# Freezing Effects in a free-Merge System[*]

## A Configurational Approach

**Abstract:** This paper reconsiders so-called freezing effects within Chomsky's (2004 and sub.) Phase Theory. I argue that freezing (or halting) should not be seen as the consequence of an exocentric {XP,YP} structure in which the heads of XP and YP share some feature (cf. Chomsky 2013, 2015) or as the invisibility of X' projections (cf. Rizzi 2015). Instead, I submit that A-freezing (Chomsky's 2000, 2001 *Activity Condition*) and A-bar freezing (Rizzi's 2006 *Criterial Freezing*) should be dealt with by different principles: the former follows from an independently motivated rule of efficient computation (the application of cyclic Transfer; cf. Chomsky 2000, Uriagereka 1999), coupled with Labeling Theory (cf. Chomsky 2013, 2015), whereas the latter is simply syntactically vacuous. In line with previous proposals (cf. Gallego 2009, Epstein, Kitahara & Seely 2016), I claim that XPs in edge positions are not frozen in the narrow syntax (they can always move, unless affected by cyclic Transfer). Nevertheless, such XPs may be part of a configuration and thus receive an interpretation at the semantic component (cf. Chomsky 2001, 2004). Therefore, if they move *from an edge*, the relevant interpretation *at that edge* (be it topic, focus, etc.) will be lost, as interpretations of the relevant kind (theta-roles, criterial-roles, etc.) cannot accumulate, which I ultimately attribute to a *Principle of Interface Freezing*, whose effects can be subsumed under the *Principle of Full Interpretation* (cf. Chomsky 1986a).

**Keywords:** freezing, halting, interface conditions, labeling, Merge, Transfer.

# 1    Introduction

A long-noted fact about context-sensitive transformations is that they cannot apply under certain conditions. The relevant conditions vary and the literature has offered different approaches to such locality constraints under well-known labels, like island, freezing, minimality, or anti-locality effects (cf. Ross 1967, Chomsky 1964, 1973, 1986a, 2000, 2008, Grohmann 2003, Huang 1982, Lasnik & Saito 1992, Rizzi 1996, 2006, 2007, 2015, Uriagereka 1999, Stepanov 2001; cf. Freidin 1999, Lasnik 2006, and Uriagereka 2011, for overviews). Consider, for instance, the impossibility to A-move subjects from inflected (φ-complete, sensu Chomsky 2001) clauses, or A-bar move them across an overt C (that), or extract from within them in (1):

(1) a. *[CP Trump seems [TP tTrump was voted t by many people ] ]
    b. *[CP Who did you say [CP that tWho voted Trump ] ]?
    c. *[CP Who do you think [CP that [ [the voters of tWho ] had serious doubts ] ] ]?

Similar effects hold of wh-phrases more generally. They cannot stop-and-go from a [+Q] C, nor move from another wh-phrase in [Spec, CP], nor undergo topicalization after having reached the CP, as shown in (2):

(2) a. *[CP Which voters do you wonder [CP tWhich voters Trump convinced tWhich voters ] ]?
    b. *[CP Which city do you wonder [CP [which voters of tWhich city ] Trump convinced tWhich voters of which city ] ] ?
    c. *[CP Who, did [TP many voters have doubts about tWho ] ]?

The facts in (1) and (2), and many more, have been studied by most approaches to locality conditions, and although the empirical basis is well described, a unitary explanation is missing.[1] Part of these conditions on transforma-

---

tions fall within what has been recently referred to as "freezing" (or "halting") effects. The gist behind the notion of freezing within minimalism goes back to Chomsky (2000) (although the notion itself was already used before; Epstein 1992 was an important precursor of it, as well as an early economy-based attempt to deal with freezing phenomena), where it is argued that subjects in a φ-complete [Spec, TP] become inactive (that is, frozen) after having checked their Case feature: they cannot move, and they also block movement from any XP within them.

In the recent literature, freezing effects have been the focus of situations where a dislocated XP (*d*XP, from now on) reaches a position in the Left Periphery (a criterial position, in Rizzi's 2006, 2007 terms) and stops there. The effect is typically regarded as the A-bar version of A-freezing, so by parity of reasoning the *d*XP can no longer move nor allow extraction form within. Importantly, the effect is related to a last-resort logic: some feature F triggers movement of *d*XP, and once it is checked, *d*XP becomes inert. Though popular in the literature, this approach becomes implausible once Move is actually a variant of Merge (Internal Merge, as Chomsky 2004 argued), an operation that is feature-free (Chomsky 2005 and sub.).[2] Here I will assume this streamlined view of Merge, which I take to be subject to principles of computational efficiency alone, the *No Tampering Condition* (NTC; cf. Chomsky 2000) and the *Phase Impenetrability Condition* (PIC; cf. Chomsky 2000).

Following Gallego (2009), I assume that *d*XPs in criterial positions are not syntactically frozen (*pace* Rizzi 2006, 2007, 2015, Bošković 2008). If they are in the edge of a phase, some special UG-enriching device would be required to make sure that *d*XPs cannot move. Movement (IM) will thus be granted (as it always is from a phase edge), but the interpretation will be lost. I will thus submit that *d*XPs cannot leave the position they occupy under the assumption that such position (such configuration) is mandatory for them to receive an interpretation at the semantic component. To be specific, I will explore the consequences of (3), which I regard as an instantiation from Chomsky's (1986b) *Principle of Full Interpretation* (PFI, henceforth) that also assumes ideas in Chomsky (2001, 2004) about interpretive principles applying at phase edges.[3]

(3) Principle of Interface Freezing (PIF)
    A *d*XP is assigned INT at SEM if *d*XP occupies a phase edge
    [INT = an interpretation, SEM = semantic component]

The PIF entails that the relevant interpretation of *d*XPs is lost if these move, destroying the surface configuration, an option that cannot be prevented. This conclusion is similar to Epstein, Kitahara & Seely's (2016) claim that halting effects can be derived from morphological or semantic principles, but not syntactic ones. Consequently, the fact that which incident can stop in the matrix and embedded [Spec, CP] positions, yielding different interpretations, falls into place (I use bold letters to signal the wh-phrases that freeze):

(4) a. [CP **Which incident** C [ did the media say [CP t_{which incident} that Trump was involved in t_{which incident} ] ] ] ?
    b. [CP The media said [ **which incident** C [CP Trump was involved in t_{which incident} ] ] ]?

The logic above carries over to customary freezing situations studied by Rizzi (2006, 2007), like that in (5):

(5) *[CP Which book does Bill wonder [CP t_{Which book} C [TP she read t_{Which book} ] ] ]?

[from Rizzi 2006:112]

Here, the wh-phrase *which book* cannot move once it has landed in the [Spec, CP] of an interrogative clause. Rizzi (2006, 2007) attributes this effect to an A-bar version of Chomsky's (2000, 2001) A freezing (the *Activity Condition*, AC), hence invoking feature checking. In the approach adopted here, the problem in (5) follows from the PIF, and is not syntactic in nature. As we will see, an analysis based on the PIF accounts for most A-bar movement cases (including dislocation) and ECP effects, to which I return in section 4.1.[4]

——

[1] A reviewer notes, rightly, that the examples in (1) do not belong, strictly speaking, to the same paradigm, at least historically. The same holds for the data in (2). All I want to emphasize is that both A and A-bar transformations are subject to similar constraints, all of which seem to be restatable in freezing terms.

[2] See Pesetsky & Torrego (2006) for the possibility that every application of Merge is parasitic on feature checking.

[3] An anonymous reviewer wants me to clarify what the principle (3) is and to explain what INT is supposed to convey. As noted, I take (3) to be nothing but a specific formulation of the PFI. The label INT is left vague (on purpose), but in the case at hand it should be taken to cover interpretive aspects related to information structure and discourse-related notions.

[4] In this paper I put aside the possibility that dislocation is analyzed as involving a hidden biclausal structure (cf. Ott 2014).

The paper is organized as follows: section 2 introduces the technical assumptions that this paper will make concerning the structure building operation Merge, Labeling Theory, and Phase Theory; section 3 reviews Chomsky's (2013, 2015) approach to labeling and how it can be used to deal with freezing effects; section 4 puts forward a feature-free approach to freezing effects that will be largely based on conditions playing a role at SEM interface (labeling and the PIF); section 5 summarizes the main conclusions.

# 2  Perspectives on Merge and Labels

This section reviews the main developments of the basic structure-operation, Merge, and its relation with features and labeling. With Chomsky (2004 and sub.), I take Merge to be feature-free, a hypothesis I ground on the idea that its duality (EM and IM) is correlated with interpretive properties of the Conceptual-Intentional systems.

## 2.1  Feature-free Merge

Following Chomsky (cf. 1995 and sub.), I assume here that a computational system (a Narrow Syntax, NS) of discrete infinity must assume both a mechanism of combination (Merge) and a list of atomic elements (lexical items, LIs) to which such mechanism applies. Merge can thus be conceived of as an operation that takes two syntactic objects (SOs), X and Y, and creates the set {X,Y}—what Epstein, Kitahara, and Seely (2014) call "simplest Merge." Since there is no set-theoretical condition requiring labels (non-terminal symbols), these entities can only be added by stipulation (going beyond Merge, as Collins 2002 noted). Although much current literature on phrase structure still assumes labels in one form or another, notice that this seems to be a residue of X-bar theory, which endorsed a restrictive view of compositionality: endocentric compositionality (i.e., composition regarded as successive attachment to a head, yielding endocentricity, distinction between complements and specifiers, etc.). Once X-bar theory is dispensed with, there is no reason why Merge should impose such constraint. The simplest formulation just says that two SOs can be merged, with no additional symbols or features being added or projected, in accord with the NTC (cf. Chomsky 2001, 2005, 2008), which subsumes the *Inclusiveness Condition* of Chomsky (1995):

> A natural requirement for efficient computation is a ''no-tampering condition'' (NTC): Merge of X and Y leaves the two SOs unchanged. If so, then Merge of X and Y can be taken to yield the set {X, Y}, the simplest possibility worth considering. Merge cannot break up X or Y, or add new features to them. Therefore Merge is invariably ''to the edge'' and we also try to establish the ''inclusiveness principle,'' dispensing with bar levels, traces, indices, and similar descriptive technology introduced in the course of derivation of an expression. It seems that this desideratum of efficient computation can also be met within narrow syntax at least.          [from Chomsky 2008:138]

Although labels in the X-bar-theoretic sense are dispensed with, already Chomsky (2004) hinted at the possibility that efficient computation requires for the nature of SOs to be determined somehow. The idea is developed in Chomsky (2008) and further sharpened in Chomsky (2013), where a label algorithm (LA) is put forward. For Chomsky (2013), LA operates through "Minimal Search," hence locating the most accessible (i.e., minimal) SOs: Heads.

> Projection is a theory-internal notion, part of the computational process [Generative Procedure]. For a syntactic object SO to be interpreted, some information is necessary about it: what kind of object is it? Labeling is the process of providing that information. Under [Phrase Structure Grammar] and its offshoots, labeling is part of the process of forming a syntactic object SO. But that is no longer true when the stipulations of these systems are eliminated in the simpler Merge-based conception of UG. We assume, then, that there is a fixed labeling algorithm LA that licenses SOs so that they can be interpreted at the interfaces, operating at the phase level along with other operations. The simplest assumption is that LA is just minimal search, presumably appropriating a third factor principle, as in Agree and other operations. In the best case, the relevant information about SO will be provided by a single designated element within it: a computational atom, to first approximation a lexical item LI, a head. This LI should provide the label found by LA, when the algorithm can apply.          [from Chomsky 2013:43]

Operating without bounds, Merge applies as indicated in (6) to construct a derivation: X and Y are selected to construct a new object, which we can call Z for expository purposes. Subsequent applications of Merge target Z, which is the only object in the derivation (Chomsky 1995:243), to yield Z', and then Z'', and so on:[5, 6]

(6) a. Merge (X,Y) = Z = {X,Y}
    b. Merge (W,Z) = Z' = {W,Z}
    c. Merge (K,Z') = Z'' = {K,Z'}

Let us concentrate on (6a). After the application of Merge, the workspace contains Z and nothing else. At this point, we may want to merge W and Z. W is either internal to Z or external to Z. If W is external, then W must be taken from the lexicon. This is External Merge (EM). If W is internal (suppose W = X), then it is a term of Z; if the NTC applies, then Z must be unchanged, so still {X,Y}, yielding {X,{X,Y}} two copies (occurrences) of X. This is Internal Merge (IM, previously Move). A question arises about how copies are obtained. Many approaches assume an additional operation, Copy, which duplicates the relevant element. As Noam Chomsky observes through personal communication, "duplication" (copying) is not needed with IM, just like it is not with EM: In particular, suppose we select Z from the workspace, then select X to merge it to Z. If X is external to Y (so, taken from the lexicon), we don't have to first copy X in order to merge it to Z.[7] The same holds for IM.

Before concluding this section, I would like to discuss another assumption of the derivational approach developed in Chomsky (2000 and sub.). For reasons attributed to computational complexity (reduction of computational load), Chomsky (2000) assumes that SOs assembled by Merge are handed over to the external components at some point. The relevant Transfer-units are called "phases" and are defined as the loci of uninterpretable $\varphi$-features and structural Case.[8] These features are encoded in the lexicon in dedicated functional heads (C and v), which act as Probes seeking a matching Goal within its complement. This is illustrated in (7), where H is a phase head (a Probe), and $\alpha$ and $\beta$ correspond to the complement and specifier respectively.

(7) [ $\beta$ [ H (Probe) [$\alpha$ . . . XP (Goal) . . . ] ] ]

For empirical reasons, Chomsky (2004) assumes that Transfer only affects the complement of H, leaving H itself and $\beta$ visible for subsequent X and XP movement. The combination of H and $\beta$ is dubbed edge.[9] The main effect of cyclic Transfer concerns the periodic forgetting of phase complements, hence rendered unaccessible in subsequent stages of the derivation.[10] This make it possible to capture strict cyclicity and some version of compositionality. Chomsky (2000) formulates a *Phase Impenetrability Condition* (PIC) to capture the idea that only the complement of H is transferred.[11]

----

[5] That X and Y are no longer available was expressed in the following passage: "Applied to two objects $\alpha$ and $\beta$, Merge forms the new object K, *eliminating $\alpha$ and $\beta$*" (Chomsky 1995:243, my emphasis). For further discussion, I refer the reader to Chomsky et al. (2017).

[6] Chomsky (2007:11, 2008:139) assumes that the free nature of Merge follows from LIs having an edge feature (EF) that is undeletable and can thus give rise to an unbounded application of Merge—the term "edge" is used to capture the fact that the operation always adds structure to the already assembled SO, on its root / edge. I will put EFs aside in this paper, as I regard them as a purely theory internal device. This will allow me to dispense with the technical problems discussed in Narita (2014) (related to the lack of EF percolation).

[7] The problem is more general if X and Y remained in the workspace, along with Z, in (6). As Noam Chomsky (p.c.) points out, it has always been assumed that they do not, for the generative procedure constructs a single object, not a multiplicity of objects. Changing that convention would mean that instead of a generative process for expressions, we would be designing a generative process for an arbitrarily large collection of expressions, unrelated expressions, and it would cause more specific problems. For instance, suppose that we hold that after EM (X,Y) = Z = {X,Y}, the workspace contains X, Y, Z. We then have a new question: what's the relation between X in the workspace (call it X1) and X in Z = {X,Y} (call it X2). They are either copies or repetitions. If they are copies, everything goes haywire. Thus, if we continue to Merge to X1 finally yielding the finite clause FC, and to Z yielding the finite clause FC', then the two clauses would contain the two copies X1 and X2, so one should be deleted, and if one enters into some relation (say anaphora) then the other does, etc. Things get much worse if, as this proposal allows, we construct simultaneously indefinitely many finite clauses. This is not only dubious, and in fact makes the notion of "copy" collapse.

[8] This is the assumption made in Chomsky (2008:155), although the kind of SOs that qualify as phases is subject to debate (cf. Gallego 2012 for discussion).

[9] Chomsky (2004) points out that Transfer of phases in full is limited to root clauses.

[10] Chomsky (2008:143) suggests that Agree can search the domain of an already transferred complement domain, thus accounting for DAT-NOM constructions like (i), where T $\varphi$-agrees with the internal argument after Transfer:

  (i)      A  Trump le        gustan  las murallas                     (Spanish)
                to  Trump CL-dat.sg like-e.pl the  walls
                'Trump likes walls'

(8) Phase Impenetrability Condition (PIC)

In phase $\alpha$ with head H, the domain of H is not accessible to operations outside $\alpha$;
Only H and its edge are accessible to such operations

[from Chomsky 2000:108]

Applied to (7), the PIC renders $\alpha$ no longer accessible:

(9) [ $\beta$ [ H (Probe) [$\alpha$ . . . XP (Goal) . . . ] ] ]

Having considered how Merge and Transfer work, let us now discuss the possibility that the two variants of the former (EM and IM) can be given a principled explanation—ideally if they follow from interface conditions. The proposal to be developed below departs in serious respects from feature-bound approaches to IM and, consequently, to the idea that A-bar dependents participate in Spec-Head agreement, a hallmark of "criteria" (cf. Rizzi 2006).

## 2.2  The Duality of Merge: C-I conditions and types of freezing

We have just seen that, under the NTC, Merge can apply in two ways, a possibility that Chomsky (2004 and sub.) relates to conditions imposed by the Conceptual-Intentional systems.

> In a well-designed FL, lacking arbitrary stipulations, both EM and IM should be permitted, and the two kinds of Merge should be expected to yield different interface properties. That is obviously true at the SM interface—the ubiquitous property of "displacement"—and appears to be true at CI as well. The two types of Merge correlate well with the duality of semantics that has been studied from various points of view over the years. EM yields generalized argument structure, and IM all other semantic properties: discourse-related and scopal properties.        [from Chomsky 2007:19]

> The hypothesis is that C-I incorporates a dual semantics, with generalized argument structure as one component, the other being discourse-related and scopal properties. Language seeks to satisfy the duality in the optimal way, EM serving one function and IM the other, avoiding additional means to express these properties.        [from Chomsky 2008:141]

To be fair, the correlation is far from clear (especially in the case of discourse-related properties, where IM can give rise to different types of interpretations, left open in the syntax by Chomsky 2001 and Uriagereka 1995), but it does provide a rationale for the EM / IM divide. Assuming well-established ideas of the GB framework, most approaches to movement have built on the hypothesis that this is a *Last Resort* (LR) operation, hence morphologically driven (cf. Chomsky 1986a). It is important to bear in mind that the questions LR was designed to address are: (i) why the DP *Trump* must move to [Spec, TP] in (10a) (the EPP position) and (ii) why, once there, it cannot move further, as (10b) shows.

(10)  a. *[$_{TP}$ T was [$_{vP}$ v nominated Trump ] ]
    b. *[$_{TP}$ Trump$_i$ T seems [$_{CP}$ (that) t$_i$ was nominated t$_i$ ] ]

Chomsky (1995:256-257) answered both questions by arguing that movement is compulsory due to morphological (checking) reasons: *Trump* moves to [Spec, TP] to check nominative Case, and once there, it needs no further checking—thus, since checking is not necessary, it is not allowed.

The idea that features and the operation Move (nowadays IM) go hand in hand quickly became popular within minimalism, giving rise to the assumption that all movements were triggered to satisfy morphological needs. Some

——

This raises the possibility that the SOs created by Merge are not literally 'expunged' from the computation, but kept in some form (as expected, under NTC). Independent evidence shows that this is in fact needed. Thus, $\alpha$ in (ii) is spelled out with the wh-phrase *which book*, although it must have been transferred at an earlier stage of the derivation:

(ii)        [ [ Which book [$\alpha$ that John bought ] ] did you read t ] ?

[11] Noam Chomsky (p.c.) observes that the PIC has a stronger and a weaker version: what has been processed is either totally inaccessible or alternatively cannot be changed. Given the evidence in the previous footnote, what is inaccessible cannot be completely deleted, so it must be retained in some form. One possibility is that it is simply retained without any change at all, and further computation is constrained by the PIC. Another possibility, expressed as Transfer, is that it is retained as a pair <X,Y>, where X is in a form accessible only to S-M [SensoriMotor] and Y only to C-I [Conceptual-Intentional].

years later this picture became more general, after Chomsky (2004) claimed that Move and Merge were variants of the same structure-building operation: External Merge and Internal Merge, as we have seen in section 2.1. The moment IM was postulated (and even before, cf. already Contreras & Masullo 2000), each and every instance of Merge was taken to obey LR, and thus involve feature checking. Consequently, along with Case and φ-features, authors postulated Θ-features (cf. Bošković 1994, Hornstein 2001, Lasnik 1999), selectional features (cf. Chomsky 1965), criterial features (cf. Rizzi 1997, 2004, 2006, 2007), operator features (cf. Bošković 2008), linearization features (cf. Biberauer et al. 2010), pragmatic features (cf. Haegeman & Hill 2010, Speas & Tenny 2003), and so on and so forth, making every application of Merge legitimate.

Although this approach may be helpful in circumstances where Merge-mates share some inflectional feature, it is easy to see that pushing it to *every* application of Merge forces one to postulate new features, some more *ad hoc* than others (cf. Adger & Svenonius 2011, Šimík 2011). Departing from this widespread reasoning, nowadays hegemonic in the field, Chomsky proposed to turn the argument upside down: If EM is free (involving no feature checking whatsoever), then IM should be free too.[12]

> It has always been presupposed without comment that EM comes free: no one has postulated an "EPP property" for EM or stipulated that it satisfies the [No Tampering Condition]. IM, in contrast, has been regarded (by me, in particular) as a problematic operation, an "imperfection" of language that has to be postulated as an unexplained property of UG unless it can be motivated in some principled way [through feature checking] […] A few years ago, it became clear that this is a misunderstanding. IM […] is as free as EM. […] It follows that any alternative device [feature checking] to deal with the displacement property and the duality of semantics requires double stipulation: to ban IM, and to justify the new device.                [from Chomsky 2008:140-141]

In the same breath, Chomsky (2008:141) conjectures that "we thus expect language to use IM rather than other mechanisms that can be devised [features, feature checking, etc.] to express semantic properties apart from generalized argument structure". This brings us directly to the controversial notion of feature checking and LR, and their relevance for freezing effects. In the GB framework, feature checking was understood as a dependency between two SOs in a Spec-Head configuration, which required making use of a broad version of c-command, namely m-command.[13] Chomsky (2000, 2001) refines the notion of checking, defining it as a valuation procedure that can operate at a distance: Agree. A central assumption of this take on feature checking is that some functional heads are drawn from the lexicon with their φ-features unvalued. Since they do not have a value, features act as a seeker (a Probe) that looks for a value-inducing element (a Goal) in its c-command domain (cf. section 2.1.).

Under an Agree-based approach to agreement, two problems emerge for GB-rooted Spec-Head checking: (i) some features (typically, semantic/pragmatic ones; e.g., [+topic], [+focus], [+Q]) cannot be treated as "attributes with values," and (ii) feature checking does not always need displacement. Examples like (11a,b) reveval that checking of φ-features is independent of movement (cf. Chomsky 2000, 2001):

(11) a. {\*Embistieron / Embistió}  a   los molinos Don Quijote                         (Spanish)
        charge-3.{PL / SG}        to the mills     Don Quixote
        'Don Quixote charged at the mills'
    b. Mér   {\*virðist / virðast}  Þeir      vera     skemmtilegir                    (Icelandic)
        me-DAT seem-3.{SG / PL}  they-NOM  be-INF  interesting
        'It seems to me that they are interesting'

---

[12] The hypothesis that Merge is free does not of course mean that 'anything goes.' What Chomsky (2004, 2007) suggests is that Merge applies freely during the computation, conditions of the C-I interface being interpretive filters. Consequently, the system can overgenerate. I will remain indifferent as to whether this is the right scenario, for such decision largely depends on how computational works in the mind, a matter that is admittedly murky (cf. Gallistel & King 2010, Murphy 2015, and references therein for discussion).

[13] There are clear examples of this approach to agreement, especially Chomsky (1986a) and Kayne (1989). For more thoroug discussion, see Hiraiwa (2005).

Data like these, where Probe and Goal are not in a Spec-Head configuration, are pervasive, both with subjects and objetcs, and provide support to Chomsky's (2000, 2001) Agree.[14] An important consequence of this system is that after a Goal agrees it is rendered inactive (frozen), unable to agree again or be IM-ed, but capable of giving rise to intervention effects (cf. Chomsky 2000, 2001). This is what explains the facts in (12), where *the athletes* cannot long-distance agree with matrix T nor raise to it:[15]

(12) a. *It seem [$_{CP}$ that **the athletes** won the gold ]
    b. ***The athletes** seem [$_{CP}$ that t$_{the\ athletes}$ won the gold ]

The idea behind "freezing" has a long and manifold tradition, going back to Chain Uniformity, Anti-Locality, Subjacency, Barriers, Minimality, and CED effects (cf. Wexler & Culicover 1981, Chomsky 1986b, 2001, Chomsky & Lasnik 1995, Rizzi 1990, Takahashi 1994, Uriagereka 1999, Stepanov 2001, Starke 2001, Grohmann 2003, and Bošković 2008). These approaches do not agree about what freezing actually amounts to, but all of them share the intuition that 'too much' movement or 'too much' checking yields locality problems.

Consider now how the overall scenario has been interpreted within the Cartographic Project, where LR is invoked in order to account for a generalized approach to criteria. More specifically, a *d*XP must be licensed by satisfying a P-Criterion, where P is a shorthand for criterial features (cf. Rizzi 2006, 2007): topic, focus, relative, Q, and the like.[16] At the heart of the criterial approach to checking is the idea that IM to the Left Periphery must be motivated. In this system, the C head is split into different functional heads endowed with a dedicated interpretive feature that must be checked with a *d*XP.

> As for Chomsky's scope-discourse positions, I will assume that they are determined by a family of principles, the criteria, which require specifier-head agreement with respect to features of the relevant class: Q, Top, Foc, R for questions, topic, focus, relatives, and so on (see Rizzi 1996 for an early formulation of this approach)      [from Rizzi 2006:101-102]

Rizzi (2006) goes on to argue that the format of the criteria is as in (13):

(13) [*d*XP]$_F$ and [P]$_F$ must be in a specifier-head configuration, for F = Q, Top, Foc, R, …

                                        [from Rizzi 2006:102]

Rizzi (2006:111) discusses if IM can apply after a *d*XP has reached its criterial position. In the case of EM (argument structure), the answer is trivially negative: XPs cannot be EM-ed more than once, by definition. Things are different in the case of IM, though. Borrowing data from Lasnik & Saito (1992), this author provides the paradigm in (14) to defend that *d*XP are frozen in place upon hitting a criterion-satisfying specifier (as above, I use bold to indicate frozen *d*XPs):

(14) a. Bill wonders [$_{CP}$ **which book** C$_Q$ [ she read t$_{which\ book}$ ] ]
    b. *Which book C$_Q$ does Bill wonder [$_{CP}$ t$_{which\ book}$ C$_Q$ [ she read t$_{which\ book}$ ] ]?
    c. *Which book C$_Q$ does Bill wonder [$_{CP}$ t$_{which\ book}$ C$_Q$ [ she read t$_{which\ book}$ ] ]?

Although the effects in (14) are clear, Rizzi (2007) ends up offering a weakened version of criterial freezing in order to account for the fact that criterial Goals allow from some of its terms to be extracted. The examples that motivated this weakened version were first discussed in Torrego (1985). As (15) reveals, the wh-phrase *de qué autora* (Eng. 'of which author') is subextracted out of the larger *qué traducciones de qué autora* (Eng. 'which translations by which author'). This is unexpected, if the latter is frozen (inactive) after IM applies.

---

[14] The Agree perspective is not unanimously entertained (cf. Hornstein 2009, and references therein). Descriptively at least, data like those in (11) suggest that some cases of checking do not involve a Spec-Head configuration. For additional discussion, see Bošković (2007) and Zeijlstra (2012).

[15] Chomsky (2007:23 fn.31, 2008:150) subsumes the AC under the PIC.

[16] For the purposes of presentation, I use the labels "topic" and "focus" in a restricted sense, simply to convey the syntacic processes of topicalization and focalization to the Left Periphery of the clause, as in (i) and (ii).
    (i)         The books, I read t$_{the\ books}$
    (ii)        THE BOOKS I read t$_{the\ books}$ ! (…not the newspapers)
Therefore I will not discuss the intricacies of the different types of foci / topics that have been studied in the literature (presentational focus, verum focus, contrastive focus, etc.). Cf. Benincà & Munaro (2010), Rubio (2014), and references therein.

(15) [$_{CP}$ De qué  autora$_i$ C [ no  sabes   [$_{CP}$ [ qué  traducciones t$_i$]$_j$  [$_{TP}$ t$_j$ ganaron  premios] ] ]?           (Spanish)
　　 of  what author    not  know-2.SG  what translations have-3.PL  won       awards
　　 'Which author don't you know what translated books have won awards?'

To accommodate data like (15), Rizzi (2007) postulates (16), which is extended so that it can cover Subject Condition effects (cf. Huang 1982):

(16) Criterial Freezing
　 In a criterial configuration, the Criterial Goal is frozen in place

[from Rizzi 2007:149]

Given (16), only the head of the bigger wh-XP, called "criterial Goal," is frozen, while the wh-internal remnant remains accessible. Although certainly consistent with the facts, the analysis raises some doubts—the most pressing one, how come the internal part of an XP doesn't freeze if the latter does.[17]

Let us stop here. In this section we have seen that the EM / IM cut can follow from C-I conditions, assuming that the semantics require for each variant to be associated with a specific type of interpretation. I have amphasized that, precisely for Chomsky's (2004) simplification, neither EM nor IM can be feature-driven. This is not saying that features do not exist or that they play no role in the grammar, but it does entail that features with a construction-like flavor (all 'relational' or 'syntagmatic' notions, like theta-roles, syntactic functions, and discourse-related notions) cannot be regarded as LIs, let alone features that drive derivations (cf. Chomsky 2001:6, 2008:151, López 2009). If features of the A-bar type are not features in the technical sense (an attribute with a value), then the entire feature-freezing logic of Chomsky's (2001) AC and Rizzi's (2006, 2007) Criterial Freezing go away. Before exploring the PIF I introduced in section 1, I would like to consider a couple of more recent, feature-free, alternatives to freezing.

# 3 The POP+ approach to freezing

This section discusses two alternative views on freezing effects based on Labeling Theory (cf. Chomsky 2013, 2015). 3.1. I introduces the basics of Chomsky's (2013, 2015) analysis of labeling. Then 3.2. explores different ways to make specifiers stable. Finally, in 3.3. I turn my attention to Rizzi's (2015) proposal, which is also different from his previous proposals.

## 3.1 Labeling Theory: the POP+ framework of Chomsky (2013, 2015)

As noted in section 2.1., Chomsky's (2004 and sub.) formulation of Bare Phrase Structure dispenses with X-bar-theoretic machinery, including labels. Thus, departing from the original proposal (cf. Chomsky 1994), the combination of X and Y, yields (17), not (18) (where K is a label, an actual "projection" of either X or Y).

(17) Merge (X,Y) → Z = {K,{X,Y}}   (Chomsky 1995, 2000, 2001)

(18) Merge (X,Y) → Z = {X,Y}       (Chomsky 2004 and sub.)

Chomsky (1995:244) took K to be identical to X or Y, not its union ($\alpha \cup \beta$) or its intersection ($\alpha \cap \beta$). Departing from this (still partially X-bar reminiscent) formulation, Chomsky (2004 and sub.) puts forward a label-free analysis in

---

[17] These data were also discussed by Lasnik & Saito (1992:102):
　　(i)　　　??[$_{CP}$ Who$_i$ C do you wonder [$_{CP}$ [which picture of t$_i$ ]$_j$ C Mary bought t$_j$ ] ]?
　　(ii)　　　??[$_{CP}$ Who$_i$ C do you wonder [$_{CP}$ [which picture of t$_i$ ]$_j$ C t$_j$ is on sale ] ]?
As discussed in Gallego (2009), the facts are misleading, since subextraction does not take place from the [Spec, CP] (criterial) position. In fact, these data may involve a process of reanalysis applying at the VP level, as Bosque & Gallego (2014) discuss, following original ideas of Bach & Horn (1976), also developed by Broekhuis (2006).

the spirit of Collins (2002). In Chomsky (2013), labels are not projected, but determined by LA, which renders SOs interpretable at the interfaces.[18] The first formulation of the LA was (19):

(19) Labeling Algorithm
    a. In $\{H, \alpha\}$, H an LI, H is the label
    b. If $\alpha$ is internally merged to $\beta$, forming $\{\alpha, \beta\}$, then the label of $\beta$ is the label of $\{\alpha, \beta\}$

<div align="right">[from Chomsky 2008:145]</div>

Chomsky (2013:43) argues that LA operates under Minimal Search (MS), a third-factor principle present in other computational operations.[19] MS locates the most accessible element within a given domain D: a minimal unit—an LI ($X^{min}$). MS operates unproblematically in $\{H,XP\}$ structures (where H is the label), but it does not in $\{XP,YP\}$, where MS leads an ambiguous result. Chomsky (2013:43) argues that this unwanted situation can be tackled in two ways: (i) either XP or YP moves, or else (ii) X and Y share some feature that can be interpreted as the label of $\{XP,YP\}$. Under the assumption that copies are invisible to computation (cf. Chomsky 2000:131, 2001:16, 23-24), (i) makes it possible for MS to determine the label in $\{XP,YP\}$, after XP raising.

(20) a. $\{XP,YP\}$           MS: ambiguous
    b. $\{YP,\{XP, t_{YP}\}\}$     MS: X labels $\{XP,YP\}$ (YP's copy is invisible)

After YP raises, the computation sees "$\{XP, t_{YP}\}$," so the head of XP is accessible to MS. Empirically, the benefits of this approach cover successive cyclic A-bar movement (which is just a case of labeling failure in $\{XP,YP\}$; cf. Blümel 2014) and subject raising to [Spec, TP] from [Spec, v*P] (and other small clauses; cf. Moro 2000).[20]

Consider next option (ii), which directly concern conditions under which *d*XPs freeze (or halt). Chomsky (2013:45) argues that the final position of a *d*XP in (21a,b) is identified through the features the heads X and Y are endowed with: $\varphi$-features in (21a), and Q features in (21b).[21, 22]

(21) a. $\{\{A, student_{\varphi}\}, \{T_{\varphi}, \{ \ldots t_{A\,student} \ldots \}\}\}$       MS: $\varphi$ features
       ↑_____↑ (Agree)
    b. $\{\{which_Q, books\}, \{C_Q, \{ \ldots t_{which\,books} \ldots \}\}\}$ MS: Q features
       ↑_____↑ (Agree)

Chomsky (2015) pushes the labeling analysis in (21) to ECP effects, which fall within what Rizzi (2006, 2007) calls "Subject Criterion" (a subcase of Criterial Freeing for him). In (22) below, extraction of *who* yields a *that*-t effect, which Chomsky (2015:9-11) relates to a de-labeling process. More specifically, Chomsky (2015) assumes that subjects must

---

18 Note that this tacitly assumes that (compositional) interpretation is endocentric. Cf. Narita (2014) for additional discussion.

19 It is not clear that the MS operating in LA is the one that operates in regular Probe-Goal dependencies (Agree). As Sam Epstein (p.c.) notes, MS in $\{H, XP\}$ finds H and stops. By contrast, when a Probe is located in $\{H,XP\}$, MS does not find H and stops—rather, MS continues to search for another head (the Goal) within XP. Although the distinction may follow from the nature of the elements involved (technically, a Probe contains uFF, which forces MS to locate a Goal), it is certainly odd for MS to both stop (in labeling) and not stop (in Probe-Goal) at H.

20 Here Chomsky (2013, 2015) departs from previous approaches to labeling in $\{XP,YP\}$ created via IM. So, in Chomsky (2007:23), it is pointed out that "questions arise about labeling only for XP-YP constructions. For IM, with XP raised from YP with head Y, Y is the probe, and the simplest assumption is that it remains the probe, so that XP need not be searched for a new Probe." This view is consistent with previous discussion on the impossibility that *d*XPs label (as already argued for in Chomsky 1995:256). See Donati (2006) for discussion that IM of heads can have that effect, which is adopted in Chomsky (2008:145), but reconsidered in Chomsky (2013:46).

21 For an early version of this approach to sublexical features (called "sublabels" in Chomsky 1995) and their role in agreement, the reader is referred to Chomsky (1995:268 and ff.). The agreement analysis of Chomsky (2013, 2015) suggests that the "intersection" option of Chomsky (1995) should not have been discarded. It is unclear, however, how to implement "feature intersection" in a system where Merge does not manipulate features, but only LIs (an option that was entertained for Move-F; cf. Chomsky 1995:262, 270-271, 383 fn.27 for relevant discussion).

22 Chomsky (2013, 2015) assumes that the Q feature on *which books* is unvalued, and must be valued by the Q feature on C. This idea was also adopted in Chomsky (2000:107), where it was suggested that "the wh-phrase has an uninterpretable feature analogous to structural Case for nouns, which requires it to move to its final position in an appropriate C." It is not clear that uninterpretability makes sense for features outside of the Case / Agreement systems, including Q, topic, focus, etc. (cf. Rizzi 1997, 2004, 2006, 2007). In fact, postulation of such construction-specific features does not seem explanatory (cf. Chomsky 2001:6, Chomsky 2008:151 for discussion).

stay in [Spec, TP], as that is the only way for the TP to be labeled—hence interpretable at the semantic component. The key assumption for this to work is Chomsky's (2015) claim that T is like roots (as already suggested in Chomsky 2001): from that it follows that T is too weak to label by itself, which is what makes subject stay in [Spec, TP].

(22) *[$_{CP}$ Who do you think [$_{CP}$ that $t_{Who}$ read the book ] ]?

[from Chomsky 2015:10]

The labeling analysis of ECP effects is carried over to A-bar freezing effects. Consider (23) to see this.

(23) *[$_{CP}$ Which book do they wonder [$_{CP}$ $t_{Which\ book}$ C$_Q$ [ he read $t_{Which\ book}$ ] ] ]?

[adapted from Chomsky 2015:12]

In agreement-based accounts (see section 2.2.), the problem with (23) is that *which book* is frozen after it agrees with the Q feature C is endowed with. In Chomsky (2015), the problem is that the embedded CP cannot be labeled Q after *which book* leaves the embedded [Spec, CP] position: IM leaves a copy, which is invisible, so the embedded clause is labeled by the Q feature of C alone, which yields a yes-no question interpretation that results in gibberish.

Although I will adopt a version of Chomsky's (2015) approach to A freezing, I believe the specifics suffer from shortcomings inherited from agreement-based accounts. Chomsky (2013, 2015) makes two crucial assumptions: first, the features that X and Y are endowed must be "the most prominent" ones, and the relevant dependency between them must be "Agree, not Match" (Chomsky 2013:45).[23] Both assumptions make the tacit claim that LA can see the internal structure of LIs (their lexical syntax; cf. Hale & Keyser 1993 and sub.), which is by no means obvious, given that LIs are "atoms of computation"—that is, units whose internal part is invisible to syntactic operations (cf. Chomsky 2007:6, 2008:135, 2013:41,46).

A second problem for Chomsky's (2013, 2015) analysis concerns the very idea that the prominent features of X and Y must agree for the label of {XP,YP} to be determined. It is unclear how this works if Agree requires c-command (another instantiation of MS) between Probe and Goal (Chomsky 2007:9, 2008:146). To be specific, suppose we have constructed the syntactic object {XP,YP}, depicted in (24) (taking XP to be a *d*XP, raised from within YP).

(24) {{X,ZP}, {Y,WP}}

Suppose now that X and Y are replaced by a structure built up of features, *x* and *y* being the most prominent ones:

(25) {{{**x**, {...}}, ZP}, {{**y**, {...}}, WP}}
    ↑_____↑ (Agree)

As the reader can see, the problem is straightforward: *x* and *y* cannot establish any structural dependency under MS—they are simply unable to communicate, unless we resort to additional devices (sisterhood, X-bar projections, feature percolation, etc.) or else we modify the definition of Agree. In fact, not only the feature *x* does not c-command the feature *y*: X does not c-command Y either. In a nutshell, Chomsky's (2013, 2015) analysis of feature agreement requires something that goes beyond a dependency between a Probe and a Goal in its c-command domain: it requires a broad notion of c-command, namely m-command (Chomsky 2007:9), which displays all the objections Spec-Head agreement suffered from.[24]

Taking stock, Chomsky's (2013, 2015) updated analysis requires a reformulation of LA along th lines of (26):

(26) Labeling Algorithm (POP / POP+ version)

————

[23] The reason why Match is not enough is empirical. If Match sufficed, small clauses such as (i) (where agreement can show up in languages of the Italian type), would be labelable.
    (i)        *Is [α the picture he cause of the riot ]
    (ii)       The picture is [α $t_{The\ picture}$ the cause of the riot ]
However, (i) is out, and IM is needed, a fact Chomsky (2008 and sub.) takes to indicate that LA fails to label α. An anonymous reviewers points out that small clauses can be headed (endocentric), but note that this does not affect the overall reasoning.
[24] An anonymous reviewer points out that the configurations we are considering do require Agree, but as he/she acknowledges, this is before IM takes place. Notice, however, that this pre-IM Probe-Goal dependency is orthogonal for labeling purposes, as MS relies on the final configuration—more precisely, it relies on the upmost occurrences of XP and YP, among which no Probe-Goal dependency can be established.

a. In {H, α}, H an LI, H is the label

b. If α is internally merged to β forming K (α = XP and β = YP), the label of K is
   the most prominent feature shared by α and β

Though elegant and in the very spirit of minimalist desiderata, the second clause of (26) raises theoretical and empirical problems. True, formation of {XP,YP} structures, either by EM or IM, yields an ambiguous SO, so something is required to stabilize (freeze, halt) the structure. In EM scenarios, IM of one of the Merge-mates kicks in, and this is also what motivates successive cyclic movement. But in IM scenarios, freezing is technically trickier. In the following section I discuss other options to make *d*XP stop.

## 3.2  Other ways to stabilize (freeze, halt) SPECs

This section explores three technical ways in which LA could account for the fact that *d*XP stop in derived positions, putting aside Chomsky's (2013, 2015) revamping of Spec-Head agreement. As we will see, some fare better than others (requiring less stipulations).

### 3.2.1  Structure elimination via Transfer

A first option would be to assume that one of the SOs in {XP,YP} (say, XP) is subject to Transfer, so it becomes X, allowing the LA to apply. This option has been explored in Uriagereka (2004), Obata (2010), Ott (2011), and Narita (2014), and all assume that NS literally destroys already created structure, making X's complement disappear:[25]

(27) In {XP,YP}, Transfer X's complement → {X,YP}   MS: X

This strategy works without involving agreement, but in so doing it violates the NTC (by destroying X's complement) and it also must increase the typology of phases (typically, *d*XPs are DPs or PPs, so D and P would have to count as phase heads). Furthermore, (27) makes unwanted empirical predictions. One was already discussed in fn. 8 and involves examples like (28), where α must have been transferred prior to wh-movement, but it is still pied-piped.

(28) [ [ Which book [α that John bought ] ] did you read t$_{Which\ book}$ ] ?

A second problem concerns the outcome of MS in Transfer-reduced {X,YP}. If X is a DP (or a PP), then the head of the entire SO would be D (or P). Although it may have interesting consequences in wh-movement scenarios, it would certainly not in sentences like *John left the room*, where the head would be *John*.

### 3.2.2  SubMerge / UnderMerge

A second option is that the *d*XP creates a new complement position after IM. The process is indicated in (29):

(29) a. Merge (Y,{ . . . XP . . . }) = Z = {Y,{ . . . XP . . . }}
     b. Merge (Y,XP) = Z' = {{Y,XP},{ . . . t$_{XP}$ . . . }}

This analysis has been explored by Pesetsky (2007) and Gallego & Uriagereka (2011) to account for the behavior of raising-to-object and Romance clitics. For reasons that will become clear in a moment, these authors dub the operation UnderMerge and SubMerge respectively.

There are two main problems with (29). The first one comes from the fact that (29) illustrates an operation that is not EM nor IM. Let us take wh-movement to see why.[26] Suppose that we have constructed Z = {C,{T,{ . . . }}}: Z is in the workspace, and we can merge it to W, where W is distinct from Z (EM) or within Z (IM). In either case, the outcome is

---

[25] I am somehow summarizing Obata's (2010) discussion. This author explores different options for Transfer, and concludes that it must be weak enough to allow for the relevant material to be available at later derivational stages, even if it has been subject to the relevant mappings.

[26] The same would hold for focus fronting, but not dislocation, which may involve a different kind of derivation (cf. Ott 2014).

Z' = {W,Z}. This would be the case of regular wh-movement. Now, (29) is crucially different, since there is no application of Merge (under NTC) that can form {C,XP}. IM entails that XP is taken from within Z and merged with Z, but XP in (29) is not merged with Z, but with C. For that to be possible, Merge would have to be ternary, involving Z, XP, and C.[27] The second problem for (29) is that, even if feasible, the analysis still brings us to a dead end. This is so because although the {Y,XP} chunk can be labeled, the entire SOs is still an {XP,YP} structure, and thus unlabelable.

### 3.2.3 IM of heads

A final alternative is outlined in Kitahara (2016). Details aside, this author suggests that XP moves (creating an ill-fated SPEC position) and then the head Y undergoes IM to label the resulting structure, as shown in (30):

(30) a. $\{Y,\{\ldots XP\}\}$
    b. $\{XP,\{Y,\{\ldots t_{XP}\}\}\}$
    c. $\{Y,\{XP,\{t_Y,\{\ldots t_{XP}\}\}\}\}$

In (30c), Y labels the structure, and XP becomes part of its complement. This is just like in the UnderMerge / SubMerge option, but without invoking a new type of Merge. Nevertheless, (30) still has the second problem we noted for (29): the outcome is labelable, but the compement of Y is still {XP,YP}.

Of course, the general question here is what triggers IM of Y in the first place. As I did in section 2.1., I will assume that Merge is not motivated (feature driven). So, by that logic, it is possible that Y does not raise (in successive cyclic scenarios, I assume).[28] Notice that (30) is in fact very close to the second clause of the LA in Chomsky (2007, 2008), repeated in (31) for convenience:

(31) If $\alpha$ is internally merged to $\beta$, forming $\{\alpha,\beta\}$, then the label of $\beta$ is the label of $\{\alpha,\beta\}$

In Chomsky (2007, 2008), the intended outcome is not what we obtain, which is an {XP,YP} configuration under the NTC. But Kitahara's (2016) offers a way to have the cake and eat it too.

## 3.3 Rizzi's (2015) reformulation of Criterial Freezing

In this final section, I would like to sketch the most recent account of Criterial Freezing put forward by Rizzi (2015), who follows Chomsky in taking the problem to have a labeling basis. Reviewing data like (23), Rizzi (2015:21-22) argues that XP movement can only involve maximal projections, but not intermediate (X') projections, a condition he expresses as in (32):

(32) Maximality: Phrasal movement can only involve maximal objects with a given label

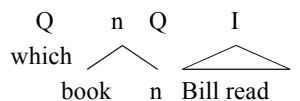[from Rizzi 2015:22]

Rizzi (2015:22) further argues that maximality of a given SO is determined "by the label of its immediately superordinate node δ: if the label of δ is different from the label of γ, then γ is maximal; otherwise it is not." Taking (23) as the case study, Rizzi (2015) provides the structure in (33), where he argues that *which book* has ceased to be a maximal object, as the mother node is also Q:

(33) I wonder . . .



---

[27] Similarly, the analysis of subject raising in Chomsky (2007:17, 2008:143,155) is also not binary. The derivation assumed in Chomsky (2008) entails that once T is merged, then C is (before the subject raises), yielding {C, {T,{EA,v*P}}}, and then {C, {EA {T,{t_EA,v*P}}}}. The key thing here is that, though ternary (since TP is substituted by {EA,TP}), subject raising does qualify as IM: EA is a term of TP, to which it merges.

[28] As Hisa Kitahara (p.c.) tells me, Chomsky (1995) was assuming something like this. In his original, projection-based, proposal, merger of XP and YP yields {H,{XP,YP}}, where H is the head of XP or YP.

```
        Q    n  Q      I
     which    ╱╲    ╱——╲
           book    n  Bill read __
```

<div align="right">[from Rizzi 2015:22]</div>

Rizzi (2015) is thus reducing freezing to a syntactic problem having to do with a constraint on Merge (IM) application. As Epstein, Kitahara, & Seely (2016) (EKS 2016 henceforth) convincingly argue, Rizzi's (2015) solution must at least assume the principles in (34):

(34) a. Every SO appearing at CI must have a label
     b. Labeling takes place in NS obligatorily and immediately whenever applicable
     c. Only maximal objects with a given label can be moved

<div align="right">[from Epstein, Kitahara, & Seely 2016:228]</div>

EKS (2016) note that (34a) and (34b) may well be necessary for independent reasons, but (34c) puts a dedicated syntactic constraint on IM, which looks far-fetched in a system where Merge is free. In addition to that, notice that the prediction embodied in (32) must take labels (*qua* projections) to be created in NS, which again deviates from the most basic form of Merge (cf. Collins 2002, Chomsky 2004).

# 4  No freezing in the syntax

This section puts forward an approach to freezing (halting) effects that does not resort to a Spec-Head agreement logic. I argue that *d*XPs in A-bar / criterial positions are not frozen in the syntax—given unbounded free Merge, this would simply require a specific stipulation on IM (*pace* Rizzi 2015). Instead, I will argue that freezing is subject to two independently needed principles: (i) the PIC, which is a direct consequence of cyclic Transfer and renders SOs in the complement domain of phase heads inaccessible and (ii) the PIF. With respect to the latter, I will follow Gallego (2009) in assuming that *d*XPs can skip a criterial position, but if they do, the relevant discourse-oriented interpretation will be lost under a configurational approach (cf. Hale & Keyser 1993 and sub., Chomsky 2001, 2008, Uriagereka 1995). As advanced in section 1, I will endorse (3), repeated below as (35), which I take to follow from Chomsky's (1986a) PFI:

(35)  Principle of Interface Freezing (PIF)
     A *d*XP is assigned INT at SEM if *d*XP occupies a phase edge

The PIF is to be related both to the duality of semantics that Chomsky (2004) relates to the two variants of Merge (EM and IM) and to the idea that dXPs in phase edges give rise to discourse-related and scopal properties. The latter proposal played an important role in Chomsky's (2001) discussion of operations of the object shift sort, which was formalized as in (36):

(36)  The EPP position of a phase Ph is assigned Int

<div align="right">[from Chomsky 2001:33]</div>

The PIF, as well as (36), is consistent with EKS's (2016) claim that halting effects can be derived from morphological or semantic principles, but not syntactic ones. If something like the PIF is entertained, then it will be necessary to assume that some {XP,YP} structures must in fact be generated and stay that way. Differently put, not all structures must be of the {X,YP} form, unless we know of some C-I principle imposing endocentricity. I do not. Recall that Chomsky (2013:43) reasonably conjectures that the LA must be satisfied so that C-I can determine the nature of SOs generated by Merge: "For a syntactic object SO to be interpreted, some information is necessary about it: what kind of object is it? Labeling is the process of providing that information." What this is saying is that {X,YP} structures may be required by some C-I principle when the relevant information involves determining whether a SO is verbal, nominal, etc. However, there is nothing in this logic, in and of itself, that precludes {XP,YP} insofar as this structure provides the right kind of interpretation at C-I. For the most part, {XP,YP} structures illustrate discourse-related constructions: questions, relative clauses, and phenomena involving new-old information. Under the reasonable assumption that a "construction" is complex, whereas the "type" of an SO is not necessarily so, I explore the hypothesis that

{XP,YP} can be generated—and perhaps must, if we want C-I to convey interpretations that go beyond the type of an SO. In this vein, I will suggest that, for the most part, {XP,YP} structures emerge in root structures, for which there are independent reasons to assume that they are unlabeled (see fn. 7).

In what follows, I will first explain how Transfer and Labeling Theory can account for A cases of freezing, and then I will consider the A-bar cases, which are those corresponding to Rizzi's (2006, 2007) Criterial Freezing.

## 4.1  A-Freezing: C and T as a discontinous element

In section 2.1., I introduced a design trait of Chomsky's (2000 and sub.) derivational model, namely the idea that there is a periodic forgetting of the structure that Merge builds. Assuming the general phase-based architecture, the operation Transfer cashes out the complement domain of a phase head H, leaving the head itself and its specifier (the edge) for subsequent operations in the next phase. This is shown in (37), where $\alpha$ is the complement of H.

(37) [ $\beta$ [ H (Probe) $[_\alpha \ldots$ XP (Goal) $\ldots$ ] ] ]

The consequence of this is that Transfer renders $\alpha$ (and all it contains) no longer accessible. Although this does not correspond to standard derived-island effects, it does have the desired effect without additional stipulations: it covers the Subject Criterion, the Activity Condition, etc. Let us consider subject freezing (EPP, ECP, *that*-t effects, etc.) in more detail and how it can be accounted for under a non Spec-Head-agreement based approach. To do this, let us go back to Chomsky's (2015) suggestion that subjects halt in [Spec, TP] because T is too weak to label the TP. Before going ahead, a question that immediately arises is how a question like (38) can be formed:

(38) Who voted Trump?

The relevant aspect of (38) is that "TP" cannot be labeled, given the logic of Chomsky (2015). If *Who* is in [Spec, CP], and T cannot label on its own, TP should remain label-less and yield deviance. But (38) is fine. The problem does not arise with non-subject wh-questions (where the subject must move to [Spec, TP] and stay there), so (38) qualifies as an ECP effect. Notice that (38) is also problematic under Chomsky's (2008) analysis of parallel movement, whereby the wh-phrase moves from the vP to [Spec, TP] and [Spec, CP] in parallel, as depicted in (39)—this is so as the A-bar occurrence is pronounced in [Spec, CP]  (where labeling takes place, unproblematically), but the A occurrence in the [Spec, TP] is not pronounced, which should render it invisible.

(39) [$_{CP}$ Who [$_{TP}$ Who [$_{vP}$ Who [ v [ voted Trump ] ] ] ] ]?
        ↑      ↑_____|
        |_____|

In a sense, Chomsky (1986b) already considered the problem in (38) when accounting for the asymmetry in (40):

(40) a. Who likes John?
    b. Who does John like?

[from Chomsky 1986b:48]

Chomsky (1986b) assumed that wh-movement to the CP does not take place in the case of subjects, so *Who* moves to [Spec, TP], but not to [Spec, CP], in (40a), which would also explain the fact that non-subject wh-phrases can circumvent wh-island effects, as (41) reveals:

(41) What do you wonder [$_{CP}$ who saw *t* ] ?

[from Chomsky 1986b:48]

Chomsky's (1986b) analysis made it possible for *What* to move to [Spec, CP], since this position is not occupied by *who*. Interesting as this is, it does does not immediately solve the problem in (38), as [Spec, TP] is indeed occupied, but by a copy. In order to solve this puzzle, and extend the solution to other ECP cases, I would like to argue that, in structures like (38), C and T are the same category in the lexicon of English. Let me elaborate. C and T are typically analyzed as distinct functional items, but much research has shown that the interaction between C and T is manifold,

both syntactically and morphologically (cf. Pesetsky & Torrego 2001 and references therein). In Chomsky (2004), it is argued that nominative is actually assigned by the C-T configuration (not T alone), which Chomsky (2007) elaborates on to suggest that φ-features are generated in C and then passed down to T through a process of feature inheritance.

Gallego (2014) assumes this much and suggests that T is actually a copy of C in languages like English. In other words, Gallego (2014) suggests that what is usually regarded as two independent heads should be conceived of as a non-trivial chain.[29] The proposal nicely recasts Chomsky's (2004) claim that C and T work together to assign nominative Case and, at the same time, dispenses with feature-inheritance (simply because the features that are in C must also be in all its copies). With this in mind, the representation of (38) should actually be as in (42):

(42) $[_{CP}$ Who $[$ C $[_{vP}$ t$_{Who}$ v $[$ voted Trump $]$ $]$ $]$ $]$?

In this analysis, *Who* raises directly to the specifier of C (taking T and C to be one and the same). Under Chomsky's (2013, 2015) approach, we could argue that the TP (CP, for us) is labeled because the φ-features of the overt DP can be used to match those of T (C), but in the present account this is not necessary: *Who* raises from [Spec, vP] for labeling reasons, and the first available position is [Spec, CP]. The question now is why (42) is not out, given that the final SO is {XP,YP}. But notice that the problem is more general: Why isn't (43) out?

(43) Many people voted Trump

One way in which (43) could be licensed is by making C undergo IM so that it creates a discontinuous object. This is in accord with the proposal in 3.2.3. and should be seen as the clause-typing strategy by default, when no discourse-related or scopal properties are involved. The analysis can be invoked in order to account for (38), as shown in (44):

(44) $[$ C $[$ Who $[$ t$_C$ $[_{vP}$ t$_{Who}$ v $[$ voted Trump $]$ $]$ $]$ $]$ $]$?

Notice that this analysis opens the door to understand EPP and ECP effects as two sides of the same coin. When C moves, it leaves a copy, which is invisible to LA. Because of that, the DP must stay in [Spec, TP], or else TP would be unlabeled. Notice that this departs from Chomsky's (2015) analysis, which argues that English T is too weak to label. What I suggest is that this labeling inertness does not come from feature-strength (in the sense of Chomsky 1993), but from T being a copy.[30]

Let us consider next ECP effects involving subject extraction across C. Suppose we have generated (45):

(45) Someone says $[$ C $[$ many people t$_C$ $[$ t$_{many\ people}$ voted Trump $]$ $]$ $]$

Again, *many people* must remain in [Spec, CP], and C, which can be realized as *that*, may or may not attract the subject to the phase edge. This would wrongly predict that *many people* can be extracted across *that*, contrary to fact. Therefore, *that*-less clauses must involve a different derivation. Let us suppose that C movement is followed by an optional PF insertion rule that is sensitive to the root / embedded distinction. Thus, if C moves in a root clause, it cannot be spelled-out as *that*, but it can in an embedded domain. Following this reasoning, suppose C does not move in embedded *that*-less clauses. This predicts the absence of *that* and the fact that the subject can abandon the [Spec, TP] position: since C stays in-situ (the C-T discontinuous object is not created), it labels the structure.

This analysis covers the Romance facts discussed in the literature, which are ECP-free (Gallego 2010 and references therein). Thus, following Gallego's (2014) reasoning, if C and T are in fact different lexical items in, say, Spanish, then T will never count as a copy. This explains why the EPP does not hold in Romance (subjects move to [Spec, TP] optionally, giving rise to a discourse-related interpretation, discussed by Rizzi 2006, 2007 and Uriagereka 2008) and, therefore, why subjects can be extracted.[31]

---

[29] This takes both heads, though not identical on conceptual grounds, to be one and the same *in the lexicon*. Similar ideas have been explored in the vP domain, where certain heads have been said to 'bundle' (cf. Pylkännen 2008, Harley 2013).

[30] A prediction made by this analysis is that {XP,YP} structures cannot be generated. As Chomsky (2008, 2013), the standard cases require for either XP or YP to move, but not both. I leave open what this might follow from.

[31] An anonymous reviewer asks what head movement boils down to in this account, and refers to EKS (2016) for a specific formulation of head movement in Internal Pair-Merge terms. The proposal in Gallego (2014) presupposes that head movement is just an instance of standard IM, with no need to resort to Pair-Merge, whose dependency-specific nature (adjunction) raises yellow flags. Space limitations prevent me from going into details, but as Chomsky et al. (2017) suggest, the label "head move-

Summarizing so far, I have argued that, in a system adopting some version of Phase Theory, the most natural way to account for why SOs occupying the complement domain of phase heads is by invoking cyclic Transfer. This said, nothing in the current system precludes that a *d*XP in [Spec, TP] raises to [Spec, CP], so cyclic Transfer does not provide an ultimate answer. In order to account for the fact that a *d*XP cannot abandon [Spec, TP], I have argued that C and T are the same element in the lexicon of certain languages (English), and that what we call T is actually a copy of C. Chomsky (2015) argues that English T is weak to label, so that the subject must stay there to label. I have kept the basics of his analysis, but without invoking feature strength. Instead, I have proposed that T's copy status is what makes it unable to label. If nothing else, the alternative I am suggesting does not need assumptions beyond Merge and the copy theory of movement.

## 4.2  Criterial Freezing: Principle of Interface Freezing

So far, I have not discussed A-bar freezing much. Let us go back to the representative case in (5), repeated here:

(46) *Which book does Bill wonder [$_{CP}$ t$_{Which book}$ C [ she read t$_{Which book}$ ] ]?

[from Rizzi 2006:112]

We have already seen how standard approaches to freezing tackle (46) and similar facts (see sections 2.2. and 3.3.). Here I argue that there can be no syntactic freezing (halting): like any other case of IM, wh-movement in (46) is allowed to apply in NS, so if a problem emerges it must be due to independent S-M or C-I requirements. Gallego (2009) outlines such an approach by capitalizing on Chomsky's (2001) analysis of discourse-oriented interpretations emerging in phase edges, which was phrased as follows:

(47) The EPP position of a phase Ph is assigned Int

[from Chomsky 2001:33]

If one endorses (47), *d*XPs in the specifier of a phase head are assigned a discourse-oriented interpretation—not because of feature checking, but simply because of their position. In brief, (47) amounts to *d*XPs being interpreted as Q-operators, Rel-operators, focus-operator, etc. for the same sort of reason an XP receives a theta-role in Hale & Keyser's (1993 and sub.) framework, namely because of the structure they are part of. Clearly, things cannot be so simple, for the specific "discourse-oriented" and "thematic" interpretation are not easy to dertermine. Thus, for instance, the interpretation of *Horatius* is not the same in (48a), (48b), and (48c), although the positions this DP occupies plausibly qualify as a specifier in a {XP,YP} configuration:

(48) a. Horatius held the bridge against a whole army
    b. Horatius did not fear the army in the bridge
    c. There was Horatius in the bridge against a whole army

As noted in the literature (cf. Acedo-Matellán & Mateu 2014 and references therein), the interpretation of a given XP depends on the nature of its Merge-mate: AGENT is assigned to XP if it is merged with a vP headed by v$_{CAUSE}$ or v$_{DO}$. GOAL is assigned to XP if it is first merged with a terminal-coincidence preposition, and so on and so forth. The same is trivially true in the CP domain, where the interpretations assigned to *X Roman* in (49) are all different:

(49) a. Which Roman defended the bridge?
    b. THIS ROMAN defended the bridge, not THAT ONE
    c. The Roman who defended the bridge

However the interpretations at the vP and CP levels obtain, I assume that this happens in the interpretive components. That is to say, NS is not sensitive to notions like topic, focus, agent, theme, subject, object, and the like: there is a computational system that can take elements from a lexicon to yield SOs. Interestingly, these notions are traditionally regarded as "syntagmatic," as they only appear once the syntactic computation has generated some structure—a lexicon does not

——

ment" probably covers different empirical scenarios, and thus require different technical implementations. In any event, and just to address the reviewer's concerns, head movement of C should be regarded here as any other instance of IM.

contain LIs that are inherently themes or foci, let alone features that collapse those notions. The point was already made in Chomsky (1965):

> The notion "Subject," as distinct from the notion "NP," designates a *grammatical function* rather than a *grammatical category*. It is, in other words, an inherently relational notion. We say, in traditional terms, that in (1) *sincerity* is an NP (not that it is the NP of the sentence), and that it is a (functions as) the *Subject-of* the sentence (not that it is a Subject). Functional notions like "Subject," "Predicate" are to be sharply distinguished from categorial notions such as "Noun Phrase," "Verb," a distinction that is not to be obscured by the occasional use of the same term for notions of both kinds.   [from Chomsky 1965:68]

From this perspective, criterial freezing can be seen as a ban on *d*XPs receiving multiple interpretations.[32] That is to say, just like an XP does not receive more than one theta-role, a *d*XPs does not receive more than one discourse-oriented interpretation. This makes sense in the case of thematic interpretations, as XPs are EM-ed only once—which would be a way to recast the Theta Criterion (cf. Chomsky 1981). However, in the case of discourse-related interpretations, things are different, as IM can apply more than once. It is at this point that the Deep Structure / Surface Structure distinction becomes relevant: a discourse-oriented interpretation piggy-backs on the final surface position of *d*XP (the one feeding the S-M systems).[33]

All of this is compatible with a rather conservative view of the PFI. As already advanced, in the case that concerns us, I would like to approach the data from the PIF, a subcase of the PFI:

(50)  Principle of Interface Freezing (PIF)
       A *d*XP is assigned INT at SEM if *d*XP occupies a phase edge

The PIF is an interface condition, not a constraint on derivations. Such constraints are in fact unstatable in a free-Merge system, unless we introduce additional stuff (features, projections, indices, etc.). Rizzi (2006) provides one empirical argument to defend that (46) is not ruled out by interpretve principles. For him the problem is syntactic. Rizzi (2006) thus shows that Italian allows for contrastive focus to be assigned either in-situ or ex-situ, in the Left Periphery (which is possible Spanish too). Rizzi (2006) offers the pair in (51) to claim that criterial freezing cannot be reduced to interpretive matters:

(51) a. Mi       domandavo [$_{CP}$ quale  RAGAZZA$_i$ C avessero   scelto t$_i$], non  quale  ragazzo              (Italian)
       CL-me wonder-1.SG    which GIRL            had-3.PL  chosen    not  which  boy
       'I wonder which GIRL they had chosen, not which boy.'
     b. *[$_{CP}$ Quale  RAGAZZA$_i$ C mi       domandavo  [$_{CP}$ t$_i$ avessero  suelto t$_i$], non  quale  ragazzo              (Italian)
           which GIRL            CL-me wonder-1.SG        had-3.PL  chosen    not  which  boy
       'Which GIRL do I wonder had chosen, not which boy?

[from Rizzi 2006:113]

Rizzi (2006:113) argues that "a wh-phrase in an embedded question can be contrastively focused in its criterial position, in the embedded C system, but it cannot be moved to the left periphery of the main clause […] as contrastive focus is clearly compatible with a wh-phrase (see [51a]), it does not seem plausible to assume that [51b] is ruled out for interpretive reasons." As Gallego (2009) argues, the point is well-taken, but not conclusive. Rather, the specific status of sentences like (46) follows form the fact that the lexical intricacies of *wonder*, which is unique in selecting

---

[32] Gallego (2009) discusses the possibility that both PF and LF may impose some uniqueness constraint, requiring that XPs are interpreted at both components only once: at PF, only one copy is pronounced, and at LF, chains contain only one EM-interpretation and one IM-interpretation. Cf. Rizzi (2006:128) for discussion of situations where multiple interpretations of the criterial type are allowed through a process of head movement that creates a cluster of criterial heads.

[33] An anonymous reviewer alludes to Rizzi's (2006) reasoning about there being a lower bound below which no argument can be merged, from which he derives the lack of movement into theta-positions. Rizzi (2006) further suggests that a similar constraint applies in the upper domain of a clause (the criterial positions), beyond which movement cannot take place. The reviewer wants to know if I adopt a similar position. The answer is negative. As just pointed out, to the extent that Chomsky's (2004) duality of interpretation conjecture is on track, then it follows that theta-roles are restricted in the same manner EM is: It can only apply once, which provides a clear way to recast the Theta Criterion. The case of IM is murkier, as its application is potentially unbounded. What this paper argues is that discourse-oriented (scopal, etc.) interpretations are restricted as there must be a final application of IM (one affecting the upmost copy of the chain) that feeds Transfer.

interrogative clauses. In Spanish, only two verbs display such behavior (cf. Suñer 1999:2154), and their meaning is that of *wonder* / *ask*. To see how this is relevant, compare (51) with (52):

(52) a. María ha       dicho [$_{CP}$ qué   CHICA$_i$ C han       elegido t$_i$], no   qué   chico            (Spanish)
     María have-3.SG said       which GIRL     have-3.PL chosen   not  which boy
     'María has said which girl they have chosen, not which boy'
   b. [$_{CP}$ Qué   CHICA$_i$ C ha       dicho María [$_{CP}$ que t$_i$ han       elegido t$_i$], no  qué    chico?   (Spanish)
        which GIRL     have-3.SG said María    that have-3.PL chosen   not which boy
     'Which girl has María said that they have chosen, not which boy?'

To my mind, (51) and (52) are analogous in the relevant respect. The only difference concerns the matrix verb: unlike *wonder* and *ask*, *say* does not necessarily take a [+Q] complement. Consider next (53), from Gallego (2009), which is ungrammatical in Spanish:

(53) *Me     pregunto    [$_{CP}$ han       elegido a qué CHICA], no   a qué   chico            (Spanish)
     CL-me wonder-1.SG   have-3.PL chosen  to what girl      not to what  boy
     'I wonder what GIRL they have chosen, not what boy'

[from Gallego 2009:47]

Here, the wh-phrase with the focused NP *CHICA* (Eng. 'girl') stays in-situ. Taking Rizzi's (2006) view on freezing at face value, it is not immediately obvious what the problem in (53) is—the NP has not raised to a criterial position, and Spanish can have both wh-phrases and contrastive focused XPs in-situ. Again, this suggests that the problem has to do with the idiosyncratic nature of *preguntarse* (Eng. 'wonder'), not with a general constraint on derivations.

Interestingly, EKS (2016) provide an explanation of the facts that is along the lines of the PIF. They start by making the following assumptions about the C of interrogative sentences like that *wonder* selects:

(54) a. There is only one C$_Q$ in the (English) lexicon, appearing in both yes / no and wh-interrogatives,
     b. every syntactic object must be labeled at CI,
     c. a CP with the label C$_Q$, unaccompanied by a wh-specifier, is interpreted as a yes / no -question at CI; and
     d. a CP with the label Q, when Q is shared by the two heads C$_Q$ and WH$_Q$ is interpreted as a wh-question at CI

[from EKS 2016:229]

They thus assume that a yes / no-question has the underlying representation of (55):

(55) [$_α$ C$_Q$ [$_{TP}$ Horatius held the bridge ] ]

EKS (2016) argue that the label of α is C, but the S-M representation of (55) is ruled out if it has a neutral or falling intonation. In English, matrix yes / no-questions require either T-to-C movement or rising (question) sentential prosody, as in (56) (intonation is signaled in square brackets):

(56) a. Did Horatius hold the bridge?
     b. Horatius held the bridge [↑↑]?

Consider, for the punch line, what happens if α is embedded. In those cases, EKS (2016) argue that since C does not contain a wh-specifier, α must be interpreted as a yes / no-question.

(57) *I wonder [$_α$ C$_Q$ [$_{TP}$ Horatius held the bridge ] ]

The problem, for these authors, is also interface-rooted: (57) is out for reasons ultimately rooted in the S-M systems, since both T-to-C movement and rising intonation in English embedded clauses are not possible. EKS (2016) reason that α in (57) is also ruled out for C-I reasons, since the composed representation of the matrix clause and embedded α yields gibberish. In their own words:[34]

---

[34] EKS (2016) suggestion is sensible, but the alleged C-I problem could also be attributed to an S-M constraint requiring that, when embedded, the C$_Q$ of yes / no-questions needs to be occupied by an overt Q element.

One possibility regarding its status as gibberish is as follows: The CP headed by C Q is itself interpreted as a yes / no- question and so would be interpreted as: "Answer me this: Does John like this dog?" that is, a performative request made of the speaker's interlocutor for a specific kind of information As such, embedding it, as in I wonder John left yields an interpretation like: "I wonder, 'Answer me this, Did John leave?'" This is anomalous to the extent that one cannot wonder a request for information.                    [from EKS 2016:236, fn.18]

Turning now to (46), EKS (2016) follow Chomsky (2013, 2015) in taking it to be a C-I (labeling) problem: the copy of the wh-phrase is invisible to LA. Consequently, the label of $\alpha$ is $C_Q$, which satisfies selection by *wonder*, but cannot be interpreted as a wh-question, given the assumptions in (54).[35]

In sum, the line of reasoning EKS (2016) advocate for is, though technically different, consistent with the PIF in taking the relevant problem of sentences like (46) to be found at the interfaces, not in the syntax.


# 5 Conclusions

This paper has discussed the nature of freezing (halting) effects. I have argued that freezing (especially A-bar / criterial freezing) should be regarded as an interface issue, not a syntactic one. If unbounded free Merge is adopted (cf. Chomsky 2004 and sub.), then its application is unconstrained as long as it adheres to efficiency principles, such as NTC, cyclic Transfer, and the like. From this, it follows that, if a *d*XPs occupies a phase edge, then it can be subject to further applications of Merge. Notice that this is not to say that the criterial facts are wrong, but that the problem is not syntactic, unless we enrich syntax (and thus UG) with features, Spec-Head agreement, projections, and similar devices that raise technical qualms. The relevant trouble-making (freezing) IM configurations can indeed be created, and I assume they yield discourse-related interpretations at the semantic component (cf. Chomsky 2004, 2007, 2008), in the same way EM configurations yield thematic ones. This has been formalized as the PIF:

(58)  Principle of Interface Freezing (PIF)
      A *d*XP is assigned INT at SEM if *d*XP occupies a phase edge

To repeat, the PIF is not constraining the way Merge applies. Under (58), alleged freezing effects should be deducible from the more general demands imposed by the C-I systems, not by Spec-Head agreement mechanisms or constraints on Merge that go beyond principles of computational efficiency. As I have shown, this naturally covers situations in which a *d*XP is removed from an edge (causing the absence of the relevant discourse-interpretation at that edge) and those where a *d*XP cannot leave a given edge (an effect I have attributed to lexical intricacies, not syntactic ones).

The interface-based solution explored in this paper opens the door to a configurational approach the cartographies. As pointed out in passing, standard lexical categories (N, V, P, etc.) and inflectional affixes (Asp, T, C, etc.) are paradigmatic, whereas topic, focus, and the like are syntagmatic—they only show up in the syntax. This very issue was brought up by Chomsky (1995:349 and ff.) in order to dispense with agreement projections, and although the literature has emphasized Chomsky's objections with the uninterpretable status of these categories, it is important to recall that he also questioned them because of their theory internal flavor: AgrP is inherently relational (paradigmatic).

We have seen that Chomsky (1965) made this point very explicity. And the point is, at the relevant level of abstraction, the one Hale & Keyser (1993) make in their approach to theta-roles, which are also relational entities: "there *are* no thematic roles. Instead, there are just the relations determined by the categories and their projections, and these are limited by the small inventory of lexical categories and by Unambiguous Projection" (p.68). Technically, there is nothing wrong

---

[35] If the CP selected by *wonder* is actually not embedded, but paratactic in nature (like all indicative dependents in Romance, for instance), then we could explore the possibility that *d*XPs only stop in root contexts, even if they appear to be embedded in the surface. True embedding (subjunctive in Romance) precludes embedded wh-sentences and also restricts the availability of contrastive focus, as pointed out in the literature (cf. Torrego & Uriagereka 1992). Topics are different (they can be embedded), but this may be a welcome result if topicalization involves a different derivation. As Noam Chomsky indicates through personal communication, in expressions like *I think that [ [that books like that], John will never read]*, the internal bracketed phrase is set off prosodically and behaves almost like a root sentence, suggesting that internal topicalization and left dislocation are both quite different from successive-cyclic movement. I leave the exploration of this possibility for future research.

with lexicalizing grammatical functions,[36] pragmatic notions, or thematic roles (this is in fact common practice; cf. Cardinaletti 2004, Haegeman & Hill 2010, Speas & Tenny 2003, Ramchand 2008; cf. Bruening 2010 for some criticism), the point is that all these categories are emergent: they only appear in a syntactic environment, so recycling them as LIs blurs the paradigmatic / syntagmatic cut, and raises non-trivial questions concerning the architecture of the Faculty of Language (cf. Uriagereka 2008 for much relevant discussion). Further research is certainly needed to clarify these matters and to achieve an account that, taking advantage of the impresive results provided by cartographic work, offers a principled explanation of the facts described.

---

[36] Notice that I do not claim that these notions are not 'grammaticalized'. There can be dedicated morphemes associated to these phenomena. What I argue against is the idea that inherently relational notions can be analyzed as LIs.

# Bibliography

Adger, D. & P. Svenonius. 2011. Features in Minimalist Syntax. In Boeckx, C. (ed.) *The handbook of linguistic minimalism*, 27-51. Oxford: Blackwells.

Bach, E. & G. Horn 1976. Remarks on 'Conditions on Transformations'. *Linguistic Inquiry* 7: 265-361.

Benincà, P. & N. Munaro. 2010. *Mapping the Left Periphery*. Oxford: Oxford University Press.

Blümel, A. 2014. Propagating Symmetry – Case Studies in exocentric Syntax. Ph.D. thesis, Johann Wolfgang Goethe University Frankfurt.

Bošković, Ž. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 27: 1-68.

Bošković, Ž. 2008. On the operator freezing effect. *Natural Language and Linguistic Theory* 26: 249-287.

Bosque, I. & Á.J. Gallego. 2014. "Reconsidering Subextraction: Evidence from Spanish". *Borealis* 3: 223- 258.

Broekhuis, H. 2006. Extraction from subjects: some remarks on Chomsky's 'On phases'. In *Organizing grammar. Studies in honor of Henk van Riemsdijk*, H. Broekhuis et al. (eds.), 59-68. Berlin: Mouton de Gruyter.

Chomsky, N. 1964. *Current Issues in Linguistic Theory*. The Hague: Mouton.

Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. 1973. Conditions on transformation. In *A festschrift for Morris Halle*, S. Anderson & P. Kiparsky (eds.), 232-286. New York: Holt, Renehart and Winston. [reprinted in Chomsky 1977b]

Chomsky, N. 1986a. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. 1986b. *Knowledge of language. Its nature, origin, and use*. New York: Praeger.

Chomsky, N. 1993. A minimalist program for linguistic theory. In *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, K. Hale and S. J. Keyser (eds.), 1-52. Cambridge, MA: MIT Press.

Chomsky, N. 1995. Categories and transformations. In *The minimalist program*, 219-394. Cambridge, MA: MIT Press.

Chomsky, N. 2000. Minimalist inquiries: The framework. In *Step by step. Essays on minimalist syntax in honor of Howard Lasnik*, R. Martin et al. (eds.), 89-155. Cambridge, MA: MIT Press.

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

Chomsky, N. 2004. Beyond explanatory adequacy. In *Structures and beyond. The cartography of syntactic structures (vol. 3)*, A. Belletti (ed.), 104-131. Oxford, NY: Oxford University Press.

Chomsky, N. 2005. Three factors in language design. *Linguistic Inquiry* 36: 1-22.

Chomsky, N. 2007. Approaching UG from below. In *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, U. Sauerland & H-M. Gärtner (eds.), 1-30. Berlin: Mouton de Gruyter.

Chomsky, N. 2008. On phases. In *Foundational issues in linguistic theory. Essays in honor of Jean-Roger Vergnaud*, C. Otero et al. (eds.), 134-166. Cambridge, MA: MIT Press.

Chomsky, N. 2013. Problems of Projection. *Lingua* 130: 33-49.

Chomsky, N. 2015. Problems of Projection. Extensions. In *Structures, Strategies and Beyond*, E. di Domenico et al. (eds.), 1-16. Amsterdam: John Benjamins.

Chomsky, N., Á.J. Gallego & D. Ott. 2017. Generative Grammar and the Faculty of Language: Insights, Questions, and Challenges. Ms. available in Lingbuzz [http://ling.auf.net/lingbuzz/003507]

Cinque, G. 1999. *Adverbs and functional heads. A cross-linguistic perspective*. Oxford, NY: Oxford University Press.

Cinque, G. 2004. Issues in adverbial syntax. *Lingua* 114: 683-710.

Collins, C. 2002. Eliminating labels. In *Derivation and explanation in the minimalist program*, S. Epstein & T. Seely (eds.), 106-132. Malden, MA: Blackwell.

Donati, C. 2006. "On wh-head movement". In *Wh-movement: moving on*, L. Cheng and N. Corver (eds.), 21-46. Cambridge, MA: MIT Press.

Epstein, S. 1992. Derivational Constraints on A-bar Chain Formation. *Linguistic Inquiry* 23: 235-259.

Epstein, S. & D. Seely. 2002. Rule applications as cycles in a level-free syntax. In *Derivation and explanation in the minimalist program*, S. Epstein & D. Seely (eds.), 65-89. Malden, MA: Blackwell.

Epstein, S. & D. Seely. 2006. *Derivations in minimalism: Exploring the elimination of A-chains and the EPP*. Cambridge: CUP.

Epstein, S., H. Kitahara, & D. Seely. 2016. What do we wonder is not syntactic? In S. Epstein et al. (eds.), *Explorations in Maximizing Syntactic Minimization*, 222-239. London: Routledge.

Freidin, R. 1999. Cyclicity and Minimalism. In S. D. Epstein & N. Hornstein, *Working Minimalism*, 95-126. Cambridge, MA: MIT Press.

Gallego, Á.J. 2009. On Freezing Effects. *Iberia* 1: 33-51.

Gallego, Á.J. 2010. *Phase Theory*. Amsterdam: John Benjamins.

Gallego, Á.J. 2012. *Phases. Developing the Framework*. Berlin: De Gruyter.

Gallego, Á.J. 2014. Deriving Feature Inheritance from the Copy Theory of Movement. *The Linguistic Review* 31: 41-71.

Gallego, Á.J. & J. Uriagereka. 2011. Dos tipos de argumentos y la distinción indicativo vs. subjuntivo. *Cuadernos de la ALFAL* 3: 188-199.

Gallistel, C.R. & A.P. King. 2009. *Memory and the Computational Brain: Why Cognitive Science will Transform Neuroscience*. Malden, MA: Wiley-Blackwell.

Grohmann, K. 2003. *Prolific Peripheries*. Amsterdam: John Benjamins.

Haegeman, L. & V. Hill. 2010. The syntacticization of discourse. Ms., Ghent University and University of New Brunswick-SJ.

Hale, K. & S. Keyser. 1993. On the argument structure and the lexical expression of syntactic relations. In *The view from building 20: Essays in linguistics in honor of sylvain bromberger,* K. Hale & S. Keyser (eds.), 53-109. Cambridge, MA: MIT Press.

Hale, K. & Keyser, S. 1997. The limits of argument structure. In *Theoretical issues at the morphology-syntax interface*, A. Mendikoetxea & M. Uribe-Etxebarria (eds.), 203-230. Bilbao: Servicio de Publicaciones de la UPV.

Hale, K. & S. Keyser. 1998. The basic elements of argument structure. In *MIT working papers in linguistics* 32: Papers from the Upenn/MIT roundtable on argument structure, H. Harley (ed.), 73-118. Cambridge, MA: MIT Press.

Hale, K. & S. Keyser. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, MA: MIT Press.

Hiraiwa, K. 2005. Dimensions of symmetry in syntax: Agreement and clausal architecture. PhD Dissertation, MIT.

Hornstein, N. 2001. *Move! A minimalist theory of construal*. Oxford: Blackwell.

Hornstein, N. 2009. *A theory of syntax: Basic operations and the minimalist program*. Cambridge: Cambridge University Press.

Huang, J. 1982. Logical relations in Chinese and the theory of grammar. Ph.D Dissertation, MIT.

Kayne, R. 1989. Facets of Romance past participial agreement. In *Dialectal Variation and the Theory of Grammar*, P. Benincà (ed.), 85-103. Foris: Dordrecht

Kitahara, H. 2016. External Merge of Internal Argument DP to VP and Its Theoretical Implications. In S. Epstein et al. (eds.), *Explorations in Maximizing Syntactic Minimization*, 195-200. London: Routledge.

Lasnik, H. 1999. *Minimalist analysis*. Oxford: Blackwell.

Lasnik, H. 2006. Conceptions of the cycle. In *Wh-movement: moving on*, L. Cheng & N. Corver (eds.), 197-216. Cambridge, MA: MIT Press.

Lasnik, H. & M. Saito. 1992. *Move α: conditions on its applications and outputs*. Cambridge, MA: MIT Press.

López, L. 2009. *A derivational syntax for* information structure. Oxford: Oxford University Press.

Moro, A. 2000. *Dynamic antisymmetry. Movement as a symmetry breaking phenomenon*. Cambridge, MA: MIT Press.

Murphy, E. 2015. The brain dynamics of linguistic computation. *Frontiers in Psychology* 6: 1515.

Narita, H. 2014. *Endocentric Structuring of Projection-free Syntax*. Amsterdam: John Benjamins.

Nunes, J. 2004. *Linearization of Chains and Sideward Movement*. Cambridge, MA: MIT Press.

Obata, M. 2010. Root, successive-cyclic and feature-splitting internal Merge: Implications for feature-inheritance and transfer. Ph.D. Dissertation, U.Michigan.

Ott, D. 2011. A note on free relative clauses in the theory of phases. *Linguistic Inquiry* 42: 183-192.

Ott, D. 2014. 2014. An ellipsis approach to Contrastive Left-dislocation. *Linguistic Inquiry* 45: 269-303.

Pesetsky, D. 2007. Undermerge . . . and the secret geniti ve inside every Russian noun. Handout of a talk at Formal Approaches to Slavic Linguisitcs 16.

Pesetsky, D. & E. Torrego. 2001. "T-to-C movement: Causes and consequences". In *Ken Hale: A life in language*, M. Kenstowicz (ed.), 355-426. Cambridge, MA: MIT Press.

Pylkännen, L. 2008. *Introducing Arguments*. Cambridge, MA: MIT Press.

Richards, M. 2007. On feature inheritance: An argument from the phase ompenetrability condition. *Linguistic Inquiry* 38: 563-572.

Rizzi, L. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.

Rizzi, L. 1996. Residual verb second and the *wh*-criterion. In *Parameters and functional heads: essays in comparative syntax*, A. Belleti & L. Rizzi (eds.), 63-90. Oxford, NY : Oxford University Press.

Rizzi, L. 1997. "The fine structure of the left periphery". In *Elements of grammar. Handbook in generative syntax*, L. Haegeman (ed.), 281-337. Dordrecht: Kluwer.

Rizzi, L. 2000. *Comparative syntax and language acquisition*. London: Routledge.

Rizzi, L. 2001. On the position "Int(errogative)" in the left periphery of the clause. In *Current studies in Italian syntax offered to Lorenzo Renzi*, G. Cinque & G. Salvi (eds.), 287-296. Amsterdam: Elsevier.

Rizzi, L. 2004. Locality and left periphery. In *Structures and beyond. The cartography of syntactic structures (vol. 3)*, A. Belletti (ed.), 223-251. Oxford, NY: Oxford University Press.

Rizzi, L. 2006. On the form of chains: Criterial positions and ECP effects. In *wh-movement: Moving on*, L. Cheng & N. Corver (eds.), 97-133. Cambridge, MA: MIT Press.

Rizzi, L. 2007. On some properties of rciterial freezing. *CISCL working papers on language and cognition* 1: 145-158.

Rizzi, L. 2015. Notes on Labeling and Subject Positions. In *Structures, strategies and beyond*, E. di Domenico et al. (eds.), 17-46. Amsterdam: John Benjamins.

Ross, J. 1967. Constraints on Variables in Syntax. PhD Dissertation, MIT.

Rubio, C. 2014. Syntactic Constraints on Topicalization Phenomena. PhD Dissertation, UAB.

Šimík, R. 2011. The elimination of formal wh-features and a theory of free wh-movement. Ms., Universität Potsdam.

Speas, M. & C. Tenny. 2003. Configurational properties of point of view roles. In *Asymmetry in grammar*, A.M. Di Sciullo (ed.), 315-344. Amsterdam: John Benjamins.

Starke, M. 2001. Move dissolves into Merge: A theory of locality. Ph.D. Dissertation. University of Geneva.

Stepanov, A. 2001. Cyclic domains in syntactic theory. Ph.D. Dissertation, UConn.

Suñer, M. 1999. La interrogación indirecta. In I. Bosque I. & Demonte (coords.), *Gramática Descriptiva de la Lengua Española*, 2149-2195. Madrid: Espasa.

Takahashi, D. 1994. *Minimality of Movement*. PhD. Dissertation, University of Connecticut.

Torrego, E. 1985. On empty categories in nominals. Ms., UMass Boston.

Torrego, E. & J. Uriagereka. 1992. Indicative dependents. Ms., UMass Boston/UMD.

Uriagereka, J. 1995. An F position in Western Romance. In *Discourse configurational languages*, K. Kiss (ed.), 153-175. Oxford, NY: Oxford University Press.

Uriagereka, J. 1999. Multiple spell-out. In *Working minimalism*, N. Hornstein & S. Epstein (eds.), 251-282. Cambridge, MA: MIT Press.

Uriagereka, J. 2008. *Syntactic anchors. On semantic structuring*. Cambridge: Cambridge University Press.

Uriagereka, J. 2011. Derivational cycles. In *The handbook of linguistic minimalism*, C. Boeckx (ed.), 239-259. Oxford: Oxford University Press.

Wexler, K. & P. Culicover. 1981. *Formal principles of language acquisition*. Cambridge, MA: MIT Press.

Zeijlstra, H. 2012. There is only one way to agree. *The Linguistic Review* 29: 491-539.