Merge, Move, and Contextuality of Syntax: The Role of Labeling, Successive-Cyclicity, and EPP Effects

Željko Bošković
University of Connecticut

1. Introduction

Minimalist revolution within the Principles & Parameters approach may be looked at as an Occam razor purging of the theoretical assumptions of the Government & Binding theory (GB). While the theoretical changes it has brought to GB were rather vast, even spectacular in their scope, Minimalism at its heart was a methodological revolution, which came about as a result of taking Occam's razor seriously. Take e.g. the Inclusiveness Condition (Chomsky 1995), a fundamental principle of Minimalism, which is essentially an Occam-razor mechanism. Under the Inclusiveness Condition, the computational system can only manipulate lexical items that enter the computation—it cannot introduce anything new. The Inclusiveness Condition is thus a ban on creationism in the syntax. It greatly constrains the power of syntax since it prevents syntactic representations from including elements like traces, indexes, bar-level distinctions…, which are not present in the lexicon. The simplification of the syntax this has resulted in is particularly evident in the way structure building works (Bare Phrase Structure).

One of the minimalist tenets is that language faculty is characterized by efficient design. Occam's razor as a research methodology in fact leads to the efficient design hypothesis—it is a dictum that your subject of inquiry should be only as complex as it needs to be, hence the efficient design hypothesis.

Recent research in syntax has simplified the syntactic design of language essentially to a simple merge operation and a locality domain.[1] The operation Merge, which combines X and Y into {X,Y}, is minimally required for language. The notion of the locality domain contributes to efficient design. It was noticed early on that syntactic dependencies can span only a limited amount of structure. In the current theory, the locality of syntactic dependencies is treated in terms of phases, the goal being to have an optimal and efficient computational system. The phase theory, combined with the multiple spell-out hypothesis, on which syntax interacts with the interfaces throughout the syntactic derivation, accomplishes this by limiting the number of syntactic objects the derivation is working on, where this is achieved by transferring parts of syntactic structure to the interfaces during the derivation, the transferred parts not being accessible for further syntactic operations (see Uriagereka 1999).

The goal of this paper is to discuss two basic mechanisms that characterize the minimalist approach, operation Merge and the locality domain, the emphasis being on the interaction between the two, which occurs with movement, the most interesting case of such interaction involving successive-cyclic movement. As a result, understanding the mechanism of successive-cyclic movement is in many respects the key for understanding the syntactic design of language. We will see that understanding the nature of projection and labeling in the merger of two elements is in turn the key to understanding the special status of successive-movement. More generally, we will see that labeling plays a heavy role in the interaction between Merge and Move; it also heavily affects both the locality and motivation for movement. The most interesting case where the locality and motivation for movement interact involves contexts where the traditional EPP is supposed to be satisfied during successive-cyclic movement, i.e. where the element that satisfies it needs to undergo further movement. A discussion of that case will lead to a new contextual approach to the EPP, where the satisfaction of the EPP is not tied to a particular

[1]An important point here is often overlooked: efficient design should make language easily learnable. Generative syntax started with (the logical problem of) language acquisition: how can children learn something as complex as language so easily. From this perspective, the simplification of the syntactic design of language in recent research raises a serious issue: if we simplify the design too much, and say that only that is innate, a question arises how can the child learn the rest?

position. This in turn will be put into a broader perspective involving broader movement toward contextuality in the grammar in general: the contextuality of syntax will be shown to manifest itself in the locality of movement/islandhood (there are in fact no islands as this notion has been traditionally understood—there are no phrases that by their nature, independently of their syntactic context, disallow extraction), the motivation for movement, structure building, labeling, as well as the new conception of the EPP.

The paper is organized as follows. Section 2 gives the basic background regarding structure building through Merge and the role of labeling in the operation Merge. Section 3 discusses interaction between Merge and Move, where labeling plays a crucial role in the way the two interact in that unlabeled elements cannot undergo movement, cannot function as interveners, and cannot be targeted by movement. More broadly, labeling provides a new insight on, and a uniform account of, all island/ locality-of-movement effects. Sections 4-5 discuss the interaction of Move and Merge in traditional EPP (and related) effects, which will lead to a more general discussion of the contextuality of syntax.

2. Basic structure building: Merge and labeling

GB and early minimalism made a distinction between basic structure building and movement. Current minimalism treats the two in essentially the same way: movement also builds structure—in fact, when it comes structure building there is really nothing special about it: the difference between movement of X and lexical insertion of X in the merger of X and Y in (1) concerns where X comes from: from inside the structure where Y is already present (which is from inside Y except with head-movement), or from the lexicon (or numeration, a collection of items taken from the lexicon which serves as the starting point for the derivation, in Chomsky 1995)—there is really no difference between the two when it comes to the point of merger itself, i.e. (1) (this is also the case if X is a separately constructed object). The terminology that has been used since Chomsky (2004) to refer to the two, internal Merge (for movement) and external Merge (for lexical insertion), is in fact a reflection of that.[2]

(1) Merge(X,Y)

Consider now the result of the merger in (1). It has been a long-standing assumption that the nature of the resulting object needs to be specified. To illustrate, when a verbal element like *arrive* and a nominal element like *Mary* are merged in (2), information needs to be provided regarding whether the resulting object is verbal or nominal in nature. In (2), it is the former, which means *arrive* projects, labeling the resulting object.

(2) {arrive, Mary}

In GB, labels were provided by the X-bar theory as part of structure building. Early minimalism kept the gist of this approach. Thus, for Chomsky (1995) labeling is part of the definition of Merge: when X and Y merge, either X or Y projects, labeling the resulting object. In (3), X is the label of the object formed by merging X and Y.

[2]Chomsky (1995) argued for a Merge-over-Move preference, where satisfying a requirement by Merge (internal merge) is preferred to satisfying it by Move (external merge). The preference, which is not easy to maintain if Move is a subcase of Merge, has however, been questioned (e.g. Bošković 2002, Castillo, Drury, and Grohmann 1999, Epstein and Seely 1999; in fact, the considerations discussed in Chomsky 2019b would lead to preferring internal merge to external merge).

(3) Merge(X,Y)={X, {X,Y}}

Regarding what projects, for movement (internal Merge), it was assumed that the target projects; for external Merge, the issue was left open in Chomsky (1995), the implicit assumption being that in principle either X or Y can project, which introduces a potential difference between external and internal merge regarding the merge itself. However, that with movement the target always projects has been questioned. Thus, Donati (2006) argues that in free relatives like *she ate what I made*, the wh-element projects after movement, this being the reason why free relatives behave like nominals.

In a more radical departure from GB, Collins (2002) argues for a label-free theory, where structure building through Merge takes place without recourse to labels, as in (4).

(4) Merge(X,Y)={X,Y}

Chomsky (2013) also assumes that labeling is not part of the definition of Merge, adopting (4) rather than (3). Still, he keeps the traditional assumption that the nature of the resulting object needs to be specified. However, he argues that the interfaces, not syntax, requires this information: there is nothing wrong with unlabeled objects in the syntax, but such objects are uninterpretable. Chomsky provides an algorithm specifying labels which applies at the point of transfer to the interfaces, labeling being interface-driven. Under this view, labeling does not apply as part of Merge, in contrast to Chomsky (1995), but syntax is still not fully label-free, in contrast to Collins (2002), labeling taking place at the phasal level, when the structure is sent to the interfaces. Early approaches essentially stipulated that labeling is necessary to permit further applications of Merge, Merge applying only to labeled structures. This is not the case in Collins (2002) and Chomsky (2013), where labeling is not part of Merge and Merge can apply to unlabeled objects.

It is actually not obvious that labels are needed for interpretation—formal semantics models typically don't employ labels. Even if they are needed for interpretation, a question arises if this holds for all cases. In this respect, Chametzky (2000), Hornstein and Nunes (2008), Hunter (2010), Bošković (2015), and Blumel (2017) argue that adjunction doesn't require labeling for interpretation, which under Chomsky (2013) would mean that the result of adjunction is not labeled. In fact, Bošković (2015) suggests that the lack of labeling is the defining property of traditional adjunction, segmentation being dispensable.

Putting this issue aside, Chomsky's interpretation-driven approach leads to labeling taking place at the phasal level since phases determine when structure is sent to the interfaces. If labeling occurs for interpretive reasons, this is where it should apply. As Bošković (2016a) notes, this raises a serious chicken-or-the-egg question. Bošković (2015) argues that phasehood determination requires labeling: phases are objects like CPs, DPs…—they are labeled objects. Phases do not really exist prior to labeling: to know whether something is a phase we need to know its label. Since phases determine the points of spell out, without any labeling structure cannot be sent to the interfaces. But sending structure to the interfaces is necessary for labeling to apply under a purely interpretative approach to labeling. We are then stuck (in a chicken-or-the-egg manner). Bošković (2016a) observes that the problem dissolves if head-complement merger is labeled immediately since this is all that is needed to determine spell-out points. There then needs to be a syntactic reason for labeling to apply in this case. Bošković (2016a) suggests that in this case labeling is required by subcategorization: satisfying subcategorization, which is a syntactic requirement (hence needs to be satisfied during syntactic computation when the relevant object is created), requires that the element with the requirement to take a complement project (i.e. determine the label of the resulting object), otherwise, there would be no head-complement relation here. Based on such considerations, Bošković (2016a) argues that in a head-complement merger, labeling is

done immediately, which also resolves the chicken-or-the-egg problem regarding determining spell-out points, while with a merger of two non-minimal projections, labeling occurs when structure is sent to the interfaces. (Chomsky 2013 in fact treats labeling with head-complement/head-phrase and phrase-phrase mergers rather differently in that in the former one element, the head, essentially automatically projects (which makes sense if labeling here is subcategorization-driven), while in the latter neither of the elements undergoing merger projects by itself (see below), i.e. neither of them is dominant (see also Chomsky 2019b)—Chomsky thus essentially keeps the automatic nature of labeling for the former, substantially changing it only for the latter.[3])

3. Interaction between Merge and Move

3.1. Unlabeled elements cannot move

Having discussed the Merge operation, I turn to its role in movement. Chomsky (2000, 2001) gives a number of criteria that differentiate phases. One of them is (5).

(5) Only phases can undergo movement.

As Bošković (2018) observes, given that only phases can undergo movement and that unlabeled elements cannot be phases (see above), (6) follows as a theorem.

(6) Unlabeled elements cannot undergo movement.

Bošković (2018) shows that (6) deduces the freezing effect in (7), in effect turning it into a corollary.

(7) Movement is not possible out of moved elements.

The ban on movement from subjects in SpecIP (see (8)) is one case of (7), given the VP-Internal-Subject hypothesis. (9) also illustrates (7), given that objects preceding particles undergo A-movement (Johnson 1991, Lasnik 1999).

(8) ?*I wonder [$_{CP}$ who$_i$ [$_{DP}$ friends of t_i]$_j$ [$_{vP}$ t_j hired Mary]].
(9) ?*Who$_j$ did you call [friends of t_j]$_i$ up t_i?

The freezing effect also holds for extraction from elements in A'-positions (10). The impossibility of P-stranding in (11) also illustrates (7).

(10) *Whose books$_i$ do you think that [reviews of t_i]$_j$ John hates t_j?
(11) *Which table$_i$ did you think that [on t_i]$_j$ John put the book t_j?

Now, Chomsky (2013), where, as discussed above, labeling is not part of Merge, proposes a labeling algorithm where when a head and a phrase merge, the head projects (labeling the resulting object). When two phrases merge, there are two ways to label: through feature-sharing or traces, where traces are ignored for labeling. (12) illustrates the former: when *which book* merges with interrogative CP (this element is a CP since after C, a head, merges with IP, C projects), both the wh-phrase and the CP have the Q-feature; what determines the label of the resulting object is then the Q-feature. (This is similar to Spec-Head agreement.)

[3]Dadan (2019) capitalizes on this difference by providing an account for a number of cases of diachronic change.

(12) I wonder [which book$_i$ C [John bought t$_i$]].

Turning to the non-feature-sharing case, as argued in a series of works by Bošković (1997, 2002, 2007, 2008a; see also fn 13), Chomsky (2013) assumes that successive-cyclic movement does not involve feature-sharing. Since there is no feature-sharing between *that* and the wh-phrase in (13a), the embedded clause cannot be labeled when *which book* moves to its edge (indicated by ? in (13b)). When v merges into the structure, *which book* undergoes movement. The element merged with *that*-CP being a trace, it is ignored for labeling—? is then labeled as CP after *which book* moves. This is extended to <u>all</u> successive-cyclic movement: there is no feature-sharing with successive-cyclic movement, which creates a labeling problem that in turn forces movement.

(13) a. Which book$_i$ do you think [$_{CP}$ t'$_i$ that [he bought t$_i$]]?
b. v [$_{VP}$ think [$_?$ which book [$_{CP}$ that [he bought t$_i$]]]

Returning to the freezing effect, consider (8) at the point when the subject is located in its base-position.

(14) [$_{DP}$ friends of who] v-hire Mary

Given the cycle, *who* must move to the edge of the subject DP before the latter moves. Since this is a case of successive-cyclic movement, it involves non-agreeing/non-feature-sharing phrase-phrase merger, which delabels the subject.[4] This creates a configuration disallowed by (6) since an unlabelled element (the subject) has to undergo movement. Under this account, which extends to other cases of the freezing effect, there is actually nothing that is in principle wrong with movement out of a moved element, XP. Such movement is in principle allowed, but successive-cyclic movement to the edge of XP freezes XP itself for the possibility of movement (given (6)).

The labeling deduction of the freezing effect makes a prediction: the effect should only hold for successive-cyclic movement since only succcessive-cyclic movement delabels the element it targets, freezing the relevant object for further movement due to (6) (which follows from (5)). An element base-generated at the edge of a moved element, or that independenty of successive-cyclic movement moves to the edge of a moved element, is expected to be allowed to move from a moved element. Bošković (2018) provides a number of such cases. The former is illustrated by Serbo-Croatian (SC) possessors, which have been argued to be base-generated at the traditional NP (TNP) edge (e.g. Bošković 2012, Despić 2013) based on the fact that they can undergo extraction and bind out of their TNP. Importantly, the possessor can be extracted out of a moved element, as (15) shows.

(15) Jovanovu$_i$ je ona [t$_i$ sliku]$_j$ vidjela t$_j$.
John's.acc.fem.sg is she picture.acc.fem.sg seen
'She saw John's picture.'

Dutch *r*-pronouns illustrate the case where an element that obligatorily moves to the edge of a moved element extracts out of it. Dutch *r*-pronouns are exceptional regarding word order within PP: they must precede the P although Dutch adpositions are otherwise always prepositional.

(16) daar op/*op daar
there on

[4]The movement does not delabel the DP from (14). It creates a new structural layer on top of it which itself lacks a label.

This is analyzed as involving obligatory movement of *r*-pronouns to SpecPP (or a higher position in the traditional PP). *R*-pronouns can also move out of a PP, stranding the P. Crucially, they can move out of moved PPs.

(17) Waar$_i$ had jij dan [t_i mee t_i]$_j$ gedacht dat je de vis t_j zou moeten snijden?
where had you then with thought that you the fish would must cut
'What did you think you should cut the fish with?' (Barbiers 2002)

The acceptability of (17) is rather dramatic in light of (11): in contrast to English, P-stranding in a moved position is possible in Dutch with *r*-pronouns. The crucial difference is that before extraction from the PP, the *r*-pronoun undergoes obligatory movement to SpecPP which is independent of successive-cyclicity.

All cases of acceptable movement from moved elements given in Bošković (2018) show the same pattern as (15)/(17): the element which exceptionally moves from a moved element is either base-generated at the edge of the moved element or obligatorily moves there (independently of successive-cyclicity). Furthermore, ill-formed examples of movement from moved elements do not involve these two configurations. In fact, under Bošković's (2013, 2014) approach to phases, where the extended projection of every lexical category is a phase, in all the unacceptable examples given above the moved element is a phase, which means these examples must involve successive-cyclic movement via the edge of the moved element, forced by the PIC, which requires movement out of phases to proceed via phasal edges. The account of (8) then extends to all of them.

Under this approach, which quite radically departs from the traditional freezing ban, there's nothing wrong in principle with movement out of moved elements (as confirmed by (15)/(17a))–even in unacceptable cases the problem does not arise when YP moves out of moved XP–the problem arises with the movement of XP itself, i.e. moving XP does not freeze the internal structure of XP for movement–movement of YP to the edge of XP prevents movement of XP. Any later movement out of XP is then trivially disallowed. The reason why the effect is restricted to successive-cyclic movement is the way it affects labeling (which brings brings (6) into play). Since the relevant cases involve successive-cyclic movement, they also involve movement of YP out of XP since it's the very nature of successive-cyclic movement that a phrase undergoing it cannot stay in the intermediate Spec for independent reasons. The reason why the relevant examples involve movement out of a moved element is thus accidental, due to the nature of successive-cyclic movement. This has, however, led to the illusion that this later movement is the reason for their unacceptability. The traditional freezing ban was thus misguided: there is no problem with movement "from" (movement of YP in YP_i $[_{XP}$ t_i $]_j$ t_j) but with movement "of" (movement of XP).

That movement "from" does not matter is confirmed by the fact that the account extends to an otherwise puzzling case which does not involve movement out of a moved element at all, namely the immobility of V-2 clauses in German (e.g. Webelhuth 1992, Reis 1997), illustrated by (18), where a V-2 clause moves to SpecCP.

(18) *[Er$_i$ sei unheimlich beliebt], möchte jeder$_i$ gern glauben.
he is.subj immensely popular would.like everyone like believe
'Everyone would like to believe he is immensely popular.' (Wurmbrand 2014)

V-2 clauses are notorious for the promiscuity of their Spec position: anything can fill it. This has led to proposals that they do not involve agreement—they involve EPP without Agree (e.g. Haegeman 1996,

Roberts 2004). Since feature-sharing involves agreement, a natural interpretation of this is that V-2 clauses do not involve feature-sharing, which means they are not labeled (see also Blümel 2017). The immobility of V-2 clauses then falls out from (6). What is important here is that under accounts like Roberts (2004), movement to SpecCP of V-2 clauses is treated like successive-cyclic movement in Chomsky (2013) in that neither involves an agreement relation. Since the crucial ingredient of Bošković's (2018) account of the freezing ban is that phrases with non-agreeing Specs cannot undergo movement (since a non-agreeing Spec delabels the relevant phrase),[5] it is natural that, just like phrases "hosting" successive-cyclic movement, V-2 clauses cannot undergo movement.

The generalization in (6) (which follows from (5)) thus enables us to unify the traditional freezing ban and the immobility of V-2 clauses in German, while also capturing a number of exceptions to the freezing ban, restricting its effect to successive-cyclic movement (the unification with the immobility of V-2 clauses capitalizes on V-2 movement to SpecCP being formally the same as successive-cyclic movement in the relevant respect).

The cases of movement from a moved element blocked by (6) involve successive-cyclic movement through the edge of the moved element. This actually confines the effect to phases since only phases must involve such movement. Now, while (6) allows movement from a moved element under all conditions in (19), its deduction based on Chomsky's (5) disallows it for (19c). Since XP in (19) moves, it must be a phase given (5).

Given the PIC, YP then cannot move out of XP without moving to SpecXP, which blocks (19c).[6]

(19) YP can move out of moved XP iff:
- a. YP is base-generated at the edge of XP.
- b. YP must move to the edge of XP independently of successive-cyclic movement.
- c. YP does not move to the edge of XP.

There have, however, been proposals where non-phases move. Thus, Collins (2005) proposes such a derivation for passives, which furthermore involves movement from a moved element (Collins suggests VoiceP, not PartP, is the phase in (20)).

(20) $[_{IP}$ John$_i$ was $[_{VoiceP}$ $[_{PartP}$ t_i arrested $t_i]_j$ $[_{Voice'}$ by $[_{vP}$ Mary $[_{v'}$ v $t_j]]]]]$

(20) conforms to the reformulation of the freezing ban in (6) (it does not involve unlabeled-element movement) but not the deduction of (6) that is based on (5)— (20) is ruled out by (5) because the moving PartP is not a phase. (6) itself, however, does not depend on only phases being mobile. (6) takes effect with successive-cyclic movement because such movement delabels the element it targets. The effect is confined to phases because of the PIC since only phases must be targeted by successive-cyclic movement. However, if non-phases can in principle move, they would not be subject to the freezing ban under its reformulation (6) (or (i)) in fn 5), since movement need not proceed through their edge.

[5]Bošković essentially replaces the traditional freezing ban with (6). Without appealing to unlabeled objects, it would be restated as in (i). ((i) has the same empirical coverage as (6); recall, however, that (6) is deducible from an independent principle.)

(i) Phrases with non-agreeing Specifiers cannot undergo movement.

[6]The approach to the PIC in Bošković (2015) where phasal complements (but not what they dominate) are accessible from the outside would allow (19c) if YP is the complement of phasal head X since movement of phasal complements need not proceed via phasal edges under this approach.

Bošković (2018) briefly outlines a deduction of (6) independent of (5), which allows non-phases to move, based on Bošković's (2007, 2011a) proposal that movement is driven by an uninterpretable/unvalued feature (uK) of the moving element (see also fn 8). Note first that the proposal fits the labeling framework quite naturally. The natural expectation there is that movement is labeling-driven: it takes place to resolve labeling problems. This is in fact what happens when XP and YP merge without feature-sharing: this creates a labeling problem, with movement taking place to resolve it. What happens here is that the problem, and the reason for movement, is present in the pre-movement structure, i.e. the base-position of movement: something would go wrong in the base-position of movement if movement does not occur—there is nothing in the higher structure that motivates it. This is also the central characteristic of Bošković's (2007) approach to movement, implemented through the presence of a uK feature on the moving element, which forces movement (without movement, a crash would occur); both the labeling approach of Chomsky 2013 and Bošković 2007 thus involve base- rather than target-driven movement.[7] It then seems natural to adopt Bošković's uK assumption here: X moves only if X has a uK.[8]

Consider then the case where X and Y merge and the resulting object ? moves, which means it must have a uK. For the movement to occur, either X or Y must have a uK and pass this feature to ? by labeling it. If X has the uK feature, then X must project and label ?. The upshot of this is that labeling is necessary for ? to have a uK and be able to move, which means that unlabeled elements cannot move. The assumption that X moves only if X has a uK thus makes it possible to deduce (6) independently of (5), i.e. without requiring that only phases move.

Under both of the above deductions of (6), movement from moved elements is in principle allowed: the relevant violation does not occur at the point of movement from a moved element but earlier in the derivation. Both deductions allow movement from a moved element under (19a-b), but only the uK deduction allows it under (19c). More research is needed to determine whether (19c) should indeed be allowed.

3.2. Unlabeled elements do not function as interveners

There is another case where labeling and movement interact. Bošković (in press b) argues that unlabeled elements are not only unable to undergo movement, they also do not function as interveners.

(21) Unlabeled elements do not count as interveners.

---

[7]See below for evidence for this overall approach based on subject questions, where *who* in *who left* moves to SpecCP without moving to SpecIP. See also Bošković (2011b) for other cases where movement is clearly not target-driven. One is QR—QR must be moving-element driven since there is nothing about its target that would require adjunction of a quantifier. **(**Quantifiers are assumed to be uninterpretable in-situ. This can be tied to the presence of a uK feature which makes them uninterpretable in-situ, requiring movement.)

Note that Bošković (2007) and Chomsky (2013) also treat successive-cyclic movement in the same way. The crucial property of the former is that there is no feature-checking/agreement with successive-cyclic movement and that for each step of successive-cyclic movement, it is something about the base-position of movement that drives it: something goes wrong there if the movement does not take place. Thus, in Bošković (2007), there is no feature-checking/agreement between *which book* and *that* in (13). Moreover, if *which book* does not move from the embedded SpecCP, a problem arises in this part of structure. This is very different from Chomsky (1995), but these are precisely the crucial ingredients of Chomsky's (2013) approach to successive-cyclic movement.

[8]The assumption is actually also there in Chomsky (2000) as part of the Activation Condition, which essentially sneaks in base-driven motivation into a target-driven system. A result of the Activation Condition is that X can only move if it has a uK. Given this, the analysis about to be presented can be implemented in a target-driven system like Chomsky (2000), where the EPP/edge feature drives movement (in fact, simply tying satisfying the EPP to a particular feature would be enough). At any rate, what is crucial below is that X moves only if X has a uK—whether this is a result of the Activation Condition or a deeper property of the system is not important.

If unlabeled elements cannot undergo movement it is not surprising that they do not function as interveners, since they are not candidates for movement themselves.

(21) is rather natural theoretically even independently of this. *Intervention* depends on the nature of the intervener. In Rizzi (1990), this involved the A/A' distinction; current work appeals to the intervener's featural properties (e.g. Rizzi 2004, Starke 2001). Labeling plays a crucial role here. Consider the situation where X and Y merge, and the resulting object ? functions as an intervener. For an intervention effect to obtain, either X or Y must have the feature involved in the intervention effect and pass this feature to ? by labeling it (so that the resulting object can function as an intervener). If X has the relevant feature, X must project and label ?. What this boils down to is that labeling is necessary for ? to function as an intervener, which means that unlabeled elements cannot function as interveners. In other words, since intervention is feature-sensitive, the intervener must have the relevant feature. This is trivially not possible with unlabeled elements since due to the lack of projection in general the relevant feature is also not projected. Unlabeled elements then cannot undergo movement and function as interveners for essentially the same reason.

3.3. Movement cannot target unlabeled elements

Bošković (2016b) argues that movement also cannot target unlabeled elements (see also Yoo 2015). This is e.g. what is behind Richards' (2001) tucking-in effect, illustrated by Bulgarian multiple wh-fronting in (22). Here, the nominative wh-phrase moves first to SpecCP, given Superiority, with the second wh-phrase then moving to a lower SpecCP.

(22) a. Koj$_i$ kogo$_j$ t$_i$ e udaril t$_j$? b. cf. *Kogo koj e udaril?
who whom has hit
'Who hit whom?'

Consider (22) before the second wh moves. Given that the result of a head-phrase merger is labeled immediately and the result of a phrase-phrase merger only after the phase is completed, as discussed in section 2, the merger of C-Q (i.e. +wh-C) with IP results in labeling, this object being labeled as CP, but the merger of this CP and the nominative wh-phrase *koj* does not result in immediate labeling (it's labeled only after the phase is completed, which means after the second wh-phrase moves), as shown in (23). As a result, if *kogo* were to merge on top of *koj* in (23), the movement would target an unlabeled element. However, this is not the case if *kogo* merges under *koj* (the movement then targets CP). The tucking-in effect then follows from the requirement that movement target only labeled elements.[9]

(23) [$_?$ *koj* [$_{CP}$ C-Q [$_{IP}$ …*kogo*…]]]

Movement and labeling, then, interact rather strongly: unlabeled elements cannot undergo movement, unlabeled elements do not function as interveners, and movement cannot target unlabeled elements (movement also must cross a labeled projection under Bošković's 2016a approach to antilocality).[10]

3.4. Movement is labeling-driven: uninterpretable features block labeling

[9]Bošković (2016a) actually deduces the requirement, making it a theorem. In particular, he deduces it from a labeling-based approach to antilocality. This also explains why the requirement holds only for movement, i.e. internal merge (external merge with an unlabeled element is possible): antilocality is a constraint on movement.
[10]As noted below, the labeling theory also provides a uniform account of all island/locality-of-movement effects.

Returning now to the uK assumption (present in both Bošković 2007 and Chomsky 2000), where a moving element must have a uK (i.e. X moves only if X has a uK), there is another case where it can be profitably used in the labeling framework, which concerns Bošković's (2020) proposal in (24).

(24) The presence of an uninterpretable feature blocks labeling via feature-sharing in XP-YP configurations

As noted above, the natural expectation in the labeling framework is that all movement is labeling-driven—it takes place to resolve labeling problems. This is what happens when XP and YP merge without feature-sharing (as with successive-cyclic movement): movement takes place to resolve a labeling problem. But what about cases like (25), where labeling must be possible pre-movement, given (26).

(25) Jovanove$_i$ on voli [t_i knjige].
John's he loves books
(26) On voli Jovanove knjige. (SC)

Given Bošković's (2007) uK assumption discussed above, *Jovanove* in (25) must have the uK feature which drives the relevant movement operation (otherwise it could not move). However, given (24), the uK feature blocks feature-sharing in the base-position of *Jovanove*: movement then occurs to resolve a labeling problem (i.e. the movement here is now labeling-driven, just as with successive-cyclic movement). The labeling problem does not arise in (26), where the relevant uK feature is not present (if it were, *Jovanove* would have to move).

Now recall that in Bošković (2007, 2011a), movement in general is driven by an uninterpretable/ unvalued feature, uK, on the moving element.[11] In Chomsky (2013), it is driven by labeling problems. As noted above, the labeling and uK-driven movement systems are very similar (thus, movement in both is base-, not target-driven, see also fn 7), and in fact easily combinable (see also Takita, Goto, and Shibata 2016 on combining the two, though they do it in an opposite way from what is done directly below, by having uK, rather than labeling, drive movement). However, adopting both would result in double motivation for movement: labeling problems and the uK feature-induced crash. The double driving force can, however, be eliminated in a way that actually reconciles the two systems: the suggestion is that uK itself doesn't drive movement, only labeling issues drive movement, but uK causes a labeling issue.

All this should then apply to the head-complement configuration too: uK should also be blocking labeling in this case—even complement movement, as in (27), will then be labeling-driven.

(27) What did Mary buy?

To make movement in general labeling-driven, (24) should then be generalized to apply to all cases, not just XP-YP feature-sharing cases, as in (28).
(28) The presence of an uninterpretable feature blocks labeling.

All movement is then labeling-driven in the sense that it takes place to resolve labeling problems.

4. The *who left* puzzle

The above discussion also allows to address the *who left* effect puzzle.

(29) Who left?

[11]In Chomsky (2000), where uK is also present, the uK makes movement possible but does not actually drive it.

There is a lot of evidence that subject wh-movement from SpecIP to SpecCP is crosslinguistically banned. Thus, in languages with both SV and VS order (in the latter the subject does not move to SpecIP) where these orders are associated with different verbal morphology, what we get in (29) is the VS order morphology (e.g. some dialects of Italian). This shows not only that the subject in subject questions does not remain in SpecIP, but that wh-movement to SpecCP cannot proceed via SpecIP, otherwise we would get the SV word order morphology. The point can also be made with languages where agreement morphology associated with subjects in SpecIP must be dropped in (29) (e.g. Kinande).

Consider also West Ulster English (WUE) (30).

(30) a. Who$_i$ was arrested all t_i in Duke Street?
b. *They$_i$ were arrested all t_i last night. (McCloskey 2000)

Although, in contrast to standard English, WUE allows Q-float under wh-movement, like standard English it disallows (30b). (30b) shows that a subject in SpecIP cannot float *all* postverbally in passives. *Who* in (30a) then cannot move to SpecIP before moving to SpecCP since *all* would then be floated under movement to SpecIP, which (30b) shows is disallowed. McCloskey (2000) concludes based on (30) that *who* moves directly to SpecCP, without moving to SpecIP (for additional arguments, see Bošković in press a and Messick 2020). This can actually be interpreted as an argument against the traditional assumption that the EPP is a requirement on the target head, I (namely, that I must have a Spec). On the other hand, this is easily captured in both Bošković (2007) and Chomsky (2013), where the traditional EPP effect has nothing to do with I: the subject moves because a problem would arise in the base-position of the subject if it does not move (the movement is not driven/required by I). Since neither system requires the subject to move to SpecIP, in both Bošković (2007) and Chomsky (2013) the relevant inadequacy can be satisfied if the wh-phrase in (29)/(30a) moves to SpecCP.

But why is it that subject wh-movement cannot proceed through SpecIP? Recall the ingredients of the above account of the driving force of movement. Bošković (2007) argues that uK on the moving element drives movement; what we have taken from there is simply that every moving element has a uK (see also Chomsky 2000). Furthermore, Bošković (2020) suggests that uK blocks feature-sharing, which was generalized to uK blocking labeling (28). These assumptions make it possible to make movement in general labeling-driven, with all movement taking place to resolve labeling problems. This is in fact all we need to address the *who left* effect puzzle—these independently motivated assumptions give us for free an explanation for why subject wh-movement cannot proceed via SpecIP. Given the assumptions in question, *who* in (29) must have a uWh feature. If *who* moves to SpecIP, uWh will block feature-sharing/labeling. This then prevents movement through SpecIP in (29)—in essence, the movement would not do anything and it is also not needed by the PIC (recall the movement is not required in Bošković 2007 and Chomsky 2013, see also below on this issue).

In fact, it turns out that the *who left* effect is much more general than the ban on subject wh-movement through SpecIP. It is not only that subject wh-movement cannot proceed through the usual derived subject A-position, object wh-movement also cannot proceed through the usual derived object position. Thus, in Kinande, not only subject agreement, but also object agreement cannot be present if the object undergoes wh-movement (Schneider-Zioga 1995, Bošković 2008a, 2016a). Given that in Kinande an agreeing object precedes an agreement marker, indicating overt object shift, the obligatory drop of the agreement marker indicates impossibility of object shift on the way to SpecCP (see Bošković 2008a, 2016a). Bošković (2020) provides evidence that the same holds for English.[12] The above account

[12]The main argument is based on the contrast in (i) regarding the Coordinate Structure Constraint (see Bošković 2020 for details and what this entails for phasehood of the middle field. Note also that while the impossibility of

of the *who left* effect extends to the impossibility of object shift/movement to SpecAgroP on the way to SpecCP. Just like uWh-feature blocks feature-sharing/labeling with movement to SpecIP, which prevents wh-movement through SpecIP, it also blocks feature-sharing/labeling with movement to SpecAgroP, preventing wh-movement through SpecAgroP.

All this is even more general. Bošković (2008a) argues that there is a general ban on multiple feature-checking, where X moves to a feature-checking (feature-sharing in Chomsky 2013) position, followed by movement to a different feature-checking position. The *who left* effect and the object shift effect are illustrations of this; they show that feature-checking A-movement cannot feed feature-checking A'-movement. There is also a great deal evidence that feature-checking A'-movement cannot feed another feature-checking A'-movement, i.e. A'-movement of type X feeding A'-movement of type Y is banned—this concerns the criterial freezing effect, where Op-variable creating movements like wh-movement, topicalization, focalization, and QR cannot feed each other (see (31)), see e.g. Epstein 1992, Rizzi 2006, Bošković 2008b). Bošković (2008b) in fact argues that what is behind successive-cyclic movement not involving agreement is actually the ban on multiple feature-checking, where a feature-checking movement feeds another feature-checking movement.[13]

The above discussion of the *who left* effect generalizes to deduce Bošković's broader generalization regarding the ban on multiple feature-checking, where a feature-checking movement feeds a feature-checking movement (the *who left* effect being just one illustration of this broader effect). To undergo two feature-checking movements, α would need to have two uKs. Say we are dealing with a case where a phrase undergoes wh-movement followed by topicalization, as indicated by topic intonation in (31a). ((31b) shows the landing site of the latter is higher than the former.)

(31) a. *Which professor, does Mary detest?

---

subject wh-movement passing through SpecIP might be attributable to antilocality, under Bošković's (2016a) labeling approach to antilocality where movement must cross a labeled projection and phrase-phrase merger labeling takes place at the phasal level (see also Erlewine 2016 as well as Bošković 1994, 1997, Abels 2003, and Grohmann 2003, among others, for antilocality more generally), antilocality clearly cannot be relevant in the object wh-movement case since there are a number of phrases between the landing site of object shift and SpecCP).

(i) a. *Who$_i$ did you believe for a long time now [t$_i$ to be a liar] and [Peter to be trustworthy]?

b. ?I've believed John$_i$ for a long time now [t$_i$ to be a liar] and [Peter to be trustworthy].

[13]See Bošković (1997, 2002, 2008a) for evidence for the position that successive-cyclic movement does not involve agreement, which Chomsky simply assumes. One argument given there concerns the Lobeck (1990)/Saito and Murasugi (1990) observation that functional heads license ellipsis of their complement only when they undergo Spec-Head agreement (SHA), i.e. feature-checking, as illustrated by (i).

(i) a. John liked Mary and [$_{IP}$ Peter$_i$ [$_{I'}$ did t_i ~~*like Mary*~~]] too.

b. John's talk about the economy was interesting but [$_{DP}$ Bill [$_{D'}$'s ~~*talk about the economy*~~]] was boring.

c. *A single student came to the class because [$_{DP}$ [$_{D'}$ the ~~*student*~~]] thought that it was important.

d. John met someone but I don't know [$_{CP}$ who$_i$ [$_{C'}$ C ~~*John met*~~ t_i]].

e. *John believes that Peter met someone but I don't think [$_{CP}$ [$_{C'}$ that ~~*Peter met someone*~~]].

Bošković (1997, 2002) notes that an intermediate C cannot license ellipsis of its IP complement (ii), which follows if passing through an intermediate SpecCP does not imply feature checking/SHA with the C.

(ii) *John met someone but I don't know who$_i$ Peter said [$_{CP}$ t$_i$ [$_{C'}$ C/that ~~*John met t$_i$*~~]].

Similar argument can be given for the impossibility of feature-checking A-movement feeding feature-checking A-movement based on the ellipsis contrast between control and ECM infinitives in (iii), noted in Martin (2001) (see also Bošković 2008a for an account of languages with overt reflexes of agreement with intermediate heads under wh-movement that does not involve feeding feature-checking movements).

(iii) a. John was not sure he could leave, but he tried [$_{IP}$ PRO$_i$ [$_{I'}$ to ~~t_i *leave*~~]].

b. *John believed Mary to know French but Peter believed [$_{AgroP}$ Jane$_i$ [$_{IP}$ t$_i$ [$_{I'}$ to ~~t_i *know French*~~]]].

b. ?To Peter, what should Mary give? (Grohmann 2003)

To undergo topic and wh-movement, the moving phrase must have uTopic and uWh. The former will block feature sharing with +whC (this essentially blocks Spec-Head agreement with the C, which is necessary in a wh-movement language). More generally, since movement is labeling-driven and uK blocks labeling via feature-sharing/feature-checking, X can undergo feature-sharing/feature-checking only once.

5. Contextuality

Returning to the *who left* effect, we have seen why movement to SpecCP cannot proceed via SpecIP. We have also seen that while the lack of movement to SpecIP is a problem for standard approaches to the EPP, which tie it to a property of IP/I, in a way that requires filling the SpecIP position, it is not a problem for approaches to EPP effects in works like Bošković (2007) and Chomsky (2013), which do not require filling the SpecIP position. Let us, however, see how the lack of SpecIP in constructions like (29) could be captured in more standard approaches to the EPP effect, which tie it to a property of the target head. A significant modification of those approaches will obviously be required, since IP does not have a Spec in (29).

Somewhat analogously to Rizzi's (1997) claim that there are two topic positions, in work in prep I argue that there two wh-positions, a higher and a lower one, where the lower wh-position is occupied by wh-moved subjects. I provide a number of arguments that wh-moved subjects and wh-moved objects do not move to the same position. One argument comes from Kaisse's (1983) observation that there is a one-word host restriction on contracted auxiliaries hosted by moved wh-phrases, but, crucially, only with non-subject wh-phrases, which I interpret as indicating that the wh-phrases/auxiliaries are not in the same position in (33a)-(34a) and (33b)-(34b). (Recall wh-subjects do undergo wh-movement.)

(32) a. What's Mary buying? b. When's dinner? c. How's your old man?
(33) a. *Whose food's the dog eating? b. Whose food's burning?
(34) a. *Which man's she the fondest of? b. Which man's leaving first? (Kaisse 1983)

Another argument for the lower position of subject wh-movement comes from interaction with topicalization: (35)-(36) show that only the landing site of non-subject wh-movement is above the topic.

(35) a. ?Mary wonders which book, for Kim, Peter should buy.
b. *Mary wonders which student, for Kim, should buy that book.
(36) a. ??I wonder under which table, that book, Mary put.
b. *I wonder which man, that book, put on the table.

Bringing the usual EPP position into all this, what we then have is this: *Who* in (37b) is not as high as *who* in (37a) but is higher than *Mary* in (37c).

(37) (a) I wonder **who Mary met** vs (b) I wonder **who left** vs (c) I think **Mary left**

As for why the lower wh-position is confined to subjects, I argue that this is a mixed A/A'-position on the border of the traditional A and A' fields: it is the landing site of wh-movement, but also the position where the EPP is satisfied (and nominative licensed). This explains what appeared to be puzzling voiding of the EPP effect in (30a) (if *who* had to pass through SpecIP *all* would be floated from the same position in both examples in (30)) and more generally (29). The EPP is satisfied in the lower wh-position, a mixed A/A' position confined to A'-moved subjects. Focalized subjects also seem to move to that position,

which is not surprising given that wh and focalized elements have been argued to pattern together regarding movement in a number of languages. Consider (38a-c).

(38) a. Only pennies did we share with a soul.
b. *We shared only pennies with a soul.
c. Only Francis gave pennies to a soul. (Branigan 1992)

The licenser *only pennies* c-commands *with a soul* in (38a-b): the contrast can be taken to indicate that *with a soul* cannot be licensed from a purely A-position. (38c) can then be captured if the focalized subject here moves to the mixed A/A' position like *who* in (29).

All this leads to a new conception of the EPP on a par with the contextual approach to phases in Bošković (2013, 2014), where there are phasal domains and the highest phrase in a phasal domain is a phase (e.g. DP is a phase in the nominal domain in English, but in languages without articles where DP is lacking, a lower projection in the nominal domain is a phase). In particular, there is an EPP domain, with the highest projection in this domain being the locus of the EPP.

So far we have (39) for different subjects (and the EPP domain; A/A'P is used for ease of exposition):

(39) [$_{A/A'P}$ wh-moved subject [$_{IP}$ Mary

There is evidence for additional subject positions in the EPP domain. First, Bošković (2020) argues for a return to split IP. One argument concerns coordination: given that bar-level coordination is disallowed, (40), where the subject is outside and the modal is inside of the coordination, provides evidence that the subject and the modal are not in the same phrase, the modal being lower than the phrase whose Spec the subject occupies. Note that this can be captured in early minimalist clausal structure, which split IP into AgrsP and TP, where in the perspective of the current discussion both would belong to the EPP domain.[14]

(40) John [travels to Rome tomorrow] and [will fly for Paris on Sunday]

Furthermore, Bošković (2019) and Cardinaletti (2004) argue that non-agreeing quirky subjects are lower than agreeing subjects, which then gives us (41) (XP would be AgrsP and YP TP under the split suggested above—non-agreeing subjects would naturally not be located in SpecAgrsP).

(41) [$_{A/A'P}$ wh-moved subject [$_{XP}$ Mary [$_{YP}$ quirky subjects

All of this can be extended to many constructions, e.g. the controversy regarding whether subject V-2 clauses in Germanic are CPs or IPs. The gist of that controversy is that subject V-2 clauses in several respects differ from non-subject V-2 clauses (see Travis 1991, Zwart 1991), but they are also not exactly the same as regular non-V-2 subject clauses: what this essentially indicates is that the subject in subject V-2 clauses is in a position that is lower than SpecCP but higher than SpecIP, which can be captured if the subject in such clauses is located in SpecA/A'P from (41) (cf. also (38) for focalized subjects in English).

In work in prep, I extend the analysis to a number of other elements that show mixed subjecthood properties, like clausal subjects, the subject of Locative and Quotative Inversion constructions, and the subject of Singlish no-agreement constructions (see Lee 2022 on the last one)—they are all argued to be

[14]One of original arguments for splitting IP was conceptual: Infl was a strange element that contained two kinds of rather different information, agreement and tense (looking at the two semantically it is quite obvious how different they are). The argument applies to the current assumption that Tense has phi-features.

located in SpecA/A'P (more generally, non-agreeing subjects cannot be located in what is labelled as SpecXP in (41), which is actually equated with AgrsP from early minimalism).[15,16]

To summarize, there is a height hierarchy regarding the wh-phrase/subject in (42a), shown in (42b).

(42) a. I wonder **what Mary bought** vs I wonder **who left** vs I think **Mary left**
b. what > who > Mary

In this hierarchy, wh-subjects are lower than wh-non-subjects but still higher than regular subjects. The analysis is extendable to other constructions, like Germanic V-2 subjects, focalized subjects, clausal subjects, locative inversion, and quotative inversion.

Based on this, I have argued for a contextual approach to the EPP (on a par with the contextual approach to phases), where the highest projection in the EPP domain is the locus of the EPP (on a par with the highest projection in a phasal domain being a phase). The hierarchy of the subject positions discussed here is given in (43).

(43) wh(A/A')-moved subjects>regular subjects>quirky subjects

A more general point about contextuality is in order. Consider the history of the locality-of-movement/ islandhood research. Early on, in the bounding node approach (see Chomsky 1973), the trouble-makers for movement were defined rigidly: NP and IP were bounding nodes regardless of the syntactic context in which they occur. While the *Barriers* system (Chomsky 1986) is quite different from the bounding node approach, the importance of one particular difference has generally gone unnoticed, namely the contextuality of *Barriers*. One cannot even ask whether e.g. CP is a barrier. Actually, one could, but the answer would be *it depends*: in particular, it depends on the syntactic context in which the CP occurs. In *Barriers*, trouble-makers for movement were defined contextually.[17] In the current theory, trouble-makers for movement are defined in terms of phases. Chomsky's original (2000, 2001) approach to phases essentially went back to the bounding node approach in that it defined phases rigidly: thus, CP and NP (ignoring the DP hypothesis) are phases regardless of their structural position. The lesson of *Barriers* regarding the contextuality of the locality of movement was lost in that approach. However, the original approach to phases was soon followed by various contextual approaches to phasehood, where whether XP is a phase or not depends on the syntactic context in which XP occurs (on a par with *Barriers* and in contrast to the bounding node and the early phasehood approach), see e.g. Bošković 2005, 2014, 2015, Bobaljik and Wurmbrand 2005, den Dikken 2007, Despić 2011, Gallego & Uriagereka 2007, M. Takahashi 2011 for different approaches.[18] Pushing the contextuality even further, Bošković (2016b)

[15]See the work in question for the reason why quirky subjects are lower and other non-nominative subjects higher than this position.

[16]The account is extendable to Japanese scrambling. Miyagawa (2003) argues that the scrambled O in the OVS order satisfies the EPP when the subject stays in SpecvP (a possibility in Japanese). Now, it is well-known that the scrambled O in the OVS order shows both A and A'-properties. All this can be captured if the scrambled object that satisfies the EPP is located in SpecA/A'P (we would essentially be dealing here with a non-nominative subject, as Miyagawa essentially argues).

[17]The relevance of contextuality is quite clear: e.g. the bounding node approach, which considered NP a bounding node, had problems with making a distinction between subject, object, and adjunct NPs with respect to extraction out of them, which was done easily in *Barriers*.

[18] One important point that is emphasized in the works of Bošković but still often missed is that the contextual approach to phasehood enables us to capture crosslinguistic variation with respect to extraction without positing parametric variation regarding what counts as a phase, which would involve positing parametric variation in the

argues that not only can the phasal status of a phrase be affected by the syntactic context in which it occurs, but the concept of *phasal edge*, i.e. the status of a Spec regarding the PIC, can also be affected by the syntactic context in which the Spec occurs (the gist of it is that, just like the highest phrase in a phasal domain functions as a phase, the highest edge in multiple-edge contexts functions as the phasal edge). In other words, to know whether XP is a phase or not we need to look at the syntactic context in which XP occurs, and even once we know that XP is a phase, to know whether SpecXP is at the phasal edge, hence accessible from the outside, we need to look at the local syntactic context in which that Spec occurs (in particular, the presence of other Specs of XP can effect the edgehood of a particular SpecXP). There has thus been a consistent move toward contextuality in the locality of movement. The contextual approach to the EPP gains theoretical significance within this broader picture. It shows broader relevance of contextuality, contextuality now also being relevant in defining the EPP (in fact, in the same way as for phases and phasal edges—there is a domain for phases/phasal edges/EPP, with the highest phrase in the relevant domain functioning as a phase, phasal edge, locus of the EPP effect). Given that the EPP is involved in determining the driving force of movement, the contextuality then becomes crucial in determining both the locality of movement and the driving force of movement.

In fact, the scope of the contextuality of syntax is even broader than that. In Chomsky (2013), labeling is actually also contextual—the same element behaves differently regarding labeling depending on the context it finds itself in (a phrase behaves differently in a phrase-phrase and a head-phrase merger, it also behaves differently in different phrase-phrase contexts), and its status regarding labeling changes during the derivation (see (13b-a)).

The Bare-Phrase Structure system is also obviously extremely contextual—whether a particular element is a head, a phrase, or an intermediate projection depends on its syntactic context—its status in this respect also changes during the syntactic derivation: thus, what is a maximal projection after a head and a phrase merge becomes an intermediate projection upon further merger.[19] Furthermore, Takita et al 2016 argue that spell-out of a phasal complement essentially removes it from the derivation, which turns the Spec of phase head X into a complement (after the spell-out).[20] There is a parallelism to be made here with the trouble-makers regarding locality of movement. In the bounding node/rigid phasehood approach, one can look at a node itself, without paying attention to anything around it, and determine whether it is a bounding node/phase or not. This is not possible in the *Barriers* system/contextual phasehood approaches. Similarly, in the GB phrase structure, one only needs to look at a particular node to determine its phrase structure status, whether it is a phrase or a head (see e.g. (44), which ignores DP, and vP for the VP-Internal Subject Hypothesis). This is not the case in the Bare-Phrase structure system; looking at any of the *like*-s in (45) itself does not help in determining whether that *like* is a head or a

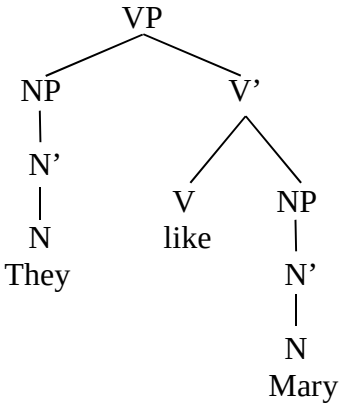
computational system itself. To illustrate, while Abels (2003) argues for a parametric variation regarding whether PP is a phase or not to capture crosslinguistic variation regarding P-stranding within a rigid approach to phasehood (along the lines of Chomsky's original approach), Bošković (2014) shows that the relevant crosslinguistic variation can be captured without any parametric variation regarding phasehood, with a crosslinguistically uniform approach to phases, under a contextual approach to phases, whose structural sensitivity gives it the necessary flexibility to capture crosslinguistic variation of this sort.

[19]See Rizzi (2016) on interaction between Bare Phrase Structure and labeling in this respect. Note that Chomsky (2019a:275) suggests a "condition of *coherence* or *stability* that says that the properties of a syntactic object can't change in the course of the derivation". However, he does not seem to have in mind the kind of issues discussed regarding contextuality above.

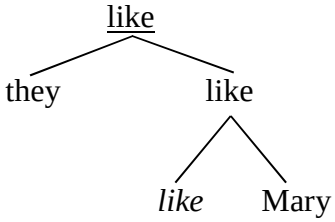
[20]Bošković (in press a) uses this to explain why phasal specifiers are more resistant to diachronic loss than non-phasal specifiers (cf. the general loss of specifiers discussed by Dadan 2019).

phrase, the relevant status being determined contextually (an element that does not project any further (underlined *like*) is a phrase, and an element that is not a projection (*like* in italics) is a head).

(44)



(45)



Furthermore, while the status of a position with respect to the A/A' distinction used to be determined non-contextually (similarly to its phrase structure status), in the phase system the status of a position with respect to the A/A'-distinction is also determined contextually. Thus, as discussed in Bošković (2007), SpecvP is not always treated as an A-position when it is a landing site of successive cyclic movement. In the phase-based system, any movement out of vP has to stop by SpecvP. The status of a SpecvP with respect to the A/A' distinction depends on the nature of movement that stops by SpecvP: if we are dealing with A-movement (meaning the position below and above SpecvP in the relevant chain is an A-position), the SpecvP created by the movement counts as an A-position (the same holds when SpecvP is the landing site of object shift), and if we are dealing with A'-movement (as in the case of wh-movement of adjuncts or long-distance movement of objects out of vP), the SpecvP created by the movement counts as an A'-position. We thus need to look at the larger syntactic context to determine the status of a particular SpecvP with respect to the A/A'-distinction.

The general contextuality is reinforced by the approach to the traditional freezing ban discussed above. Moved elements have been standardly considered to be islands. We have, however, seen that extraction is actually possible out of moved elements (including moved subjects) under certain conditions. Bošković (2018) argues that the same holds for the traditional Adjunct Island, and Bošković (2020) for the Coordinate Structure Constraint (i.e. conjunct/ConjP islands). Bošković (2015, 2016a) extends this to the Complex-NP Constraint, the Specificity Effect, and the Comp-trace effect. Taken together, all this provides us with a broader insight into the nature of islandhood in general; there is actually no islandhood as that notion has been understood traditionally—there are no phrases that by their very nature (and independently of their syntactic context) disallow extraction out of them. In other words, being e.g. a subject or an adjunct does not by itself make the relevant phrase an island. This leads to a heavily contextual view of islandhood, which is abstractly in line with the contextual approach to phases (the contextuality of phases being essentially the same as the contextuality of the EPP). In fact, in a series of works Bošković (2015, 2016a, 2018, 2020) proposes a uniform account of pretty much all island/locality-of-movement effects based on a contextual approach to phases, whose crucial component

is the labeling theory, which is itself heavily contextual. More generally, Bošković shows that successive-cyclic movement has a special status in many islandhood effects, where many islands are confined to successive-cyclic movement (in a similar way in which the traditional moved-element island is confined to successive-cyclic movement), with the labeling theory putting us in a position to understand why that is the case, by providing an explanation for the special status of successive-cyclic movement in this respect.

Putting everything regarding labeling and movement together, we are left with a unified picture where movement and labeling interact rather strongly: unlabeled elements cannot undergo movement, unlabeled elements do not function as interveners, and movement cannot target unlabeled elements (movement also must cross a labeled element under the labeling approach to antilocality). Additionally, the labeling theory provides a uniform account of all traditional islands and locality-of-movement effects. Much of this was made possible because successive-cyclic movement has a special status with respect to a number of locality effects, which could not have been even seen before the labeling theory, the reason being that only the labeling theory treats successive-cyclic movement in general differently from other instances of movement—in a way, it gives it a special status. The labeling theory itself, where labels are essentially contextually determined, is part of a broader move toward context-sensitivity, which permeates many domains, including structure-building and labeling, the A/A' distinction, formulation of locality domains (traditional islandhood as well as the status of phasal projections and their edges), and, as argued here, the EPP (the contextuality of the EPP is essentially the same as the contextuality of phases and phasal edges, being defined in the-highest-phrase-in-the-relevant-domain terms).

References

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

Barbiers, Sjef. 2002. Remnant stranding and the theory of movement. In *Dimensions of movement: from features to remnants*, ed. by Artemis Alexidou, Elena Anagnostopoulou, Sjef Barbiers and Hans-Martin Gartner, 47-67. Amsterdam: John Benjamins.

Blümel, Andreas. 2017. Exocentric root declaratives: Evidence from V2. In *Labels and roots*, ed. by Leah Bauke and Andreas Blümel, 263-289. Berlin: Mouton de Gruyter.

Bobaljik, Jonathan and Susie Wurmbrand. 2005. The domain of agreement. *Natural Language and Linguistic Theory* 23:809-865.

Bošković, Željko. 1994. D-structure, theta-criterion, and movement into theta-positions. *Linguistic Analysis* 24:247-286.

Bošković, Željko. 1997. *The syntax of nonfinite complementation: An economy approach.* Cambridge, Mass.: MIT Press.

Bošković, Željko. 2002. A–movement and the EPP. *Syntax* 5:167-218.

Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59:1-45.

Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38:589-644.

Bošković, Željko. 2008a. On successive cyclic movement and the freezing effect of feature checking. In *Sounds of silence: Empty elements in syntax and phonology*, ed. by Jutta M. Hartmann, Veronika Hegedüs, and Henk van Riemsdijk, 195-233. Amsterdam: Elsevier.

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

Bošković, Željko. 2011a. On Unvalued Uninterpretable Features. In *Proceedings of 39th Conference of the North East Linguistics Society,* ed. by Suzi Lima, Kevin Mullin, and Brian Smith, 109-120. Amherst: GLSA, University of Massachusetts.
Bošković, Željko. 2011b. Last Resort with Move and Agree in derivations and representations. In *The Oxford handbook of Linguistic Minimalism*, ed. by Cedric Boeckx, 327-353. Oxford: Oxford University Press.
Bošković, Željko. 2012. On NPs and clauses. In *Discourse and grammar: From sentence types to lexical categories,* ed. by Gunther Grewendorf and Thomas Ede Zimmermann, 179-242. Berlin: Mouton de Gruyter.
Bošković, Željko. 2013. Phases beyond Clauses. In *The nominal structure in Slavic and beyond*, ed. by Lilia Schürcks, Anastasia Giannakidou and Urtzi Etxeberria, 75-128. Berlin: Mouton de Gruyter.
Bošković, Željko. 2014. Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45:27-89.
Bošković, Željko. 2015. From the Complex NP Constraint to everything: On deep extractions across categories. *The Linguistic Review* 32:603-669.
Bošković, Željko. 2016a. On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling. *The Linguistic Review* 33:17-66.
Bošković, Željko. 2016b. Getting really edgy: On the edge of the edge. *Linguistic Inquiry* 47:1-33.
Bošković, Željko. 2018. On movement out of moved elements, labels, and phases. *Linguistic Inquiry* 49:247-282.
Bošković, Željko. 2019. On subject positions and specifiers more generally. Paper presented at Generative Linguistics in the Old World in Asia 12, Dongguk University, Seoul, August 2019.
Bošković, Željko. 2020. On The Coordinate Structure Constraint, Across-the-Board-Movement, Phases, and Labeling. In *Recent developments in phase theory*, ed. by Jeroen van Craenenbroeck, Cora Pots, and Tanja Temmerman, 133-182. Berlin: Mouton de Gruyter.
Bošković, Željko. in press a. Generalized asymmetry. *Proceedings of 2019 International Conference on Explanation and Prediction in Linguistics: Formalist and Functionalist Approaches.*
Bošković, Željko. in press b. On smuggling, the freezing ban, labels, and *tough*-constructions. In *Smuggling in Syntax*, ed. by Adriana Belletti and Chris Collins, 53-95. Oxford: Oxford University Press.
Branigan, Phil. 1992. Subjects and complementizers. Doctoral dissertation, MIT, Cambridge, Mass.
Cardinaletti, Anna. 2004. Towards a cartography of subject positions. In *The structure of CP and IP*: *The cartography of syntactic structures*, ed. by Luigi Rizzi, 115-165. Oxford: Oxford University Press.
Castillo, Juan Carlos, John Edward Drury and Kleanthes K. Grohmann. 1999. Merge over move and the extended projection principle. *University of Maryland Working Papers in Linguistics* 8:63-103.
Chametzky, Robert. 2000. *Phrase structure: From GB to Minimalism*. Malden, Mass.: Blackwell.
Chomsky, Noam. 1973. Conditions on transformations. In *A Festschrift for Morris Halle,* ed. by Stephen Anderson and Paul Kiparsky, 232-286. New York: Holt, Rinehart and Winston.
Chomsky, Noam. 1986. *Barriers*. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89-155. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 1-52. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In *Structures and beyond*, ed. by Adriana Belletti. 104-131. Oxford: Oxford University Press.
Chomsky, Noam. 2013. Problems of Projection. *Lingua* 130:33-49.
Chomsky, Noam. 2019a. Some puzzling foundational issues: The Reading program. *Special Issue: Generative Syntax. Questions, Crossroads, and Challenges.* 263-285.
Chomsky, Noam. 2019b. The UCLA lectures. Ms. University of Arizona.
Collins, Chris. 2002. Eliminating labels. In *Derivation and explanation in the Minimalist Program*, ed. by Samuel Epstein and T. Daniel Seeley, 42-64. Oxford: Blackwell.
Collins, Chris. 2005. A smuggling approach to the passive in English. *Syntax* 8:81-120.
Dadan, Marcin. 2019. Head labeling preference and language change. Doctoral dissertation, University of Connecticut.
Den Dikken, Marcel. 2007. Phase extension contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33:1-41.
Despić, Miloje. 2011. Syntax in the absence of Determiner Phrase. Doctoral dissertation, University of Connecticut.
Despić, Miloje. 2013. Binding and the structure of NP in Serbo-Croatian. *Linguistic Inquiry* 44:239-270.
Donati, Caterina. 2006. On wh-head movement. In *Wh-movement: Moving on,* ed. by Norbert Corver and Lisa Lei-Shen Cheng, 21-46. Cambridge, Mass.: MIT Press.
Epstein, Samuel. 1992. Derivational constraints on Ā-chain formation. *Linguistic Inquiry* 23:235-259.
Epstein, Samuel and T. Daniel Seely. 1999. SPEC-ifying the GF "subject"; eliminating A-chains and the EPP within a derivational model. Ms. University of Michigan and East Michigan University.
Erlewine, Michael Yoshitaka. 2016. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language and Linguistic Theory* 34:429-479.
Gallego, Ángel. and Juan Uriagereka. 2007. A critique of phase extension, with a comparison to phase sliding. *Theoretical Linguistics* 33:65-74.
Grohmann, Kleanthes. 2003. *Prolific domains: On the anti-locality of movement dependencies*. Amsterdam: John Benjamins Publishing.
Haegeman, Liliane. 1996. Verb second, the split CP and null subjects in early Dutch finite clauses. *Geneva Generative Papers (GenGenP)* 4*:*133–175.
Hornstein, Norbert and Jairo Nunes. 2008. Adjunction, labeling, and bare phrase structure. *Biolinguistics* 2:57-86.
Hunter, Timothy. 2010. Relating movement and adjunction in syntax and semantics. Doctoral dissertation, University of Maryland, College Park.
Johnson, Kyle. 1991. Object positions. *Natural Language and Linguistic Theory* 9:577-636.
Kaisse, Ellen. 1983. The syntax of auxiliary reduction in English. *Language* 59:93-122.
Lasnik, Howard. 1999. *Minimalist analysis*. Oxford: Blackwell.
Lee, Si-Kai. 2021. Agreement-drop in Singlish: Subject to topichood. Presented at West Coast Conference on Formal Linguistics 39.
Lobeck, Anne. 1990. Functional heads as proper governors. In *Proceedings of the 20th North East Linguistic Society*, ed. by Juli Carter, Rose-Marie Dechaine, William Philip, and Timothy Sherer, 348-362. GLSA, University of Massachusetts, Amherst.
Martin, Roger. 2001. Null Case and the distribution of PRO. *Linguistic Inquiry* 32:141-166.
McCloskey, James. 2000. Quantifier float and wh-movement in an Irish English. *Linguistic Inquiry* 31:57-84.
Messick, Troy. 2020. The derivation of highest subject questions and the nature of the EPP. *Glossa* 5:1-12.

Miyagawa, Shigeru. 2003. A-movement scrambling and options without optionality. In *Word order and scrambling*, ed. by Simin Karimi, 177-200. Oxford: Blackwell.
Reis, Marga. 1997. Zum syntaktischen Status unselbständiger Verbzweit-Sätze. In *Sprache im Fokkus*, ed. by C. Dürscheid, 121–44. Tübingen: Niemeyer.
Richards, Norvin. 2001. *Movement in language: Interactions and architectures.* Oxford: Oxford University Press.
Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.
Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. by Liliane Haegeman, 281-337. Dordrecht: Springer.
Rizzi, Luigi. 2004. Locality and left periphery. In *Structures and Beyond: The Cartography of Syntactic Structures*, ed. by Adriana Belletti, 223-251. Oxford: Oxford University Press.
Rizzi, Luigi. 2006. On the form of chains: Criterial positions and ECP effects. In *Wh-movement: Moving on*, ed. by Norbert Corver and Lisa Lei-Shen Cheng, 97-134. Cambridge, Mass.: MIT Press.
Rizzi, Luigi. 2016. Labeling, maximality and the head–phrase distinction. *The Linguistic Review* 33:103-127.
Roberts, Ian. 2004. The C-system in Brythonic Celtic languages, V2, and the EPP. In *The Structure of CP and IP*, ed. by Luigi Rizzi, 297-328. Oxford: Oxford University Press.
Saito, Mamoru and Keiko Murasugi. 1990. N'-deletion in Japanese: A preliminary study. In *Japanese/Korean Linguistics* ed. by Hajime Hoji, 285-301. Stanford, Calif.: Stanford Linguistic Association, Stanford University.
Schneider-Zioga, Patricia. 1995. Specifier/head agreement in Kinande. *Cahiers linguistiques d'Ottawa* 23:67-93.
Starke, Michal. 2001. Move Dissolves into Merge: a Theory of Locality. Doctoral dissertation, Université de Genève.
Takahashi, Masahiko. 2011. Some consequences of Case-marking in Japanese. Doctoral dissertation, University of Connecticut.
Takita, Kensuke, Nobu Goto and Yoshiyuki Shibata. 2016. Labeling through spell-out. *The Linguistic Review* 33:177-198.
Travis, Lisa. 1991. Parameters of Phrase Structure and V2 Phenomena. In *Principles and Parameters in Comparative Grammar*, ed. by Robert Freidin, 339-364. Cambridge, Mass.: MIT Press.
Uriagereka, Juan. 1999. Minimal restrictions on Basque movements. *Natural Language and Linguistic Theory* 17:403-444.
Webelhuth, Gert. 1992. *Principles and Parameters of Syntactic Saturation.* Oxford: Oxford University Press.
Wurmbrand, Susi. 2014. The merge condition. In *Minimalism and beyond: Radicalizing the interfaces*, ed. by Peter Kosta, Steven Franks and Teodora Radeva-Bork, 130-167. Amsterdam: John Benjamins.
Yoo, Yong Suk. 2015. Labeling algorithm and its consequences on movement. Paper presented at the 17th Seoul Conference on Generative Grammar, Kyung Hee University, Seoul, August 5th-8th.
Zwart, Jan-Wouter. 1993. Verb movement and complementizer agreement. In *MIT Working Papers in Linguistics* 18:297-340. MITWPL, MIT, Cambridge, Mass.