_1) 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 (P_2) which is still part of the nuclear scope and only at the end of it, we can remerge the QP.

(61)



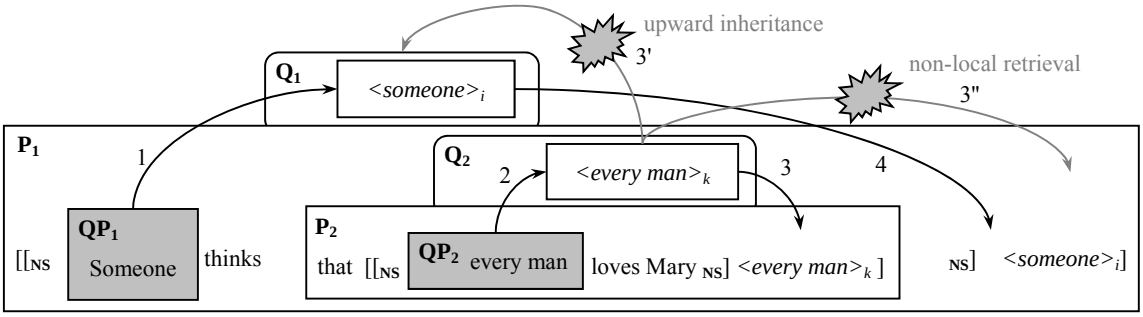
Notice that this is sufficient to explain clause-boundedness (independently predicting the Right Roof Constraint in (37)): since buffers are phase-local, a QP cannot attach to a superordinate phase, this explains the impossibility of getting inverse scope in sentences like the following one:

(62) Someone thinks [$_{CP}$ that every man loves Mary]. (∃>∀; *∀>∃)

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):

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

(63)



Then the clause-boundedness of QR follows from the computational sequencing of phases: the matrix subject QP_1 is stored in the Q-buffer (Q_1) of the containing phase P_1 (the variable insertion in the argument position is ignored in the derivation). The embedded subject QP_2 is stored in the Q-buffer of P_2 (Q_2), but it cannot "get into" the Q-buffer of the previously computed P_1. As a result, QP_2 will only have scope over the embedded P_2, whereas QP_1 will have scope over both P_1 and P_2. 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 Top-Down Left-Right Grammar

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)-(19) binding provides an important set of tests to detect the attachment point of an extraposed element.

Following Bianchi 2010 (phase-based extension of Schlenker's 2005 proposal) we can implement A-binding using a *Referential Buffer*[25] 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 of the processing for the sentence to be grammatical;
ii. We admit some degree of freedom in retrieving the referents in the memory buffer (no rigid LIFO structure is assumed[26]);

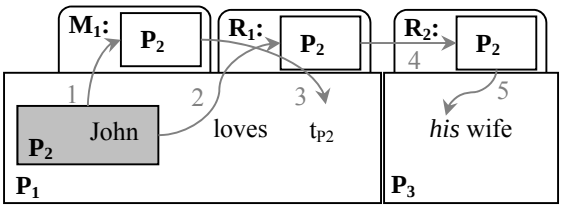
Below the schematic example on how referents are retrieved then evaluated:

[25] 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 2010 for discussion).

[26] This is because locality does not seem to play a role in binding as strong as in movement (but see Grosz et al. 1995).

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

(64)



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 is closed. Schlenker (2005) and Bianchi (2010) propose that whenever in the evaluation sequence (which roughly corresponds to our memory buffer) an element which is already 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 ("I gave him a painting yesterday of John" schematically depicted later in (65)), the "heavy" (i.e. containing select features) DP ("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, (58b)) for minimizing nesting (59). The four available solutions are:

i. Keep everything "in situ" ("*yesterday* we saw *a painting of John*");
ii. Both shift the whole (heavy) adjunct and the whole (heavy) constituent ("we saw *yesterday a painting of John*")
iii. Shift the adjunct and just phase projection of the heavy NP ("we saw a painting *yesterday of John*")
iv. Just shift the adjunct and leave the heavy constituent is situ ("we saw *a painting of John* yesterday")

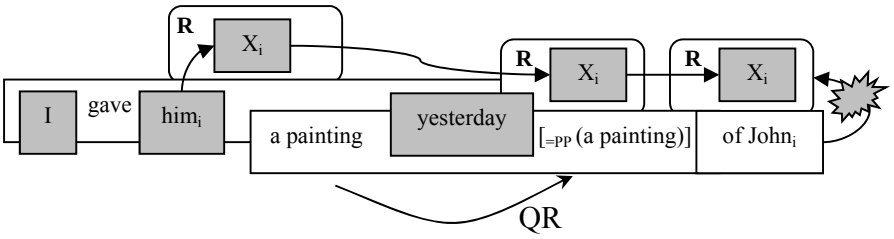
We are not interested now in ranking or excluding (for independent reasons) some of the proposed solutions[27]. What is crucial, according to the discussion in §5.3, is that, in order to get the EXT version in iii., we need some movement trigger to keep Phase Projection local to the verbal head (thus guaranteeing a sort of Complement-Principle à la Guéron and May, 1984). Following Fox and Nissenbaum (1999), I propose that the relevant 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-

[27] Here, my intent is 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). A precise account of optionality is beyond the scope of this paper.

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

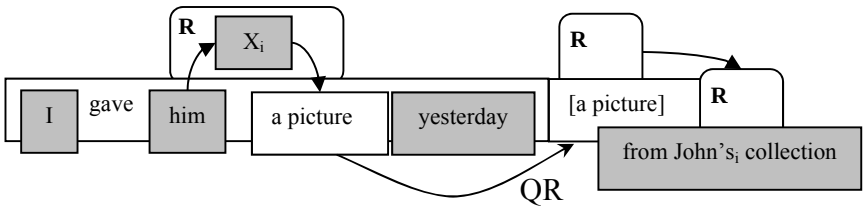
Projection, then Merging the relevant argument. Notice that the selection requirements of the nominal phase head is unsatisfied in situ and this prevents the relevant phase from being closed (the last selection requirement is not "phase projected"), and processed as discussed in §5.5. The content of 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"[28], is inserted in the local R-buffer yielding a condition C effect.

(65)



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, once the NP has been remerged after QR, the evaluation sequence (i.e. the memory buffer) is no more available.

(66)



As in the original Guéron and May (1984) and Fox and Nissenbaum (1999) proposal, EXT is ("clause"/)phase-bounded since:

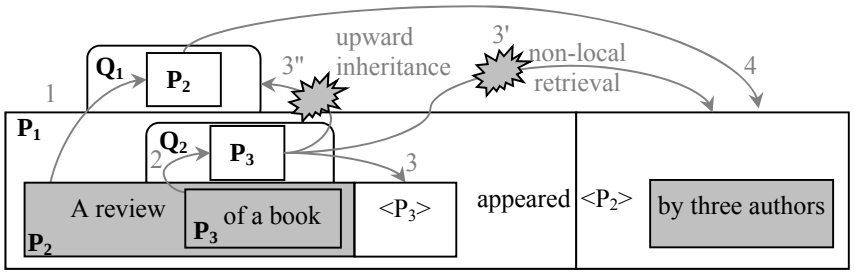
i. QR cannot scope out of the superordinate phase (Chesi and Bianchi 2010);
ii. An adjunct clause, in order to be a nested phase, needs to be attached to the first open (and compatible) superordinate phase.

For the very same reason we predict that, by default, embedded PPs cannot host EXT targeting a superordinate phase (the Q-buffer does not permit neither upward inheritance, nor non-local retrieval):

[28] The evaluation of *him* requires that some individual be already present in the referential memory buffer once the pronoun is processed the first time. This is possible because the R-buffer is populated by individuals that are in the relevant conversational *common-ground* (Bianchi 2010).

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

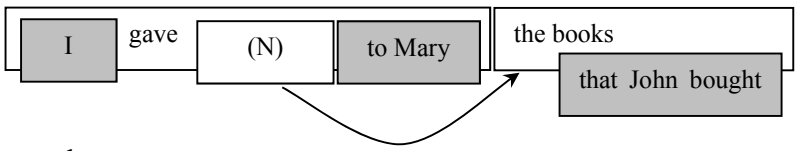
(67)



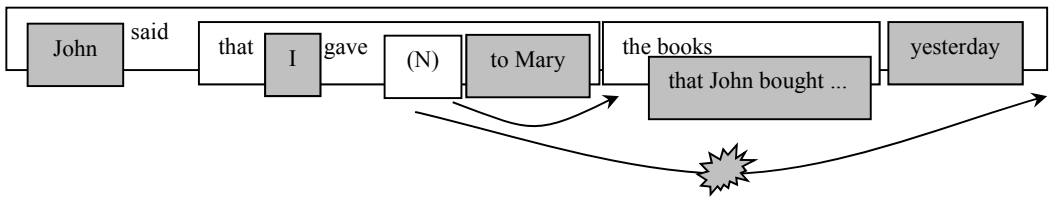
5.7. Heavy NP-Shift

As briefly suggested in the paragraph before, Minimize Nesting (59) 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:

(68) a.

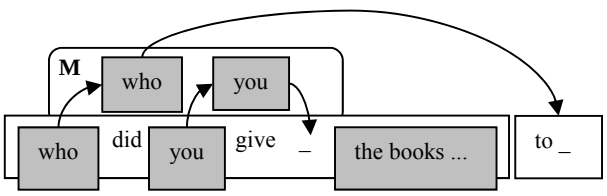


b.



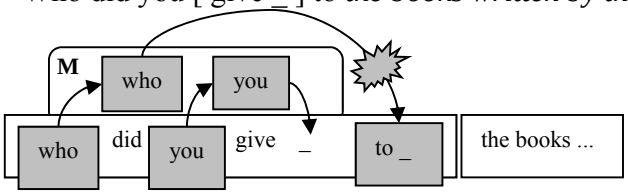
Notice that when the shifted phase is selected, it becomes the "last selected phase", that is, we can predict movement "from" this constituent ((69a) vs. movement from the internal one (69b)):

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



BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

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



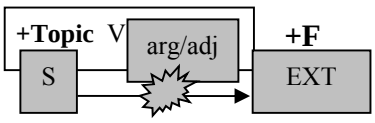
Since we can correctly predict that shifted selected constituents behave as last selected (sequential) phases, the "ne" extraction from focalized and heavy NPs is straightforwardly captured ((29).b vs. (30)).

5.8. Remaining issues

In the remaining pages I would like to sketch the tentative solution for two other issues that I did not have space to discuss in depth here: the ban on EXT from the canonical subject position in Italian and the availability of EXT from embedded constituent in Dutch and, sometimes, in English too.

We noticed in §2.2 that EXT from the subject position is not cross-linguistically uniform: English (10a), for instance, allows for such operation but Italian does not (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 (59)). As noticed by Calabrese (1982), 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 focalized[29]) makes it incompatible with the status of "focalized constituent" that is necessary in order to allow EXT of an inner constituent[30]. This is indeed possible under focalized conditions that seem to be the standard conditions for the subject in languages like English:

(70) 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 either:

[29] 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).

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

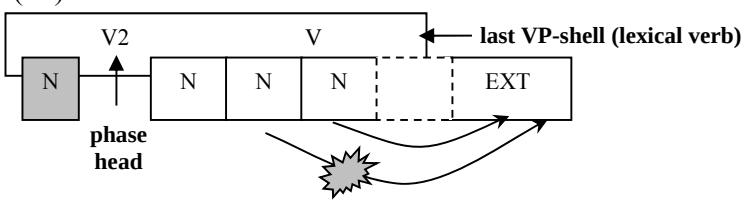
BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

(71) 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) and, sometimes, also in English as discussed in Strunk and Snider (2013). I am inclined to believe that such possibility is much more restricted than what has been reported in literature[31] and it is essentially possible only from the last selected constituent (the last VP-shell):

(72)



Despite the fact that these data deserve more scrutiny, this option could be coherent with the nature of the last selected (sequential) argument since the last complement has some degree of independence with respect to the selecting phase. Then I would prefer not to weaken the Right Roof Constraint or the (generalized) Subjacency idea (as I reframed it in Top-Down terms), but rather concentrate the analysis on this very restricted set of subcases.

6. 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, which are *clause boundedness*, insensitivity of Extraposition to *adjunct/argument asymmetry* and the *definiteness constraint*. All of these properties can be explained 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 still able to characterize as finely as necessary such non feature-driven movements and to capture how they affect scope relations (e.g. selectively bleeding condition C effects, Fox and Nissenbaum 1999, following the derivational top-down implementation of binding principles proposed by Schlenker 2005, Bianchi 2010). With respect to Extraposition, I have argued 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),

[31] "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 2013).

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

moreover, the directionality of QR does not need to be stipulated (Bianchi and Chesi 2010, vs. Fox and Nissenbaum 1999).

In this system, every long distance dependency (e.g. A'-movements, QR etc.) is regulated by specific triggers (e.g. unselected argumental features inserted in a functional position, the unavailability of a relevant domain to be computed, e.g. nuclear scope in the case of QR) using memory buffers, which are locally connected to the notion of phase and regulated by a simple discharge 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 selected complement) that allows for successive cyclic movements, with no need of look-ahead features 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 marginalize "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 of the memory buffers: they are phase-bounded (i.e. "right-roof constrained") 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.

It is important to stress, in this volume, that these empirical generalizations required a derivational assumption that contrasts with the Minimalist one, which is widely assumed to be "bottom-up". In fact, assembling phrase structures Top-Down results in a phrase structure growth that is "from left to right", since expectations must be fulfilled only after they are introduced in the derivation. This produces a precise derivational sequence that is compatible with the sequence in which the items are parsed and generated in the structure. Also crucial is the subdivision of the derivation in phases that are computational chunks driven by local selection and feature lexicalization. It is the geometry and sequence of phase expansions that permit to explain restrictions like the right-roof constraint that must be otherwise simply stipulated. Since the empirical generalizations here discussed seem to be correct and many apparently unrelated properties have now an integrated explanation, I think that this approach might be promisingly on the right track.

References

Abelson, H., & Sussman, J. (1996) *Structure and Interpretation of Computer Programs*. MIT Press, Cambridge, MA, pp. 261–264.

Adger, D. (2006) Combinatorial Variation. *Journal of Linguistics* 42, 503-530.

Adger, D. (2007) A Minimalist Theory of Feature Structure. *http://ling.auf.net/lingBuzz/000583*

Akmajian, A. (1975) More evidence for an NP-cycle. *Linguistic Inquiry* 6, 114-130.

Baker, M. C. (1988) *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Baltin, M. (1981) *Strict Bounding*. Boston MA: MIT Press.

Baltin, M. (2006) Extraposition. In The Blackwell Companion to Syntax, Martin Everaert, Henk Van Riemsdijk (eds), 237-271. Malden MA: Blackwell.

Beghelli, F. & Stowell, T. (1997) Distributivity and negation. In *Ways of scope taking*. A. Szabolcsi (ed.) Dordrecht: Kluwer. 77-109

Belletti, A. (2001) Inversion as Focalization. In *Subject Inversion and the Theory of Universal Grammar*, Hulk, A., Pollock, J. (eds):, 60-90. Oxford: Oxford University Press.

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

Belletti, A. & Rizzi, L. (1981) The syntax of "ne": some theoretical implications. *The Linguistic Review*, 1(2): 117-154.

Belletti, A. & Shlonsky, U. (1995) The order of verbal complements: A comparative study. *Natural Language and Linguistic Theory* 13(3): 489-526.

Bianchi, V. (2010) *A note on backward anaphora*. Ms., University of Siena presented at GLOW 30, 2007, Tromso < http://www.ciscl.unisi.it/doc/doc_pub/bianchi_010_back_anaph.pdf >

Bianchi, V. & Chesi, C. (2008) Phases, strong islands, and computational nesting. *STIL, Studies in Linguistics* 2: 81-100.

Bianchi, V. & Chesi, C. (2010) Reversing the perspective on Quantifier Raising, *Rivista di Grammatica Generativa*. 35 (2010):3-38.

Boersma, P. & Hayes, B. (2001) Empirical Tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45-86.

Burzio, L. (1986) *Italian syntax: A government-binding approach*. Dordrecht: Springer.

Calabrese, A. (1982) Alcune ipotesi sulla struttura informazionale della frase in italiano e sul suo rapporto con la struttura fonologica. *Rivista di Grammatica Generativa* 7: 3-79.

Cardinaletti, A. (1991) On pronoun movement: The Italian dative loro. *Probus*. 3(2): 127-154.

Cardinaletti, A. (2003) Toward a Cartography of Subject Positions. In *The Structure of CP and IP. The Cartography of Syntactic Structures Vol.2*, Luigi Rizzi (ed.), 115-165. New York: Oxford University Press.

Cecchetto, C. (2004) Explaining the locality conditions of QR: Consequences for the theory of phases. *Natural Language Semantics*. 12: 345-397.

Chesi, C. (2004) Phases and Cartography in Linguistic Computation: Toward a Cognitively Motivated Computational Model of Linguistic Competence. PhD dissertation, University of Siena.

Chesi, C. (2007) An introduction to Phase-based Minimalist Grammars: Why move is Top-Down from Left-to-Right. *STIL, Studies in Linguistics* 1: 38-75.

Chomsky, N. (1981) *Lectures on Government and Binding: The Pisa Lectures*. Dordrecht: Foris Publications.

Chomsky, N. (1995) *The Minimalist Program*, Cambridge MA: MIT Press.

Chomsky, N. (1999) Derivation by Phase. *MIT Occasional Papers in Linguistics, 18*. Cambridge MA: MITWPL.

Chomsky, N. (2008) On Phases. In *Foundational issues in linguistic theory: essays in honor of Jean-Roger Vergnaud,* Jean-Roger Vergnaud, Robert Freidin, Carlos Peregrín Otero, Maria Luisa Zubizarreta (eds), 133-166. Cambridge MA: The MIT Press.

Citko, B. (2008) An Argument against Assimilating Appositive Relatives to Coordinate Structures. *Linguistic Inquiry* 39.4: 633-655.

Culicover, Peter W. & Rochemont, Michael S. (1990) Extraposition and the Complement Principle. *Linguistic Inquiry* 21:23-47.

Culicover, P. W., & Rochemont, M. S. (1997) Deriving Dependent Right Adjuncts in English. In: *Rightward Movement*. D. Beermann, D. LeBlanc, and H. C. van Riemsdijk (eds.). Philadelphia: John Benjamins.

de Vries, M. (1999) Extraposition of Relative Clauses as Specifying Coordination. In *Proceedings of ConSole VII*, Tina Cambier-Langeveld, Anikó Lipták, Michael Redford & Erik Jan van der Torre (eds), 293-309.

de Vries, M. (2002) *The syntax of relativization*. PhD dissertation. University of Amsterdam.

de Vries, M. (2006) The Syntax of Appositive Relativization: On Specifying Coordination, False Free Relatives, and Promotion. Linguistic Inquiry 37: 229-270.

Diesing, M. (1992) *Indefinites*. Cambridge MA: MIT Press.

Emonds, J. (1979) Appositive Relatives Have No Properties. *Linguistic Inquiry*. 10: 211-243.

Fodor, J.D., (1978) Parsing strategies and constraints on transformations. *Linguistic Inquiry*. 9, 427–473.

Fox, D. & Nissenbaum, J. (1999) Extraposition and scope: A case for overt QR. *Proceedings of the 18th West Coast Conference on Formal Linguistics*. 132-144.

Göbbel, E. (2007) Extraposition as PF movement. In *Proceedings of WECOL 2006*. Bainbridge, E. & B. Abayani (eds). 132-145.

Grimshaw, J. (1991) Extended projection. In *Lexical specification and insertion*. P. Coopmans, M. Everaert & J. Grimshaw (Eds.) The Hague: Holland Academic Graphics. 115-134.

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)

Grosz, B. J., Weinstein, S. & Joshi, A. K. (1995) Centering: A framework for modeling the local coherence of discourse. *Computational linguistics* 21(2): 203-225.
Guéron, J. & 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. & Na, Y. (1990) Extraposition and Focus. Language 66: 51-77
Jackendoff, R. (1990) On Larson's treatment of the double object construction. *Linguistic Inquiry* (21): 427-456.
Johnson, D. & Lappin, S. (1997) A Critique of the Minimalist Program. *Linguistics and Philosophy*, 20(3): 273-333.
Kayne, R. (1994) *The Antisymmetry of Syntax*. Cambridge MA: MIT Press.
Larson, R. K. (1988) On the double object construction. Linguistic Inquiry (19): 335-392.
Lebeaux, D. (1988) Language acquisition and the form of the grammar. PhD dissertation, University of Massachusetts, Amherst.
McCawley, J. D. (1999) Why surface syntactic structure reflects logical structure as much as it does, but only that much. *Language* 75(1): 34-62.
Phillips, C. (1996) Order and Structure. PhD dissertation. MIT.
Pinker, S. (1994) *The Language Instinct*. New York: HarperCollins.
Pollard, C. & Sag, I. (1994) *Head-driven phrase structure grammar***.** Chicago: University of Chicago Press.
Richards, N. (1999) Dependency formation and directionality of tree construction. *MIT Working Papers in Linguistics* 34**:** 67-105.
Riemsdijk, H. van, and Williams, E. (1981) NP-Structure. *The Linguistic Review* 1: 171-217.
Rizzi, L. (1990) *Relativized minimality*. Cambridge MA: MIT Press.
Rizzi, L. (1997) *The Fine Structure of the Left Periphery*. Cambridge MA: MIT Press.
Rizzi, L. (2006) On the Form of Chains: Criterial Positions and ECP Effects. In *Wh-Movement: Moving on*. Lisa Cheng & Norbert Corver (eds), Cambridge MA: MIT Press. 97-133.
Rosenbaum, P. S. (1965) The Grammar of English Complement Constructions. PhD dissertation, MIT.
Ross, J. R. (1967) Constraints on Variables in Syntax. PhD dissertation. MIT.
Schlenker, P. (2005) Non-redundancy: towards a semantic reinterpret-ation 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.
Stabler, E. (1997) Derivational minimalism. In *Logical Aspects of Computational Linguistics***.** Christian Retoré (ed). Berlin: Springer. 68-95.
Staub, A., Clifton, C. & Frazier, L. (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. In *The theory of markedness in generative grammar***.** Adriana Belletti, Luciana Brandi & Luigi Rizzi (eds) 475–516. Pisa: Scuola Normale Superiore.
Van Dyke, J. A., & McElree, B. (2006) Retrieval interference in sentence comprehension. *Journal of Memory and Language*, 55(2), 157-166.
Vergnaud, J.-R. (1974) *French Relative Clauses*. PhD dissertation, MIT.
Wilder, C. (1995) Rightward Movement as Leftward Deletion. In *On Extraction and Extraposition in German*. U. Lutz and J. Pafel (eds.), 273–309. Amsterdam and Philadelphia: John Benjamins.
Williams, E. S. (1974) *Rule Ordering in Syntax*. PhD dissertation, MIT.
Zubizarreta, M. L. (1998) *Prosody, Focus, and Word Order*. Cambridge MA: MIT Press.

BDD-A22721 © 2013 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.217.31 (2026-11-26 23:04:33 UTC)