Larger than expected: constraints on pied-piping across languages

Cristiano Chesi, Pauli Brattico

UNIVERSITY SCHOOL FOR ADVANCED STUDIES, IUSS PAVIA

Abstract. In this paper we focus on pied-piping, a phenomenon that undermines the standard probe-goal theory of movement. Specifically, we argue that several constraints that must be stipulated under the standard bottom-to-top theory follow from a top-down redefinition of Merge, Move and Phase domain. Two universal constraints will be specifically addressed ("Don't Leave Your Complement alone" – *Don't-LYC* and "Be Accessible" - *BA constraints*) and derived together with their (apparent) exceptions. The empirical evidence considered includes alternation between preposition stranding and pied-piping, recursive genitive pied-piping and the *Was-für* split.

1 Introduction

Under the Probe-Goal approach to movement (Chomsky 2000; 2001), the Probe triggers the displacement of the closest C-commanded Goal and its re-merge into the local spec-Probe position. Details vary depending on the formulation, but they all rely on an asymmetry between the Probe and the Goal that can be expressed in terms of valuation/interpretability[1] and/or in terms of position[2]. The following definition captures what we consider to be the lowest common denominator of various definitions found in the relevant literature:

[1]The Probe bears the relevant uninterpretable (Chomsky 1998) or unvalued (Chomsky 2000) features and the C-commanded Goal, with the relevant valued/interpretable features, is attracted to create the spec-head relation necessary to value/check the unvalued/uninterpretable features.

[2]The Probe and the Goal features can be exactly the same, but the Probe licenses a criterial, scope-discourse position, while the Goal is inserted into a s-selected, thematic position (Rizzi 2004).

(1) The **Probe-Goal** dependency
A newly merged Probe, endowed with a probing feature $+p$, immediately triggers overt re-merge of the closest Goal endowed with the same feature p.

Under the standard bottom-up, Merge-based approach, this formulation captures the structural asymmetry between the C-commanding Probe and the C-commanded Goal (the Goal must be already merged at least once within the structure just merging with the Probe to be re-merged), adopts the null hypothesis that the feature in the Probe and the one probed in the Goal do not need to be different (Rizzi 2004) and, furthermore, captures locality ("closest Goal"; see §2 for details). This definition is not sufficient, however, however, to capture even the basic contrasts in (2) and (3):

(2) a. $[_{wh}$ what$]_i$ C_{wh} did you buy $_i$?
b. $[_{wh}$ which $[_{\text{NP}}$ book$]]_i$ C_{wh} did you buy $_i$?
c. * $[_{wh}$ which$]_i$ C_{wh} did you buy [$_i$ $[_{\text{NP}}$ book]] ?

(3) a. $[_{wh}$ *which* photographer$]_i$ C_{wh} did you buy $[_{\text{DP}}$ pictures $[_{\text{PP}}$ of $_i$]]?
b.?* $[_{\text{PP}}$ of $[_{wh}$ *which* photographer$]]_i$ C_{wh} did you buy $[_{\text{DP}}$ pictures $_i$]?
c. * $[_{\text{DP}}$ pictures $[_{\text{PP}}$ of $[_{wh}$ *which* photographer$]]]_i$ C_{wh} did you buy $_i$?

We would expect (2c) instead of (2b) to be the minimal configuration compliant with (1). In fact, the solution (2c) seems to be plausible, at least in some languages, as apparently suggested by (4b):

(4) a. [*Was* für Bücher$]_i$ hast Du gelesen $_i$? (German, Blümel 2012)
b. *Was*$_i$ hast Du [$_i$ für Bücher$]_j$ gelesen $_j$?
what have you for books read
'What kind of books did you read?'

The *Was-für split* configuration suggests that we need a generalization (eventually deductible from other constraints) to capture the contrast between (2b) and (2c). Let us adopt provisionally the following constraint:

(5) **Don't Leave Your Child/Complement alone (Don't-LYC) constraint**
A Goal must pied-pipe its complement to satisfy a Probe-Goal relation.

The *Don't-LYC* constraint correctly captures the contrasts in (2), the DP movement in (3a) and (4b), but not the contrasts between (3a) and (3b-c) or the option in (4b). On the other hand, the full paradigm in (3) is already predicted by the strict application of (1): since the target feature is on D(P), the selecting P(P) has no reason to be moved/pied-piped together with the Goal DP. Yet, there is both cross-linguistic (6) and intra-linguistic (7) evidence that undermines this generalization:

(6) a. * [$_{wh}$ *quale* fotografo]$_i$ hai comprato [$_{DP}$ foto [$_{PP}$ di $_i$]]?
which photographer have.2sg bought photos of
b. [$_{PP}$ di [$_{wh}$ *quale* fotografo]]$_i$ hai comprato [$_{DP}$ foto $_i$]?
c. ? [$_{DP}$ foto [$_{PP}$ di [$_{wh}$ *quale* fotografo]]]$_i$ hai comprato $_i$?

(7) a. the person [of whom]$_i$ [pictures $_i$] are on the table
(Chomsky 1986)
b. * the man who$_i$ [pictures [of $_i$]] are on the table

Italian is among those languages not permitting P stranding, (6a), and letting the Goal to pied-pipe material above it (6b-c). On the other hand, P stranding is necessarily avoided in relative clauses (7a-b) vs. *wh*-sentences (3a-b) (as discussed in Bianchi & Chesi 2015). Again, to capture (3a-b), we can provisionally formulate another constraint (that seems not to be operative neither in (6) nor in (7)):

(8) **Be Accessible (BA) constraint**
The relevant Goal feature must be at the edge of the moved constituent.

Contrary to the apparent counterexamples, we argue that *Don't-LYC* and *BA* are universal constraints that follow from a re-definition of the Probe-Goal dependency. Our hypothesis is that most of the constraints of pied-piping follow by adopting a Top-Down derivation (Chesi 2015) in which the phrase structure (and any Probe-Goal dependency) is created incrementally, from-left-to-right in a phase-chunked computation.

The paper is organized in the following way: we will first discuss the major classes of phenomena we need to derive, most of them falling under the rubric dubbed pied-piping (§2). We will then discuss counterexamples that apparently disconfirm the Don't-LYC and BA principles; we will focus on the weakest points of the recent analyses presented in the current literature (Abels 2012; Heck 2009) trying to account for these data. In the end, we will present the derivational approach (§3) in which the top-down reformulation of Merge, Move and Phase domain derives the Don't-LYC and BA constraints while accommodating the major counterexamples.

2 Pied-piping: definitions, facts and generalizations

Before discussing the empirical evidence, we will specify the linguistic framework assumed as a background. It is clear that the Probe-Goal dependency (1) cannot be primitive. The notion of "probing feature", "closest", "C-command" and also the notion of "Probe" and "Goal" must be defined precisely to avoid any ambiguity. We begin with the simplest possible definitions for the crucial components that will be used throughout the paper (see Stabler 1997; Collins & Stabler 2016 for a more complete, sometime different, set of definitions):

Definition 1. *(Abstract) features*[3]
Three kinds of abstract features are considered:

- *Category*, expressing the basic morpho-syntactic categories (e.g. C, T, D, V, N...);
- *Selection*, indicating categorial selection ($=X$ indicates that a category X is selected);
- *Licensor*, indicating criterial/scope-discourse licensing ($+Y$ indicates that a category Y is licensed).

As usual (Chomsky 1995: 30), lexical items are considered idiosyncratic clusters of phonetic, semantic and abstract features stored in the lexicon. No morphemic analysis is approached here; lexical items, with their features, are the atomic components to be combined in phrases according to the *Merge* operation:

Definition 2. *Merge*
Merge is a binary operation taking two items α and β, either atomic (i.e. lexical items) or resulting from previous merge operations, and returns the labeled, unordered set formed by α and β. The label is either α or β. (e.g. Merge(α, β) $= {}_{\alpha}[\alpha\ \beta]$ or ${}_{\beta}[\alpha\ \beta]$).

We assume here tentatively that the labeling algorithm reduces to projecting the selecting/licensing head (see Rizzi 2016 for a more articulated discussion):

[3]Also other kinds of features should be considered: *Phonetic*, expressing instructions for the sensory-motor interface; *Semantic*, expressing instructions for the conceptual-intentional interface. Here we simply consider abstract features driving combinatorial options. Moreover, we opt here for a simplistic definition of features not organized in feature-structures (as Attribute Value Matrices in HPSG, Pollard & Sag 1994). We did not include either, in the basic definitions, any notion of "valuation" or "sharing" (see §2.2.1) of morphological properties; this could be easily added in the form *LABEL:value* (e.g. GENDER:fem, CASE:nom etc.), extending the category set.

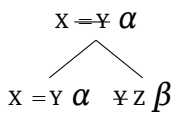
Definition 3. *Labeling algorithm*
If α merges with β, then either α selects/licenses β or β selects/licenses α. The selecting/licensing item is the one projecting, namely the label of the set created. Select, selected and licensor features are deleted after successful merge.

Notice that licensed features are not deleted after (re-)merge (this will create a favourable environment for successive cyclic movement). Assuming that selection must be local, while licensing can be triggered "at distance", forcing re-merge of a previously merged item (i.e. movement), we can adopt the following definition:

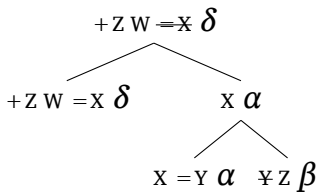
Definition 4. *Selection and licensing*
Selection is always local and preempts licensing; only licensing can trigger re-merge of a previously merged item (i.e. movement).

Roughly speaking, using X′ terminology, *selection* would correspond to the head-comp relation, while *licensing* to spec-head (both selection and licensor features should be hosted by the head). To clarify, assume we have four categorial features *X, Y, W* and *Z* and a lexicon formed by three items with the relevant categorial select and licensor features associated to them as follows[4]: $[_{X\ =Y}\ \alpha]$ $[_{Y\ Z}\ \beta]$ and $[_{+Z\ W\ =X}\ \delta]$. Then we can simulate a simple derivation:

(9) i. *Merge*($[_{X\ =Y}\ \alpha]$,$[_{Y\ Z}\ \beta]$)

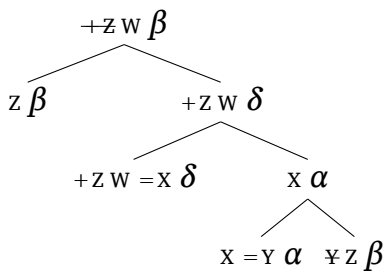


ii. *Merge*($[_{X}\ \alpha]$,$[_{+Z\ W\ =X}\ \delta]$)



iii. *Merge*($[_{+Z\ W}\ \delta]$,$[_{Z}\ \beta]$)

[4]Features are conventionally ordered in the lexical entry: <<licensors>, <categories>, <selections>>.



We can now easily define Probes and Goals (respectively δ and β in (9)):

Definition 5. *Probes and Goals*
Probes and Goals are lexical items; Probes must have a licensor "+X" feature forcing (re-)merge of the closest Goal with an X category feature.

This definition is sufficient to capture the positional asymmetry between Probe and Goal, but crucially rely on "the closest" modifier for capturing locality. We define this notion, and the notion of (derivational) C-Command, as follows:

Definition 6. *Closest*
Taking three items, α, β and γ, either atomic (i.e. lexical items) or resulting from previous merge operations, β is the closest item to α iff α C-commands β and there is no other γ such that γ C-commands β and α C-commands γ.

Definition 7. *C-command (derivational, merge-based version)*
α c-commands β and any items merged within β iff α directly merges with β.

Considering "a phrase" as the result of a Merge operation, and containing / contained_in α respectively as a synonym of "merged with α and projecting" / "merged with α without projecting", we can define *pied-piping* as follows:

Definition 8. *Pied-piping*
Movement of a larger phrase *X containing* the Goal α (a.) or *contained* in α (b.):
a. $[_{\delta}$ **[**$_{X}$ **X ...** $[\alpha_{Goal}]$**]** $_{i}$ $\delta_{Probe}]$ [... Y [... $_{i}$]]?
b. $[_{\delta}$ $[_{\alpha}$ α_{Goal} **[**$_{X}$ **X ...]** $]_{i}$ $\delta_{Probe}]$ [... Y [... $_{i}$]]?

Given the informal definition in (1), pied-piping as introduced in Definition 8 cannot be derived. A reformulation of the basic components of our grammar is then necessary. Let us schematize the set of relevant possibilities we could consider (pseudo-English examples are provided to illustrate the different options):

(10) pre-movement configuration
$\boldsymbol{\delta}_{\text{Probe}}$ … [$_{\textbf{XP}}$ [$\boldsymbol{\alpha}_{\text{Goal}}$ [$_{\textbf{YP}}$ …]]]
$\mathbf{C_{wh}}$ did John buy a book [$_{\textbf{XP}}$ of [$\mathbf{which_{wh}}$ [$_{\textbf{YP}}$ scholar]]]

i. no pied-pipe
α_{Goal} [$\boldsymbol{\delta}_{\text{Probe}}$ … (_α) [$_{\textbf{XP}}$ [_α [$_{\textbf{YP}}$ …]]]
$\mathbf{which_{wh}}$ $\mathbf{C_{wh}}$ did John buy [a book [$_{\textbf{XP}}$ of _$\mathbf{wh}$ [$_{\textbf{YP}}$ scholar]]]

ii. containing (XP) + contained (YP) strand (complement stranding)
[$_{\textbf{XP}}$ [$\boldsymbol{\alpha}_{\text{Goal}}$]] [$\boldsymbol{\delta}_{\text{Probe}}$ … [_$\mathbf{XP}$ [$_{\textbf{YP}}$ …]]]
[$_{\textbf{XP}}$ of $\mathbf{which_{wh}}$]$\mathbf{C_{wh}}$ did John buy [a book [_$\mathbf{XP}$ [$_{\textbf{YP}}$ scholar]]]

iii. containing (XP) + contained (YP) pied-pipe (full pied-pipe)
[$_{\textbf{XP}}$ [$\boldsymbol{\alpha}_{\text{Goal}}$ [$_{\textbf{YP}}$ …]] [$\boldsymbol{\delta}_{\text{Probe}}$ … _$\mathbf{XP}$]
[$_{\textbf{XP}}$ of $\mathbf{which_{wh}}$ [$_{\textbf{YP}}$ scholar]]$\mathbf{C_{wh}}$ did John buy [a book _$\mathbf{XP}$]

iv. containing (XP) strand + contained (YP) pied-pipe
[$\boldsymbol{\alpha}_{\text{Goal}}$ [$_{\textbf{YP}}$ …]] [$\boldsymbol{\delta}_{\text{Probe}}$ … [$_{\textbf{XP}}$ _α]]
[$\mathbf{which_{wh}}$ [$_{\textbf{YP}}$ scholar]]$\mathbf{C_{wh}}$ did John buy [a book [$_{\textbf{XP}}$ of _$\mathbf{wh}$]]

v. move to the edge of the containing XP (edge-move) + full pied-pipe
[[$\boldsymbol{\alpha}_{\text{Goal}}$ [$_{\textbf{YP}}$ …]] $_{\textbf{XP}}$ _α] [$\boldsymbol{\delta}_{\text{Probe}}$ … _$\mathbf{XP}$]
[$_{\textbf{XP}}$ $\mathbf{which_{wh}}$ [$_{\textbf{XP}}$ of _α [$_{\textbf{YP}}$ scholar]]]$\mathbf{C_{wh}}$ did John buy [a book _$\mathbf{XP}$]

We exclude the "no pied-pipe" option (10i) (which would be compatible with BA but incompatible with Don't-LYC) since it would make the wrong predictions both in (2b-c) and (3a) repeated below for convenience.

(2) b. [$_{wh}$ which [$_{\text{NP}}$ book]]$_i$ C_{wh} did you buy _i ?
c. * [$_{wh}$ which]$_i$ C_{wh} did you buy [_i [$_{\text{NP}}$ book]] ?

(3) a. [$_{wh}$ *which* photographer]$_i$ C_{wh} did you buy [$_{\text{DP}}$ pictures [$_{\text{PP}}$ of _i]]?
b.?* [$_{\text{PP}}$ of [$_{wh}$ *which* photographer]]$_i$ C_{wh} did you buy [$_{\text{DP}}$ pictures _i]?

For the same reason we exclude also "complement stranding" (10ii) that is incompatible both with Don't-LYC and with BA.[5] Option "full pied-pipe" (10iii) is compliant with Don't-LYC (it captures the (2b-c) contrast), but it is against BA and it would wrongly reject (3a) in favor of the (ungrammatical) (3b). Option (10iv), stranding the item merged above the moved one, and pied-piping the one merged below is compatible both with BA and Don't-Like, correctly derives (2) and (3), and, furthermore, correctly rejects (3b). Lastly, option (10v) is compliant

[5]We might rescue BA by introducing a definition of "percolation", which we will not consider here (refer to Cable 2010 and Heck 2009 for a critical discussion of this option).

both with BA and Don't-LYC. This case is attested in English (recursive) genitive pied-piping (11) and in Finnish *wh-* "snowball" movement (Huhmarniemi 2012).

(11) a. a person [[[***whose*** [lawyer's [sister's … [son]] adores you
b. *a person [the son … of the lawyer of the sister **of who**] adores you

(12) a. [[[*Mitä* kaupunkia]$_i$ kohti$_{_i}$]$_j$ virtaamalla$_{_j}$]$_k$ Seine
what city towards by.flowing Seine
pääsee valtamereen$_{_k}$?
reaches ocean
b. *Seine pääsee valtamereen [virtaamalla [kohti [*mitä*
Seine reaches ocean by.flowing towards what
kaupunkia]]]?
city
'By flowing towards which city does Seine reach the ocean?'

This provides compelling evidence in support of BA. Before discussing this issue further (§2.2), we will discuss the evidence supporting the validity of the Don't-LYC constraint (§2.1).

2.1 Evidence in favour of the Don't-LYC constraint and (apparent) counterexamples

The "Don't Leave Your Child/Complement alone" (Don't-LYC) constraint in (5) states that a Goal must pied-pipe its complement while satisfying a Probe-Goal relation. According to the definition 2 (Merge) and 3 (Labeling), the relation between the Goal and the "child/complement" will be that of selection or licensing. The selection case is represented by (2b-c):

(2) b. [$_{wh}$ which [$_{NP}$ book]]$_i$ C_{wh} did you buy $_{\_i}$?
c. * [$_{wh}$ which]$_i$ C_{wh} did you buy [$_{\_i}$ [$_{NP}$ book]] ?

The derivation of the DP "which book" is the result of merging [$_{D\ =N}$ which] and [$_N$ book], forming [$_D$ which [$_{D\ \cancel{=N}}$ which] [$_{\cancel{N}}$ book]]. Notice that we do not need to rely on an explicit phrase-head definition to distinguish between phrases and heads here: the phrase-head distinction is inferable from the fact that the "lowest copy" of [$_{D\ =N}$ which], before merge, is not "saturated" (the selecting feature indicates this) while the "highest copy" [$_D$ which] is saturated (no more selecting feature is

present/active and, as a consequence, such an item can enter another merge relation only if selected (in the example by $[_{V\ \ldots\ =D}$ buy]). In this sense, both $[_{D}$ which] subsumes the DP label, and it also corresponds to a maximal projection that will not be able to project any further (no more select, =X, or licensing, +Y, features available triggering further "active" structural expansions). More radically, according to Definition 2 (Merge) and 3 (Labeling), the two "copies" are in fact the very same item at two different times during structure building (pre- and post-merge) and this is sufficient to derive the Don't-LYC constraint: when a Probe is merged, the pre-merge Goal constituent does not exist any further and only "maximal projection" (i.e. merged items, with no more active select features) can become Goals.

Let us now turn to the licensing option: this would be the case of $[_{X\ +Y}\ \alpha\ ]$ merging with $[_{Y}\ \beta]$ resulting in $[_{X\ +\cancel{Y}}\ \alpha\ [_{\cancel{Y}}\ \beta]]$. In this case, α is a Probe (by Definition 5) and β its Goal. Looking at (pseudo-)English to formulate the specific subcase $[_{Y\ +\cancel{Y}}\ \alpha\ [_{\cancel{Y}}\ \beta]]$ (namely a Probe looking for a Goal with its very same category) we might expect, for instance, a *wh-* item to be the Probe of another *wh-* item before being a Goal of a superordinate Probe:

(13) * $[_{+\cancel{wh}\ wh}$ which $[_{\cancel{wh}}$ what]] do you want _ ?

In fact, being a Probe for a Goal that could be the Probe itself is rather odd. This is excluded by Definition 6 (Closest): the probed feature, being contained in the Probe itself, is "closer" than any other compatible Goal. We should then reject (13). The more general configuration (i.e. becoming a Goal after a successful merge of a Probe and another Goal: $[_{X\ +\cancel{Y}}\ \alpha\ [_{\cancel{Y}}\ \beta]]$) cannot be excluded on purely theoretical ground and, in fact, seems to be attested by the "snowballing movement" constructions (Aboh 2004) under the cartographic assumption of a rigid and universal distribution of the functional layers (Cinque 1999).

2.1.1 Snowball movement: Probes becoming Goals

Clear cases of snowball movements are discussed in Aboh (2004:111-113) for Gungbe and Malagasy and in Shlonsky (2004) for Hebrew and various Arabic dialects; the example below is taken from Hebrew (Shlonsky 2004) and illustrates the typical reversal ordering of functional layers within the DP.

(14) rabanim fanatim rabim/mə'atim 'elu
rabbits fanatic many$_{MPL}$/few$_{MPL}$ these
'these many/few fanatic rabbits' (Hebrew, Shlonsky 2004:1497)

Following Shlonsky, such reordering (from the basic word order (15a)) is derived by successive movements, targeting the Probe that triggered the immediately previous movement operation: first the NP targets the adjectival phrase (AP) spec, then the AP targets the (weak) Quantifier Phrase (QP) spec, finally the QP targets the Demonstrative Phrase (DemP) as illustrated in (15b):

(15) a. $[_{DemP}$ ‘elu $[_{QP}$ rabim/mə’atim $[_{AP}$ fanatim $[_{NP}$ rabanim]]]
these many/few fanatic rabbits

b. $[[[_{NP}$ rabanim] $[_{AP}$ fanatim$_{_NP}$]] $[_{QP}$ rabim/mə’atim $_{\_AP}$]]
rabbits fanatic many/few
$[_{DemP}$ ‘elu $_{\_QP}$]]
these

We consider this as an evidence of the fact that the Probe of each movement constitutes the target/Goal of the superordinate Probe. The minimal implementation able to derive the re-ordering in (15a) requires the lexicon in (16) and the assumtion that selected features are not deleted as proposed in Definition 3 (pseudo-English is used for convenience). A possible step-by-step derivation is explored in (17):

(16) $[_{+Q\,D\,=Q}$ these], $[_{+A\,Q\,=A}$ many/few], $[_{+N\,A\,=N}$ fanatic], $[_{N}$ rabbits]

(17) i. merge(fanatic, rabbits): $[_{+N\,A\,\cancel{=N}}$ fanatic $[_{N}$ rabbits]]

ii. move($[_{+N}$ … $[_{N}$ …]]): $[_{A}$ $[_{N}$ rabbits] $[_{\cancel{+N}\,A}$ fanatic $_{\_N}$]]

iii. merge(many/few, $[_{A}$ …]):
$[_{+A\,Q\,\cancel{=A}}$ many/few $[_{A}$ $[_{N}$ rabbits] $[_{A}$ fanatic $_{\_N}$]]]

iv. move($[_{+A}$ … $[_{A}$ …]]):
$[_{Q}$ $[_{A}$ $[_{N}$ rabbits] $[_{A}$ fanatic $_{\_N}$]] $[_{\cancel{+A}\,Q}$ many/few $_{\_A}$]]

v. merge(these, $[_{Q}$ …]):
$[_{+Q\,D\,\cancel{=Q}}$ these $[_{Q}$ $[_{A}$ $[_{N}$ rabbits] $[_{A}$ fanatic $_{\_N}$]] $[_{Q}$ many/few $_{\_A}$]]]

vi. move($[_{+Q}$ … $[_{Q}$ …]]):
$[_{D}$ $[_{Q}$ $[_{A}$ $[_{N}$ rabbits] $[_{A}$ fanatic $_{\_N}$]] $[_{Q}$ many/few $_{\_A}$]] $[_{\cancel{+Q}\,D}$ these $_{\_Q}$]]

The order of the operations (first select, then license) is derived by the asymmetry between select and licensor features (Definition 4). Nothing here hinges on the fact that the probing feature can be different from the selected one (this is a purely empirical matter out of the scope of

this theoretical investigation[6]). Despite the set of theoretical options one could exploit, example (15) demonstrates that a Probe can become a Goal for another Probe during the derivation. Our original guess about the Don't-LYC constraint is fulfilled: Probes never move leaving the probed Goal in situ. This is clear from any (partial) snowball movement attested in Hebrew and Arabic (Shlonsky 2004) but also in other Romance languages, showing reversed adjectival re-ordering with respect to Germanic languages (Cinque 1994):

(18) a. a huge orange fruit
b. un fruit orange énorme (French, Cinque 1994:101)
a fruit orange huge

Along the lines of (17), we derive the specific re-ordering indicated in (18b) as follows:

(18) b′. $[_{\text{D}\ \cancel{=\text{A-size}}}$ a $[_{\text{A-size}}\ [_{\text{A-color}}\ [_{\text{N}}$ fruit$]$ $[_{\cancel{+\text{N}}\ \text{A-color}\ \cancel{=\text{N}}}$ orange $_{\text{N}}]]$
$[_{\cancel{+\text{A-color}}\ \text{A-size}\ \cancel{=\text{A-color}}}$ huge $_{\text{A-color}}]]]$

This snowball analysis, whenever exploited, predicts that given a set of strictly ordered functional layers (e.g. F1, F2, F3) above a lexical category (e.g. L), only a restricted set of re-ordering options should be available cross-linguistically (as observed in Cinque to appear):

(19) a. [F1 [F2 [F3 [L]]]]
b. [F1 [F2 [[L] F3 $_{L}$]]]
c. [F1 [[[L] F3 $_{L}$] F2 $_{F3}$]]
d. [[[[L] F3 $_{L}$] F2 $_{F3}$] F1 $_{F2}$]

2.1.2 Was-für split as an (optional) extraposition

A remaining empirical challenge for the Don't-LYC constraint is represented by the was-für split construction in some Germanic langages, presented in (4) and repeated below:

(4) a. [*Was* für Bücher]$_i$ hast Du gelesen $_i$? (German, Blümel 2012)
b. *Was*$_i$ hast Du [$_i$ für Bücher]$_j$ gelesen $_j$?
what have you for books read
'What kind of books did you read?'

[6]Extra functional layers should be postulated between each Probe and its Goal for creating proper, empirically supported, anti-locality configurations (Abels 2003, 2012); see discussion in §2.2.1.

The dependency between [*was*] and [*für Bücher*] needs scrutiny. Staring from [*für Bücher*], notice that the NP "selected" by *für* cannot be definite/quantified (20) and does not get accusative case assigned by *für* (21b):

(20) [*Was* für *diese/zwei Bücher]$_i$ hast Du gelesen$_{_i}$?
what for these/two books have you read

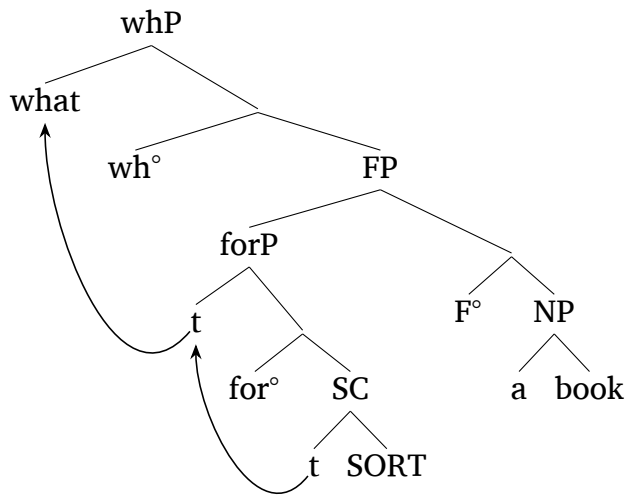
(21) a. Peter hat das Geschenk für einen/den Freund gekauft.
Peter has the present for a$_{_ACC}$/the$_{_ACC}$ friend$_{_ACC}$ bought.
Peter bought the present for a friend

b. *Mit was für einen Typen ist die unterwegs?
With what for a$_{_ACC}$ guy$_{_ACC}$ is she on-her-way

b′. Mit was für einem Typen ist die unterwegs?
With what for a$_{_DAT}$ guy$_{_DAT}$ is she on-her-way

This indicates that, contrary to the (21a) context, in which the DP is the complement of *für*, receiving an accusative case from it, the lexical entry *für* merged in the *was-für* context either cannot be [$_{P\ =D}$ für] or the $=D$ selection cannot be fulfilled by the [$_{NP}$ Bücher] directly. Leu (2008) argues that (20) indicates that the [$_P$ … für] must be hosted in a functional projection above NP, but below DP (possibly right above the kind-denoting phrase à la Zamparelli 2000) where it directly selects *was* (in fact, a small clause formed by *was* and an empty N "*SORT*") that then moves to an higher DP peripheral position, as schematically depicted in (22).

(22)



This solution is compatible with another critical property of this construction: the NP in (22) triggers agreement on the verb, as shown in (23a), not *was* (23b); the English counterpart of this construction shows the same agreement pattern, (23c) (Leu 2008:4):

(23) a. Was für Schüler *hat/haben sich beklagt?
what for pupils has/have refl complained
b. Was ist/*sind zerbrochen?
what is/are broken
c. What kind of books *is/are lying on the table?

Further evidence that *was* must have been sub-extracted from an internal DP/PP position to the higher left peripheral position is suggested by its sensitivity to island constraints: the split is not possible with the subject, (24a-b), but it is with the direct object (25a-b), as discussed by Den Besten (reported in Safir 1985)

(24) German, Safir 1985:209
a. Was für Leute haben dir geholfen
What for people have you_{dat} helped
What kind of people helped you.
b. *Was haben für Leute dir geholfen
What have for people you_{dat} helped

(25) a. Was für Museen hast Du in Italien besucht?
What of museums have you_{nom} in Italy visited
What kind of museums have you visited in Italy.
b. Was hast für Museen Du in Italien besucht?
What have of museums you_{nom} in Italy visited

Under this analysis, the split can be explained by the *forP* movement to a functional position in the matrix clause, followed by the movement of the remnant DP (i.e. $[_{whP}$ what _]). Under the definitions proposed in §2, the lexicon deriving first (22), then the split illustrated in (4b), is indicated in (26) and the relevant part of the derivation is expanded in (27).

(26) $[_{D\ wh}$ was], $[_{D\ =D}$ SORT], $[_{+wh\ \ P\ =N\ \ =D}$ für], $[_{N}$ Bücher],
$[_{+P\ F}$ Ø], $[_{+wh\ C}$ Ø]

(27) i. Merge(was, SORT) = $[_{D\ \cancel{=D}}$ SORT $[_{D\ wh}$ was]]

ii. Merge (für, [was…]) = $[_{+wh\ \ P\ =N\ \cancel{=D}}$ für $[_{D\ wh\ \cancel{=N}}$ was $[_{N}$ SORT]]]

iii. Move ($[_{+wh}$ $[_{wh}$ …]) = $[_{P}$ $[_{D\ wh\ \cancel{=N}}$ was $[_{\cancel{+wh}\ P\ =N\ \cancel{=D}}$ für $[\ _{\_D}$ $[_{N}$ SORT]]]]

iv. Merge ([für…], Bücher) = [was [[$_{P\ \cancel{=N}}$ für …] $[_{N}$ Bücher]]]

v. Merge $[_{+P\ F}$ Ø] in the matrix spine

vi. Move $[_{+P}$ $[_{P}$ …] = [[$_{P}$ für … $[_{N}$ Bücher]] $[_{\cancel{+P}\ F}$ Ø … [… $_{wh}$ was $_{\_P}$]]]

vii. Merge $[_{+wh\ C}$ Ø] in the matrix spine

viii. Move $[_{+wh}$ $[_{wh}$ …] = [[… $_{wh}$ was $_{\_P}$] $[_{\cancel{+wh}\ C}$ Ø … [[$_{P}$ für … $[_{N}$ Bücher]] $[_{\cancel{+P}\ F}$ Ø … $_{\_wh}$]]]

The derivation (27), in principle, correctly derives the predicted structure, but there are critical steps that need to be highlighted: first of all, the definition of Merge and Labeling (Definition 2 and 3) does not allow for the creation of a Small Clause as an unlabeled structure; for this reason the step (27i) and the lexical entries in (26) $[_{D\ wh}$ was] and $[_{D\ =D}$ SORT] make a necessary asymmetrical stipulation: either *was* must select or license SORT or the way around. Assuming that the phonetically empty item SORT selects *was*, we induce an expected head-complement asymmetry, enabling then possible *was* movement in total compliance with the Don't-LYC constraint[7]. This is clearly a convenient (but arbitrary) solution that could be avoided by enabling a purely symmetric Merge operation[8].

A second critical aspect is related to linearization: notice that with our purely asymmetrical definition of Merge, we are already able to subsume an equivalent versions of the Linear Correspondence Axiom (Kayne 1994):

Definition 9. *Linearization Principle based on Merge*
The selecting item always precedes the selected one

This definition could suffice to derive the (pre-movement) linear ordering without further externalization assumptions. According to the definition of Probe and Goal (Definition 5) and derivational C-command (Definition 7), we can also extend the linearization Principle to the moved constituents.

[7]Nothing here precludes the possibility of considering an inverse selection or a different formation of the SC. These other options require extra stipulations. We adopt this solution here for sake of simplicity.

[8]See Cecchetto & Donati (2010) for a proposal on labeling of small clauses.

Definition 10. *Linearization Principle based on Movement*
The licensed item always precedes the licensing one.

To solve the well know "linearization paradox"[9] we must assume that linearization is resolved incrementally and that "movement" leaves linearized traces. The consequences of this (common) assumption will not be discussed here (see Chesi 2012). We will see in §3 that the conflict between Definition 9 and 10 (in select-driven merge, the projecting item precedes the selected one, while in the licensing-driven merge, the projecting item follows the licensed one) can be resolved.

The last critical step in this analysis concerns the criterial positions: the minimalist system we have introduced so far is not able to distinguish intermediate and criterial positions from each other, possibly letting the *wh-* item stop at an intermediate, non-criterial, +*wh* licensed position (missing a criterion satisfaction) or moving it beyond the criterial position (violating criterial freezing, Rizzi 2004). Moreover no constraints on extraction (apart from locality) are formulated so far. We will address the "mainstream" solution of these issues in section §2.2 and we will provide our alternative view in §3.

2.1.3 Intermediate summary

The evidence discussed in §2.1 suggests that the Don't-LYC constraint is always respected, and that apparent counter-examples (e.g. *was-für* split) are consistent with it, given an appropriate, independently motivated, structural analysis. The universal hierarchy of functional projections predicted by cartographic investigations (Cinque & Rizzi 2008) seems, furthermore, to create the correct scaffolding for all possible rearrangements without violating the Don't-LYC constraint. An asymmetric prediction in terms of linearization can be deduced from the definitions in §2, which is compatible with cases discussed in this section. A more restrictive implementation of the movement operation is required to account for criterial freezing and proper constraints on extraction.

2.2 Evidence in favour of the BA constraint and (apparent) counterexamples

The Be Accessible (BA) constraint, defined in (8), states that the Goal feature must be at the edge of the moved constituent. This constraint is very similar to the Edge generalization proposed by Heck (2009:89):

[9] Probing a selected item leads to a linear paradox: α, first merged by selection, then re-merged as licensed is expected both to be linearly "after" the selecting head β and "before" the licensing head δ: $\alpha < \delta < \beta < \alpha$.

(28) **Edge Generalization**
If α can pied-pipe β, then α must be at the edge of β

We need an extra generalization on the recursive pied-piping behavior ((29), Heck 2009:85) to derive the facts presented in (11) and (12), repeated below for convenience, which are also compliant with BA.

(29) **Generalization on recursive pied-piping**
If α can pied-pipe β and β is in a canonical position to pied-pipe γ, then α can pied-pipe γ

(11) a. a person [[[***whose*** [lawyer's [sister's … [son]] adores you
b. *a person [the son … of the lawyer of the sister **of who**] adores you

(12) a. [[[*Mitä* kaupunkia]$_i$ kohti$_{_i}$]$_j$ virtaamalla$_{_j}$]$_k$ Seine
what city towards by.flowing Seine
pääsee valtamereen$_{_k}$?
reaches ocean
b. *Seine pääsee valtamereen [virtaamalla [kohti [*mitä*
Seine reaches ocean by.flowing towards what
kaupunkia]]]?
city
By flowing towards which city does Seine reach the ocean?

Contrary to Don't-LYC constraint, these generalizations, as well as the BA constraint, are not directly deducible from the definitions in §2: the relevant *wh-* item is "at the edge" of the structure both in (11a) and (12a), but it can be deeply embedded within the DP of which it is a proper argument.

The full-fledge theory of pied-piping developed by Heck (2009) is based on a minimalist implementation of the Probe-Goal dependency that hinges on the fact that such dependency can be satisfied also within a non-optimal configuration: first, a Probe-Goal dependency is satisfied not directly through movement, but Agree (Definition 11), second, Agree must be "local" (Definition 12).

Definition 11. *Agree (Heck 2009:78)*
A Probe P can establish *Agree* with Goal G if and only iff (a) and (b) hold:
(a) P C-commands G
(b) there is no Goal G_i such that P c-commands G_i and G_i c-commands G

Definition 12. *Local Agree (LA) (adapted from Heck 2009:80)*
No phrasal boundaries should occur between *P* and *G*

Being the most local agree configuration the spec-head relation, when the Goal moves to the SPEC position, the LA condition is satisfied. Assuming, as Heck does, that LA is in fact a "violable" constraint, it follows that if a boundary intervenes between P and G, the sentence will not be ruled out automatically. Ranking all possible alternatives produced by different possible displacement operations, compliant with these and other operative principles (à la Prince & Smolensky 2004), the most optimal configuration will be the one, for instance, reducing the number of boundaries between P and G. As proposed by Heck, we might add one penalty to each boundary crossed for establishing an Agree relation; in this way we can deduce a preference principle for preposition stranding over preposition pied piping ((3a), most local spec-head configuration, vs (3b) repeated, an extra PP boundary between the Goal and the Probe below):

(3) a. $[_{wh}$ *which* photographer$]_i$ C_{wh} did you buy $[_{\mathrm{DP}}$ pictures $[_{\mathrm{PP}}$ of $\__i$]]?
b.?* $[_{\mathrm{PP}}$ of $[_{wh}$ *which* photographer]]$_i$ C_{wh} did you buy $[_{\mathrm{DP}}$ pictures $\__i$]?

The solution is attractive for its simplicity[10]. Other ingredients necessary for predicting the correct ranking, and then inferring the correct empirical generalizations, are the notion of Phase Impenetrability Condition, Numeration and the idea of Last Resort (another violable constraint)[11]:

Definition 13. *Phase Impenetrability Condition (PIC) (adapted from Chomsky 2001:14)*
The domain of a head *H* of a phase *HP* is not accessible to operations outside *HP*. Only *H* and its edge domain are accessible.

Definition 14. *Numeration (Chomsky 1995:225)*
a set of pairs (*LI*, *i*), where *LI* is an item of the Lexicon and *i* is its index, understood to be the number of times that *LI* is selected.

Definition 15. *Last Resort (adapted from Chomsky 1995:128)*
If a Goal *G* moves within a phase, it must check a *Probe* feature.

[10]Maybe less for its psycholinguistic implausibility: shall the speaker compare all possible alternatives before going for the best solution?

[11]Other definitions are necessary to fully introduce the phase idea. The reader is invited to check Heck's 2009 original formulations for the notions of Accessibility, Edge domain etc.

Assuming that phases are *CP* and *vP* (hence *C* and *v* are phase heads, Chomsky 2000, 2001) and that the phase-based numeration (Heck must assume that each phase has its own (sub)numeration) is responsible for introducing a relevant Probe, if any, Pied-piping, under the violable constraint approach, becomes a "repairing strategy":

(30) **Repair generalization** (Heck 2009):92
Pied-piping of β by α is possible only if movement of α from β is impossible.

This explains, for example, why pied-piping is obligatory when the *wh-* item is contained in an island (or whether a relevant Probe is not included in the phase numeration). The notion of "islands" must still be defined, as well as the Left Branch Condition and other restrictions blocking the dependency between a Probe and its Goal.

In sum, the nature of the BA constraint remains ancillary with respect to the Local Agree relation which is, in turn, non-primitive, and optionally violable, rescuing long-distance dependencies that are otherwise impossible. Notice that at least the asymmetry present in English and discussed in (7) (repeated below for convenience) suggests that optionality should be restricted in some way to result in an empirically adequate formalization:

(7) a. the person [of whom]$_i$ [pictures $_{\_i}$] are on the table (Chomsky 1986)

b. *the man who$_i$ [pictures [of $_{\_i}$]] are on the table

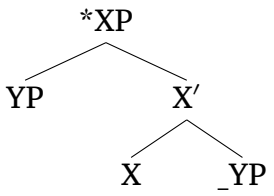
2.2.1 On preposition stranding

Preposition pied-piping vs. stranding is another piece in the puzzle, exhibiting both cross-linguistic (e.g. (3) vs (6)) and intra-linguistic (e.g. (3) vs (7)) variation. An interesting proposal for accommodating some of the critical facts is developed by Abels (2012). Abels assumes that Merge becomes possible if and only if at least one feature gets shared between the two elements merged. "Sharing", in his system, is a necessary (but not sufficient) condition for feature valuing. Features that are valued [F] and those that must be valued [uF] must occur in an appropriate "valuation domain" before being delivered at the interfaces. Each unvalued feature can be specified for the conditions under which it can be shared: [uF↓] indicating that [F] must be searched in the C-commanded domain of the head bearing the [uF↓] feature; [uF↑] indicating that the relevant [F] feature should C-command the [uF↑] bearing item; [uF↓↑] requiring

that [F] both C-commands and is C-commanded by [uF↓↑]. [uF]s, once specified for one of the possible sharing options, become Probes.

Assuming the anti-locality constraint in (31) (i.e. a ban on remerging the same item twice with the same category, see §2.2.1) and PIC (Definition 13), the stranding generalization in (32) follows:

(31) Anti-locality constraint



(32) Stranding generalization: Given a phase H and a constituent X in H's C-command domain, the following generalizations are, respectively,
   a. possible to derive: [X … [H [… $_X$ …]] …] and
   b. impossible to derive: [X … [H $_X$] …]

Given a list of phase heads, P, D, v and C (Abels 2012:122) and the reasonable assumption that they are universal, preposition pied-piping should be the default, while the "parameterized option" of P-stranding should be available only if the P phase head and its (apparent) DP complement is intervened by a phrase. Abels convincingly argues in favor of the presence of such a phrase (and its absence in language, like Italian or French, where the P + D incorporation is possible) based on the fact that, also in some German dialect, P-stranding is possible if the preposition shows the morphological incorporation of a clitic R-word "DR" (Abels 2012: 233):

(33) Wo hast du { *in | drin } geschlafen?
   where have you in | DR.in slept
   What did you sleep in?

The DR phrase (the reader should refer to Abel's original work for useful details here unnecessary) crucially must be a complement of P and, when present, it should license (islands constraints permitting) P-stranding. The generalization is then clear:

(34) a. [$_D$ Who] did he talk [$_P$ to [$_{DR}$ DR Ø $_D$] ? (stranding)
   b. [$_P$ A [$_D$ qui]] a-t-il parlé $_P$? (non stranding)

A challenge to this analysis comes from Finnish, where there is no evidence for a functional projection between P and DP, yet the DP moves to Spec,PP:

(35) a. lähellä minu-a
near I-par

b. minu-a lähellä _
I-par near

c. minun lähellä-ni _
I-gen near-1s
near me

If the DP has a *wh* feature, movement is obligatory (Huhmarniemi 2012):

(36) a. *lähellä ketä
near who.par

b. ketä lähellä _
who.par near

c. kenen lähellä _
who.gen near
near who

Movement of the partitive-marked DP complement, (35b), (36b), constitutes regular A-bar movement from COMP to SPEC (Huhmarniemi & Brattico 2013), whereas the COMP-to-SPEC movement of the genitive DP, which also triggers phi-agreement with the preposition, (35c), is interpreted as A-movement triggered by the EPP feature at P (Manninen 2003, Brattico 2011). In these cases, there seems to be no clear evidence, morphological or syntactic, for the existence of any extra projection within the PP though Finnish allows for P-stranding.

2.2.2 Intermediate summary

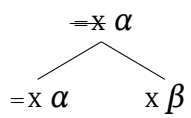
The evidence discussed in §2.2 suggests that BA is not a primitive constraint and cannot be derived from the definitions presented in §2. At least two notions must reconsidered to derive BA: the notion of *phase impenetrability condition* and a refined Probe-Goal condition. The second modification requires the definition of an *Agree* relation and a precise specification of the "agreement" and/or "sharing" domain. The necessity of establishing a Probe-Goal dependency despite some adverse configuration (e.g. the Goal is within an island) seems to justify pied-piping.

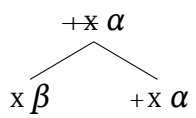
3 Deriving BA and Don't-LYC by reversing structure building

The aim of this section is to propose a simplification of the Probe-Goal dependency in order to derive both BA and the Don't-LYC constraints as byproducts of the derivation, while retaining, at the same time, the empirical generalizations discussed in the previous sections. We argue that this can be achieved by reversing the derivational procedure itself.

3.1 A natural asymmetry in Merge

The first operation we reconsider is Merge. Merge must involve some asymmetry in order to select and project the label. This is the pivot of our simplification strategy: in both selection and licensing, the selecting/licensing item projects over the licensed/selected one:

(37) a. 

b. 

From the top-down perspective (i.e. starting from the root of the tree structure), the necessity of external linearization conditions (Definition 9 and 10) disappears: selecting/licensing items should always be processed before the selected/licensed one, according to (37). We could then simplify Merge by assuming that this is a binary operation between an "expectation" (select feature in Definition 1) and a lexical item:

Definition 16. *Merge (re-defined top-down)*
Merge is a binary operation taking in input an expectation ($=X$, or $+X$ as in $[_{=X}\ \alpha]$ or $[_{+X}\ \alpha]$) and a compatible lexical item (e.g. $[_X\ \beta]$), returning the lexicalized structure headed by the expecting item (e.g. Merge($[_{=X}\ \alpha]$, $[_X\ \beta]$) = $[\alpha\ [_X\ \beta]]$)

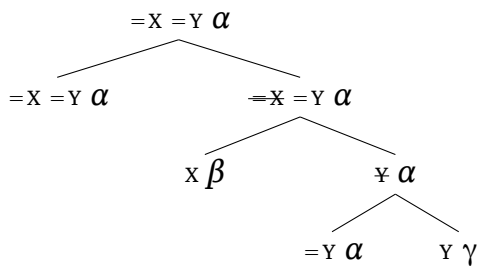
An interesting fact about language is that linear order and dominance do not always overlap[12]. This is only possible, under Definition 16, if we assume that an item α can project two expectations: $[_{=X\ =Y}\ \alpha]$. When expectations are ordered[13], the relevant configuration can be created by

[12]Within the bare-phrase structure we adopt here (labels are nothing but lexical items, in line with Collins 2002), precedence is a total order, that is, it can be defined among any item introduced (or re-introduced) in the derivation, while dominance is a partial order.

[13]Having an ordered numeric restriction on expectations is not an innocent assumption. Here we greatly simplify the discussion: considering $[_{=\mathrm{V}}\ \mathrm{v}]$ (*little v* selecting V) and V, *little v* will be lexicalized first with a lexical predicate selecting for an external

(at least) two merge operations (i. Merge([$_{=\boldsymbol{X}\ =Y}$ α], [$_X$ β]) = [$_{=Y}$ α [$_X$ β]]; ii. Merge [[$_{=\boldsymbol{Y}}$ α [$_X$ β]], [$_Y$ γ]) = [α [$_X$ β] [α [$_Y$ γ]]]), leading to this final structure:

(38)



The tree in (38) is, in fact, a representation of the history of the derivation. The α spine is the head of the constituent; *C-command* can be redefined as follows:

Definition 17. *C-command (re-defined top-down)*
α c-commands β and any items β will merge with iff α directly merges with β

Under this definition, in the derivation (38) α c-commands β, α c-commands γ but also β c-commands γ (since β merged with α and, later, α merged with γ), while γ does not c-commands β (since γ merged with α, but later α does not merge with anything else; being β merged with α *before* α merged with γ, γ cannot c-commands β; this emphasize the importance of the order of application of each Merge operation). It is not an accident that α, β and γ behave, respectively, as a predicative head (α), the external argument (β) and the internal one (γ) or, as reworded in the X-bar terms, head (α), spec (β) and comp (γ)[14]. It is also useful to highlight that feature sharing/valuation (Abels 2012, §2.2.1) should be restricted by the ordering of the operations (i.e. β can share/value features with α - e.g. SV agreement -, then α can share feature with γ, if

argument (i.e. a DP), while V will be lexicalized with a lexical predicate (possibly the same one lexicalizing *little v*) selecting the internal argument (another DP). By hypothesis, both the default category and the lexical item can have at most one select feature; this leads to the conclusion that we can have at most two select features per category: when default categorial selection is present, the lexical predicate selection merged induces binary branching and it gets expanded before the default categorial selection (i.e. =V of *little v*). On the other hand, the default categorial expansion creates the right branch. If such default categorial expansion is absent, right branching can only be created by the select feature on the lexical item.

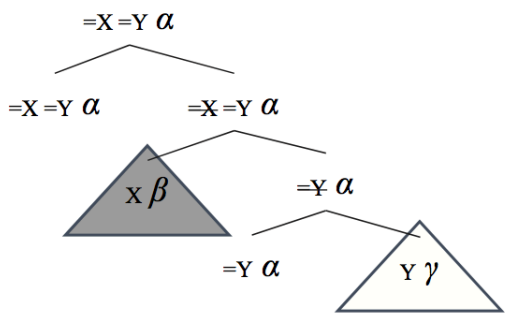
[14]We do not discuss here the fact that the two positions would be possibly available/parametrized for spelling out the selecting head (eventually producing the linear order $<\alpha, \beta, \gamma>$ or $<\beta, \alpha, \gamma>$). By default we assume that the "first merge" position, i.e. the higher one, is the one pronounced.

needed and possible: α might have some feature already valued at some point, this might explain why pre-verbal merged items, e.g. object clitics, can trigger agreement while post-verbal merged one, e.g. post-verbal full objects, cannot in Romance languages).

3.2 Phases: the domain of an expectation

The assumption that one item can project (at most, see note 13) two ordered expectations introduces the possibility of a complex left branching structure: in the example (38), for example, β could have had other selection features, projecting extra expectations and then forming a complex left branched constituent. There is a natural computational distinction we can make between the first and the last expectation: the first expectation (=X in (38)) will be processed while the item α is still active (because of the pending expectation =Y); the satisfaction of the last merge (=Y) is, on the other hand, the last operation triggered by α. We assume that a *phase* is the domain of an expectation, and that the first and the last expectations are, respectively, a nested (gray triangle in (38)) and a sequential phase (white triangle in (38), Bianchi & Chesi 2006).

(38)



Definition 18. *Phases (re-defined top-down)*
A phase is the domain of an expectation, i.e. the set of merge operations fully lexicalizing an expectation generated by a select/licensor feature.

- The last expectation constitutes a *sequential* phase.
- Any other expectation is a *nested* phase.

The nesting vs. sequential distinction has a computational / psycholinguistic impact (see Chesi 2015) that we will use to restrict the Probe-Goal dependency domain (§3.3) and derive the pied-piping constraints (§3.4).

Before going into the Probe-Goal dependency, we should observe another distinction that follows from the top-down derivation: the distinc-

tion between functional and lexical items. According to the cartographic approach (that we assume here), functional categories are universally ordered (Cinque 1999). Each functional projection then hosts the same unique selection requirement. For example, a declarative C should select a criterial subject position S, the criterial subject position S should select a finite T, a finite T should select a VP. Being universal means that there is no freedom in selection, but each language can employ phonetically empty realizations of these positions, or morphological clusters incorporating more features in the same head. What is not possible is to find either recursion in these positions[15] or licensing of an extra expectation other than the default (cartographic) one. Without lexical items (A, N, V and, possibly, some P) there is no way of projecting a nested phase along the rigid spine[16]. Lexical items, on the other hand, can introduce this "extra expectation" and, crucially, true recursion, that is, they can project an expectation for the sequentially highest functional category (e.g. CP or P/DP).

3.3 Move unexpected features

We will now rephrase the Probe-Goal dependency (1) in top-down terms. Here we simply assume that what is unexpected must be moved, that is, it must be re-merged in a proper expected/selected position. A compact version of this definition is given in 19. This includes i. a definition of memory, ii. the idea of unexpected categories triggering movement, iii-iv a storage and retrieval mechanism permitting non-local re-merge:

Definition 19. *Move (re-defined top-down)*

i. a memory buffer (*M-buffer*) is a phase-related structured repository used to store unexpected items;
ii. an item is *unexpected* if it brings into the structure categorial features that are not selected;
iii. if an item α is merged with unexpected features Y, α is moved into the M-buffer with its unexpected features $[_{\mathrm{Y}}\ \alpha]$;
iv. an item in the M-buffer with unexpected Y features will be re-merged in the structure, before any other item from the lexicon, as soon as a selecting item (e.g. $[_{=Y}\ \beta]$) is merged, introducing the relevant Y features as expectation; once re-merged, the item $[_{\mathrm{Y}}\ \alpha]$ is removed from the M-buffer.

The M-buffer is a metaphor for explaining the non-local dependency between a criterial and a selected position the very same item should

[15]Left-peripheral topic positions (Rizzi 1997) are somehow special in this respect and need some exceptional discussion.

[16]Same consideration extends to cartographic DP and CP split.

occupy. An unselected item can be licensed in its first merge position to fulfill a licensor feature (e.g. $+Y$) or "pied-piped" by another categorial expectation (if $=X$ is expected and $[_{X\,Y}\ \alpha]$ is merged, then Y categorial feature results unselected). In both cases, categorial features (i.e. Y, possibly qualifying this item as a good argument, e.g. D) are unexpected in that "criterial" position. "Moving it in memory" indicates that the item is flagged for an unsatisfied selection dependency and, unless a proper selecting feature is introduced later in the derivation, the sentence will be ungrammatical. The phase-locality of the M-buffer would prevent any item stored in a specific phase to be used, by default, in any other phase. We will maintain this assumption for the nested phases (i.e. items in the matrix memory buffer cannot be selected and discharged within a nested phase), but we assume that unexpected items are inherited from the matrix clause by the sequential phase, i.e. through the main spine of the tree. This inheritance mechanism technically closes the previous phase: no other operation can target items contained in that phase, a part from the one(s) transmitted through the M-buffer to the sequential phase.

Definition 20. *Phase closure*
Once the last selection is projected, the phase is closed and the content of the M-buffer transmitted to the M-buffer of the sequential phase.

3.4 Deriving pied-piping

An English (toy) grammar presented in (39) will be used to illustrate the approach. We distinguish between the Lexicon[17], default Categories (that are targeted by selection) and a set of Parameters (parametrized features associated to specific categories) in English (triggering overt wh-movement, aux-to-comp and subject movement respectively):

[17] Since no morphological decomposition is addressed here, the lexicon expresses minimal derivational chunks distinguishing among determiners, pronominal forms, proper names (Elbourne 2005) and predicate distinct selections (see note 13). Morphological features expressing relevant constraints are added, for convenience, under brackets (see note 3).

(39) Lexicon:

$[[_{D}\ \text{John}_{i}]\ [_{N}\ \__{i}]]$, $[[_{\text{wh D}}\ \text{what}\ [_{N}\ e]]$, $[_{D}\ \text{you}\ [_{N}\ e\ ]]$, $[_{N}\ \text{book}]$, $[_{D}\ \text{the}]$, $[_{T}\ \text{did}]$, $[_{P}\ \text{of}]$, $[_{v\ =[D(case:nom)]}\ \text{buy}\ [_{V\ =D(case:acc)}]]$ …

Categories:

$[_{=\mathbf{S\_topic}}\ \text{Root}_{C}]$, $[_{=\mathbf{C\_fin}}\ \text{Root}_{wh}]$,
$[_{=\mathbf{T\_fin}}\ S_{topic}]$, $[_{=\mathbf{V}}\ T_{fin}]$, $[_{=\mathbf{S\_topic}}\ C_{fin}]$
$[_{=\mathbf{D}}\ P][_{=\mathbf{N}}\ D]$

Parameters (English):

$[_{=\mathbf{wh}}\ \text{Root}_{wh}]$, $[_{+\mathbf{T}}\ C_{fin}]$, $[_{+\mathbf{D}}\ S_{topic}]$

The derivation starts with a Root (default) expectation (either $[_{=\mathbf{S\_topic}}\ \text{Root}_{C}]$ or $[_{=\mathbf{C\_fin}}\ \text{Root}_{wh}]$ are available (step i) then Merge applies (step ii), trying to fulfill an expectation (*wh*); this expectation indicates that a *wh*- item is selected, but also a D feature is introduced which results unselected, hence the item bearing it must be stored in memory (step iii); the default select features associated to Root_{wh} ($[_{=\mathbf{C\_fin}}\ \text{Root}_{wh}]$) trigger the following expectation (C_fin) and so on, until the final expectation of V (*buy*) is projected, closing the matrix phase.

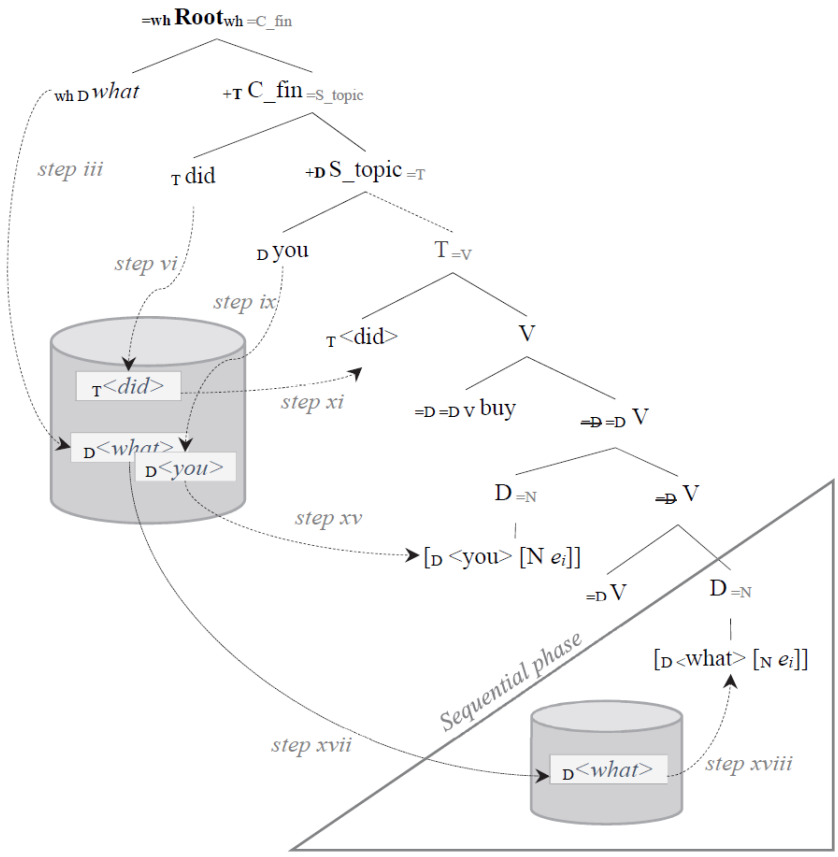
(40) $[[_{wh}\ \textit{what}]_{i}\ C_{wh}]$ did you buy $\__{i}$?

i. **Expect**(Root_{wh})
ii. **Merge**($_{\mathbf{wh}}$ Root, $_{\mathbf{wh\ D}}$ *what*)
iii. **Move**($_{\mathbf{D}}$ <*what*>)
iv. **Expect**(C_{fin})
v. **Merge**(C_{fin}, *did*)
vi. **Move**(*did*)[18]
vii. **Expect**(S_{topic})
viii. **Merge**(S_{topic}, *you*)
ix. **Move**(*you*)
x. **Expect**(T)
xi. **Merge**(T, <*did*>)
xii. **Expect**(V)
xiii. **Merge**(V, *buy*)
xiv. **Expect**(D)
xv. **Merge** (D, <*you*>)
xvi. **Expect**(D)
xvii. Close the root phase and discharge the memory buffer into the sequential phase
xviii. **Merge** (D, <*what*>)

[18]By Definition 19 both licensors +X (as in this case, +T) and unselected features trigger movement.

The tree diagram (41) summarizes the steps (bold features, +**X** are parametrized features, grey features, =**Y**, are default cartographic expectations):

(41)



The first thing to notice is that criterial features are expressed in two ways: either we license a feature that must be selected later (+T feature) or we expect it, but movement is still triggered because the merged item is "larger than expected" ("wh D" while only "wh" was expected). This simplifies the featural inventory and constitutes the simplest possible implementation of the aux-to-comp and subject movement in (40)-(41) (theoretically, at the price of making the derivations slightly less transparent, one of the two options might be discarded, Chesi 2015).

A second consideration is that elements must be stored in the memory buffer according to their top-most categorial feature; when two items stored in memory share the same feature(s), the derivation becomes ambiguous and we might expect confusion (this has been considered the source of processing difficulty for Object Clefts, Chesi & Canal 2018). This conflict is solved by assuming a last-in-first-out memory usage: the last item stored should be the first to be retrieved, if the critical features

are shared with other items in memory.

3.4.1 Deriving BA and Don't-LYC

Under the current approach, the Don't-LYC constraint follows from the local expansion of an expectation, by the definition of Merge (Definition 16): as soon as an expectation is projected, the only way to satisfy it is by merging an appropriate item, that is, an item with the properly selected categorial feature. This explains the ungrammaticality of (3c), repeated here:

(3) c. * [$_{wh}$ which]$_i$ C_{wh} did you buy [$_i$ [$_{NP}$ book]] ?

Which, with the structure [$_{\text{wh D}}$ which $_{=\text{N}}$], should be merged to the root to satisfy the *wh* root expectation ([$_{=\mathbf{wh}}$ Root$_{\text{wh}}$]); because of D, it qualifies as an item to be moved in the M-buffer[19]; because of its selection feature =N it is also a nested phase that will be expanded while the superordinate root C is still pending with its =C_{fin} selection. When =N is projected, a compatible item must be merged. Merging "did", satisfying a C_{fin} requirement, leads to an ungrammatical sentence, since =N expectation is not fulfilled and cannot be satisfied later.

On a similar vein, *BA* is subsumed under the licensing (+X) and the selection (=X) options. Considering that X is merged to the structure to fulfill a +/=X expectation, if it happens immediately after the projection of the relevant expectation, the merged item will be "at the edge" of the structure: as in (3) above, this indicates that the item brings the prominent category of the nested phase [$_{\text{D}}$ which [$_{\text{N}}$ book]], but it might also be at the spec (of the spec ...) of it [$_{\text{D}}$ [$_{\text{P}}$ whose] books]. The second case is more problematic, because it requires a sort of "sinking" of the expected item under a not-yet merged constituent: in a nutshell, the *wh*-item must be considered as moved from within the nested constituent, but since the nested constituent must still be merged, its status must be decided either on the basis of some morphological property or we might expect reanalysis effects during parsing[20]. We believe this is the role of the genitive marker "'s" in English (*whose* is then analyzed as a syncretic *who* + *'s*) that creates an expectation for the incoming item to be a

[19]Remember that the items are unique across the whole derivation; this implies that the item stored in the M-buffer and the item that must be expanded are the same; as soon as the expansion is processed, the item in the M-buffer is also "modified" (that is, its selection requirement is expanded).

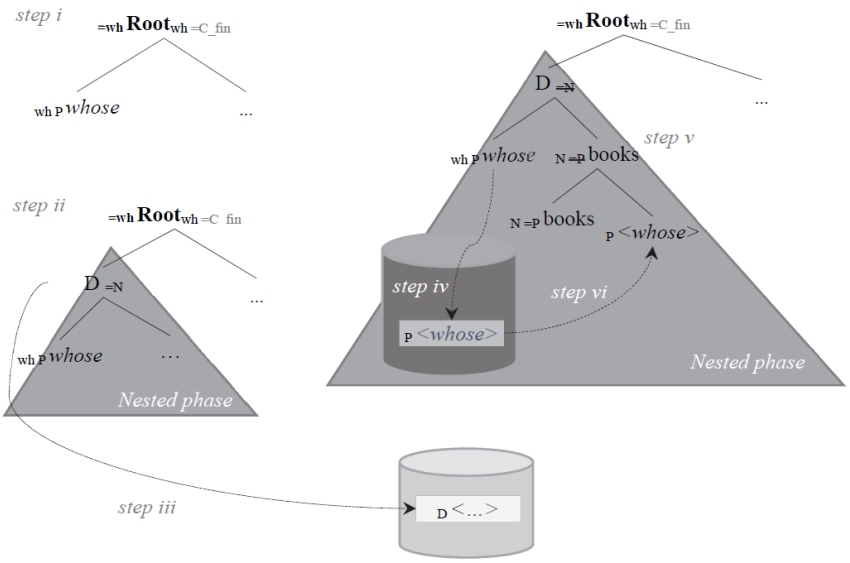
[20]An advantage of the proposed derivation is that this is transparent both with parsing and generation performance tasks. Here we mainly focus on generation, but the lack of morphological evidence for specific operations (like "sinking") might lead to difficulties in parsing that are not evident in generation (e.g. solving lexical ambiguities).

D-selected nominal predicate of which the pre-"'s" constituent is a selected complement. We call this "expectation sinking" and we define it as follows:

Definition 21. *Expectation sinking*
An item can be moved within a nested phase only if sunk within the nested phase: *sinking* embeds the expected feature within one layer whose category is predicted by the morphological features shared between the sinking item and the embedding one.

In the specific case "whose books …", illustrated in (42), [$_{wh\ \mathrm{P}}$ whose] is first merged in *wh* criterial position (or *R* for restrictive relative clause, see note 22), step i, hence becoming [$_{\cancel{wh}\ \mathrm{P}}$ whose]; then it sinks under a D layer because of the "'s" morphology ([$_{\mathrm{D}}$ [$_{\mathrm{P}}$ whose] (N)], step ii). This operation does not involve the *wh* feature that, once merged, already fulfilled the *wh* expectation then, as usual, gets deleted. The item to be moved in the M-buffer of the matrix phase will become a category D item (step iii), while the "secondary" movement within the nested phase, using its own M-buffer, will consider a category P item (step iv). This P item should be discharged after a properly selecting item, [$_{\mathrm{N}\ =\mathrm{P}}$ books], will be merged (step v) in the nested phase, then creating the relevant expectation for the pending P item (step vi).

(42)



3.4.2 Snowball movement

The sinking (Definition 21) can be considered as an alternative to "percolation": from this perspective, there is no need to percolate features,

as the relevant feature at the edge is merged as expected, while the item sinks later into the embedded structure[21]. This also derives the recursive pied-piping behavior exemplified both by genitive recursive pied-piping English (11) and *wh*-snowball movement in the Finnish construction (12).

(12) a. [[[*Mitä* kaupunkia]$_i$ kohti$_{_i}$]$_j$ virtaamalla$_{_j}$]$_k$ Seine
what city towards by.flowing Seine
pääsee valtamereen$_{_k}$?
reaches ocean

Assuming the lexicon in (43), the derivation unfolds as indicated below:

(43) Lexicon:
[$_{\text{wh D =N}}$ Mitä], [$_{\text{N}}$ kaupunkia], [$_{\text{P =D}}$ kohti], [$_{\text{P V =P}}$ virtaamalla]
[$_{\text{D}}$ Seine], [$_{\text{v =D}}$ pääsee [$_{\text{V =D}}$]], [$_{\text{D}}$ valtamereen]
(same English categories and parameterization)
i. **Expect**(Root$_{\text{wh}}$)
ii. **Merge**($_{\textbf{wh}}$ Root, $_{\textbf{wh D}}$ *Mitä*)
i. **Expect (N)**
ii. **Merge** (**N**, *kaupunkia*)
i. **Sink** *[$_D$ Mitä kaupunkia]* $_{+D}$
ii. **Move**(**D**)
iii. **Merge**([$_{\textbf{+D P =D}}$ *kohti*]);
iv. **Expect (D)**
v. **Merge**($_{\textbf{=D}}$ *Kohti* , $_{\boldsymbol{D}}$ <*Mitä*>)
i. **Sink** *[$_P$ [Mitä kaupunkia] kohti]* $_{+P}$
ii. **Move**(**P**)
iii. **Merge**([$_{\textbf{+P P =P}}$ *virtaamalla*])
iv. **Expect (P);**
v. **Merge**($_{\textbf{=P}}$ *virtaamalla*, $_{\boldsymbol{P}}$ <*kohti*>)
iii. **Move**($_P$ < *[[[Mitä … virtaamalla]*>)
iv. **…**

While the above derivation illustrates the general computational involved in the derivation of the Finnish secondary wh-movement, it is not yet sufficient to capture the phenomenon in its entirety. One additional fact we need to consider is that many of the specifier positions

[21] This solution shares some similarity with Cable's 2010 intuition that a Q-morpheme is always present over the top of the moved constituent: from this point of view, the Q-morpheme is nothing but the expected *wh* feature and it is merged at the root of the nested constituent before sinking.

targeted by secondary *wh*-movement, as in (12), are unselective with respect to the type of phrase that can occur in that position. This can be modeled by assuming that, instead of selecting for a *wh* feature (by the interrogative root), another functional layer selects for a (contrastive) *Focus* or *Topic* feature associated to the item moved to the spec. The derivation will then be totally coherent with the one illustrated in (43). A related problem is that without the presence of the *wh*-feature, movement to the specifier position is often optional. If optionality is real, and no scope/discourse effects are induced, the (contrastive) topic/focus featural approach is not a solution: this, in fact, should predict fixed positions (different from the *wh*- one) to be targeted once the secondary movement is present and the absence of such a topic/focus induced movement (and any scope/discourse effect) when secondary movement is absent. Whether the predicted topic/focus interpretations arise in connection with secondary A-bar movement remains to be studied. An alternative view is proposed by Huhmarniemi & Brattico (2013): instead of selecting for a specific label, the P/Adverbial heads are endowed with an unselective EPP feature (for a top-down implementation of this hypothesis and the derivation of all Finnish pied-piping data, see Brattico & Chesi, submitted).

3.4.3 Was-für split

The expect-then-merge mechanism predicts no discontinuities between the phase projection and its satisfaction. Theoretically, we can distinguish between two sub-steps in the projection and satisfaction procedure: first, there is the projection of an expectation that we can dub "expansion", then there is the real merge, possibly anticipated by a search for a compatible item to be merged, first in memory, then in the lexicon. We might consider the option that the select feature could either be expanded as soon as a lexical item is first merged, or after it has been remerged. In fact, independent empirical reasons (i.e. reconstruction phenomena, Bianchi & Chesi 2014, extraposition, Chesi 2013) suggest that, in specific cases, the projection of an expectation can be procrastinated until the item has been remerged a second time. This will be the key of our analysis of the was-für split and p-stranding.

Definition 22. *Delayed phase expansion*
The selection feature of an item α can be processed after α has been remerged in the structure. If α becomes, once re-merged, the sequential phase, and its select feature(s) are not yet expanded, at this point it can inherit the M-buffer of the matrix phase.

The "delayed phase expansion" is not free and often optional, but

it is linked to a delayed interpretation of the constituent under analysis (Barker 2009). We disregard here the semantic reasons forcing a delayed interpretation, but we assume that: i. this is what happen in the was-für split; ii. preposition stranding is the morphological evidence for a local evaluation, hence when p-stranding is present, no delayed phase expansion is possible (§3.4.4).

This mechanism is responsible for the (apparent) "sub-extraction" from a complement of a predicate that, after remerge, becomes sequential, as in (44) (see Bianchi & Chesi 2014 for other overt violations of the subject island constraint compatible with this intuition):

(44) a. [Of which masterpiece]$_i$ is [one reproduction $_{\_i}$]$_j$ already [available $_{\_j}$]

b.?* [Which masterpiece]$_i$ is [one reproduction of $_{\_i}$]$_j$ already [available $_{\_j}$]

The principal steps of the derivation (44a) proceed as follows ("process" is the term used for indicating the evaluation of relevant select feature(s) and the related merge steps, "re-merge" indicates the dischargement of an item from the M-buffer):

i. **process** [$_{PP}$ of which masterpiece] and **move** it in the M-buffer
ii. **process** [$_D$ one [$_{N\ =PP}$ reproducion]] and **move** it BEFORE **expanding** the PP expectation
iii. **process** [$_{A\ =DP}$ available] and expect a DP as a sequential phase
iv. **re-merge** [$_D$ one [$_{N\ =PP}$ reproducion]]
v. **expect** $_{=PP}$ as a sequential phase
vi. **re-merge** [$_{PP}$ of which masterpiece]

Similarly, assuming a "remnant" analysis (Abels 2003, Leu 2008) for (4b) here repeated (see the assumed structure in (22) obtained under reconstruction):

(4) b. *Was*$_i$ hast Du [$_{\_i}$ für Bücher]$_j$ gelesen $_{\_j}$?
what have you for books read
'What kind of books did you read?'

i. **merge** [$_{wh\ D}$ *was*] and **move** it in the M-buffer
ii. **merge** [$_T$ hast] and **move** it (see aux-to comp, in (40)-(41))
iii. **merge** [$_D$ Du] and **move** it (see subject movement, in (40)-(41))
iv. **merge** [$_{P\ =SC}$ für] and **move** it in the M-buffer without expanding it
v. **merge** [$_N$ Bücher] and **move** it in the M-buffer

vi. …
vii. **merge** [$_{v\ =D}$ gelesen [$_{V\ =D}$]]
viii. **re-merge** the first available D item [$_{D}$ Du]
ix. **expand** the last D expectation, then **transfer** the content of the M-buffer into the M-buffer of this sequential phase
x. **merge** [$_{D}$ *was*]
xi. **sink** [$_{D}$ *was*] under the FP [$_{F\ =NP}$] and **move** it in the nested M-buffer
xii. **sink** [$_{D}$ *was*] under the forP
xiii. **merge** [$_{P\ =SC}$ für]
xiv. **process** $_{=SC}$ and **expand** the SORT "small clause"
xv. **merge** the empty [$_{N\ =D}$ SORT]
xvi. **process** $_{=D}$ and **expand** a D expectation
xvii. **merge** [$_{D}$ *was*] and remove it from the M-buffer (this close the nested forP phrase)
xviii. **process** $_{=NP}$ of the superordinate [$_{F\ =NP}$] phase
xix. **merge** [$_{NP}$ Bücher]

3.4.4 Revisiting preposition pied-piping

In the previous chapter we assumed that P-stranding signals the fact that a selection requirement of the superordinate DP phase is expanded in the exact position the selecting item is first merged. This suggests that the Delayed Phase Expansion proposed in Definition 22 cannot apply to complements that must move but that show p-stranding (even if they will become sequential phases after re-merge). This derives the contrast in (44a-b) and also the original conflict p-stranding between (3a-b) and (7a-b) (first noticed by Adriana Belletti and discussed in Chomsky 1986, see Bianchi & Chesi 2015):

(3) a. [$_{wh}$ *which* photographer]$_i$ C_{wh} did you buy [$_{DP}$ pictures [$_{PP}$ of $_{\_i}$]]?
b. ?* [$_{PP}$ of [$_{wh}$ *which* photographer]]$_i$ C_{wh} did you buy [$_{DP}$ pictures $_{\_i}$]?

(7) a. the person [of whom]$_i$ [pictures $_{\_i}$] are on the table (Chomsky 1986)
b. * the man who$_i$ [pictures [of $_{\_i}$]] are on the table

Two aspects must be clarified on this point: first, how exactly a PP can be merged when a *wh-* item is expected instead; second, how to make sure that the PP is really "extracted", and that it does not form a "A′-chain" with a resumptive pronoun as discussed in literature (Cinque 1990, Ch. 3).

On the first issue, we propose that the sinking option should be available for (certain) overt prepositions as it was in the case of "whose":

in (7), *person*, in order to be modified by a restrictive relative clause, should be inserted with featural make up [$_{N}$ *person* $_{=R\ +D/+P}$] requiring a relative pronoun/complementizer[22] to be merged and sunk within the RC nested phase. If +D feature is chosen, the DP (e.g. *which*...) would wrongly merge after the predicate and never within the nested [$_{NP}$ *pictures* ...]. On the other hand, *pictures* could become the sequential phase after re-merge if and only if, its =P select feature is not expanded yet (hence a *delayed phase expansion* is operative, signaled by the absence of the preposition after "pictures"). If =R +P are present, the PP (*of which*) can be correctly reconstructed/moved into the sequential phase, but =R expectation must sink under the PP[23].

The contrast is exemplified in (45a-b); the derivation (45a) is expanded below:

(45) a. [Of which cars] were [the hoods _] damaged by the explosion?
b.?* [Which cars] were [the hoods *of* _] damaged by the explosion?

i. **Expect**(Root$_{wh}$)
ii. **Sink**(wh +D)
iii. **Merge**($_{wh}$ Root, [$_{P\ =D(wh)}$ of])
iv. **Move**([$_{P\ =D(wh)}$ of])
v. **Merge**([$_{P\ =D(wh)}$ of], [$_{wh\ D\ =N}$ which])
vi. ...
vii. **Merge** [the hoods $_{=P}$] and move it in the M-buffer BEFORE projecting the PP expectation (delayed phase expansion) (this is possible only in (45a) since in (45b) we already have P projected, indicating a phase expansion)
viii. ...
ix. **Expect** [$_{V}$ *damaged* ... $_{=DP}$] last sequential phase

[22]This is not a trivial aspect: licensors like "+R" would predict that the merged R item should be moved in the M buffer since unselected, but the R feature itself should not be re-merged; this is the role of a D/P feature, hence =R +D/+R is the compatible featural set up leading to the analysis we want to provide for the headed restrictive relative clauses. On the other hand, a =R selected item could constitute a sequential phase that would produce unwanted predictions (i.e. it would be transparent for extraction if =R is the last selected item). We then assume that the correct implementation of a "matching" analysis of the restrictive relative clauses is the one selecting correctly a R complementizer, and requiring a D/P item to be co-indexed with the restricted head. This is done by assigning to the relative head the selection =R +D/+P feature. To prevent =R from constituting the sequential phase, we need to assume that every nominal phase requesting for a restricted RC, also has a "default" selection fulfilled by a null element after the =R feature inducing nesting of the RC clause.

[23]Sinking can be easily restricted to one phase (P or D). Extra assumptions are necessary for implementing massive pied-piping. This goes beyond the scope of this paper.

x. **Re-merge** the first D compatible item in M-buffer [$_{D}$ *the hoods* $_{=P}$]
xi. **Expect** $_{=P}$ as a sequential phase
xii. **Re-merge** [$_{PP}$ *of which cars*]

On the second issue, one might be tempted to assume that the PP was not "extracted" but linked to a resumptive pronoun (instead of a trace) while forming a representational A-chain. Since such a representational chain should not be sensitive to islands, and a resumptive pronoun can only be associated to a DP category, the contrast (46a-b) would be explained:

(46) a. [$_{DP}$ Who] did they leave [before speaking to $_{\_\ DP}$]?
b.?* [$_{PP}$ To whom] did they leave [before speaking $_{\_\ PP}$]?

As discussed in Bianchi & Chesi (2015), however, Truswell (2007) has shown that an adjunct can also be selectively transparent for extraction if the event it denotes is identified with an event position in the matrix predicate:

(47) a. * [$_{DP}$ What] does John dance [whistling $_{\_\ DP}$]?
b. [$_{DP}$ What] did John arrive [whistling $_{\_\ DP}$]?

This challenges the no-movement analysis.

4 Conclusion

In the last few years a series of different implementations of the basic minimalist intuitions have led to conflicting assumptions and generalizations that are hard to compare from a neutral standpoint. Here we considered the basic minimalist notions that are commonly used to explain pied-piping. These approaches, however, are not able to account for the whole set of facts in a complete way. Here we propose that the theory can be simplified by reversing the derivational perspective and then deducing the two general constraints (Don't Leave Your Child alone, Don't-LYC and Be Accessible, BA) that explain the major contrasts involved with pied-piping: if Merge reduces to fulfilling a selection expectation, Movement is driven by the insertion of a "larger than expected" item (i.e. partially unselected) and Phases are domains of an expectation, the overall framework get simplified and it becomes more performance-friendly.

References

Abels, Klaus. 2003. *Successive cyclicity, anti-locality, and adposition stranding*. University of Connecticut dissertation.

Abels, Klaus. 2012. *Phases: an essay on cyclicity in syntax* (Vol. 543). Walter de Gruyter.

Aboh, Enoch Oladé. 2004. *The morphosyntax of complement-head sequences: clause structure and word order patterns in kwa*. Oxford: Oxford University Press.

Barker, Chris. 2009. Reconstruction as delayed evaluation. In Erhard W. Hinrichs & John Nerbonne (eds.), *Theory and evidence in semantics*, 1–28. Stanford, CA: CSLI Publications.

Bianchi, Valentina & Cristiano Chesi. 2006. Phases, left-branch islands, and computational nesting. *Upenn Working Papers in Linguistics* 12. Proceedings of the 29th Penn Linguistic Colloquium, 15–28.

Bianchi, Valentina & Cristiano Chesi. 2014. Subject islands, reconstruction, and the flow of the computation. *Linguistic Inquiry* 45. 525–569.

Bianchi, Valentina & Cristiano Chesi. 2015. On a PP/DP asymmetry in extraction. In Di Domenico, Hamann & Matteini (eds.), *Structures, strategies and beyond. studies in honour of adriana belletti*, 47–66. Amsterdam: John Benjamins.

Blümel, Andreas. 2012. Towards a novel analysis of was-für-split. *Proceedings of ConSOLE XIX*. 105–123.

Brattico, Pauli. 2011. The diehard extended projection principle. *SKY Journal of Linguistics* 24. 181–188.

Cable, Seth. 2010. *The grammar of Q: Q-particles, wh-movement, and pied-piping*. Oxford University Press.

Cecchetto, Carlo & Caterina Donati. 2010. On labeling: principle C and head movement. *Syntax* 13. 241–278.

Chesi, Cristiano. 2012. *Competence and computation*. Padova: Unipress.

Chesi, Cristiano. 2013. Do the 'right' move. *Studies in Linguistics* 6. Special issue on 'directionality of phrase structure building', 107–138.

Chesi, Cristiano. 2015. On directionality of phrase structure building. *Journal of Psycholinguistic Research* 44. 1–50.

Chesi, Cristiano & Paolo Canal. 2018. Person features and lexical restrictions in italian clefts.

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

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1998. Minimalist inquiries: The framework. *MIT Occasional Papers in Linguistics* 15.

Chomsky, Noam. 2000. Minimalist inquiries. In R. Martin, D. Michaels & J. Uriagereka. (eds.), *Step by step: essays on minimalist syntax in honor of Howard Lasnik*. 89–156. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In M. Kenstowicz. (ed.), *Ken Hale: a life in language*, 1–52. Cambridge, MA: MIT Press.

Cinque, Guglielmo. to appear. Some notes on linear ordering and externalization.

Cinque, Guglielmo. 1990. *Types of A′-dependencies*. Cambridge, MA: The MIT Press.

Cinque, Guglielmo. 1994. On the evidence for partial N-movement in the Romance DP. In G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi & R. Zanuttini (eds.), *Paths towards universal grammar. Studies in honor of Richard S. Kayne*, 85–110. Washington: Georgetown University Press.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective*. Oxford: Oxford University Press.

Cinque, Guglielmo & Luigi Rizzi. 2008. The cartography of syntactic structures. *Studies in linguistics* 2. 42–58.

Collins, Chris. 2002. Eliminating labels. In Samuel Epstein & Daniel Seely (eds.), *Derivation and explanation in the minimalist program*, 43–64. Blackwell.

Collins, Chris & Edward Stabler. 2016. A formalization of minimalist syntax. *Syntax* 19. 43–78.

Elbourne, Paul D. 2005. *Situations and individuals*. Cambridge, MA: MIT Press.

Heck, Fabian. 2009. On certain properties of pied-piping. *Linguistic Inquiry* 40.

Huhmarniemi, Saara. 2012. *Finnish A'-movement: Edges and islands*. University of Helsinki dissertation.

Huhmarniemi, Saara & Pauli Brattico. 2013. On primary and secondary movement. *Acta Linguistica Hungarica* 60. 173–216.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Boston, MA: MIT Press.

Leu, Thomas. 2008. What for internally. *Syntax* 11. 1–25.

Manninen, S. 2003. Finnish PPs and the phase impenetrability condition. In D. Nelson & S. Manninen (eds.), *Generative approaches to Finnic and Saami linguistics*, 295–320. CSLI Publications.

Pollard, Carl & Ivan A Sag. 1994. *Head-driven phrase structure grammar*. Chicago, IL: Chicago University Press.

Prince, Alan & Paul Smolensky. 2004. *Optimality theory: constraint interaction in generative grammar*. Malden, MA: Blackwell Publishing.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Haegeman L. (ed.), *Elements of grammar* (Kluwer International Handbooks of Linguistics). Dordrecht: Springer.

Rizzi, Luigi. 2004. On the form of chains.
Rizzi, Luigi. 2016. Labeling, maximality and the head–phrase distinction. *The Linguistic Review* 33. 103–127.
Safir, Ken. 1985. Missing subjects in German. In Jindrich Toman (ed.), *Studies in german grammar*, 193–229. Dordrecht: Foris Publications.
Shlonsky, Ur. 2004. The form of semitic noun phrases. *Lingua* 114. 1465–1526.
Stabler, Edward. 1997. Derivational minimalism. In Retoré C. (ed.), *Logical aspects of computational linguistics. LACL 1996. Lecture notes in computer science*, vol. vol 1328. Berlin/Heidelberg: Springer.
Truswell, Robert. 2007. Extraction from adjuncts and the structure of events. *Lingua* 117. 1355–1377.
Zamparelli, Roberto. 2000. *Layers in the determiner phrase*. New York: Garland.