Generalized Asymmetry

Željko Bošković
University of Connecticut

Abstract: The paper argues for a Maximize Asymmetric Relations preference (MAR) as a general property of language based on a number of phenomena that are independent of word order. In addition to expanding the domain of asymmetricity, a number of mechanisms and phenomena are unified from this perspective, with their reason for existence traced back to MAR, namely the diachronic loss of specifiers, the Linear Correspondence Axiom, the Phase Impenetrability Condition, the no-Spec-without-complement aspect of Bare Phrase Structure, the rarity of multiple Spec constructions (as with, e.g. multiple wh-fronting), and the *who left* effect (where subject wh-movement cannot proceed through SpecTP). MAR is also shown to favor approaches where movement is moving-element driven over those where movement is target-driven as well as Bare Phrase structure building over GB structure building, and to have consequences for the proper formulation of several mechanisms, including the Phase Impenetrability Condition, Case licensing, and the EPP. The paper also discusses MAR within a broader formalism vs functionalism setting, within a unificational perspective where both of these approaches have a place.

Keywords: Bare Phrase Structure, diachronic change, functionalist approaches, formalist approaches, LCA, Phase Impenetrability Condition, typology, Universal Grammar, wh-movement

Prolegomenon

This paper grew out of one part of a talk given at the 2019 Heidelberg formalists vs functionalists workshop. There were actually very few formalist talks; most of the talks took a functionalist perspective, and argued against formalist accounts of particular phenomena. Still, as someone who was characterized as a formalist,[1] I found many of those talks quite interesting. In many cases, the issues that were raised for particular formalist analyses were real, and in a number of cases the talks presented very interesting data that any account of the relevant phenomena should integrate. One issue was the general approach in some of the talks that the formalists and the functionalists cannot both be right in principle, that there is room for only one of these approaches to language (as a result, taking e.g. a particular formalist account of phenomenon X and showing that there were problems with it, would lead to the conclusion that all formalism is wrong). To me, that in principle seemed wrong. I would be the first one to argue against any formalist who would argue that functionalism as a whole is wrong, that there is no aspect of language that can be explained

[1] While I would prefer to be called a theoretical syntactician, even that label gets me to squirm a little. We have reached the point where linguistic phenomena need to be looked at in their totality, not compartmentalized by specific subfields. A syntactician has to be constantly paying attention to semantics, morphology, phonology, language change, and language acquisition. Doing theoretical syntax also means doing syntax-semantics interface, syntax-morphology interface, syntax-prosody interface, looking seriously into understudied languages and typology, and always be ready to bring in diachronic change and language acquisition facts. Typology is particularly important here. We have reached the point where typology should be at the center of investigation of what generativists refer to as Universal Grammar (UG). At this point of our understanding, broad crosslinguistic comparisons and Greenberg-style typological generalizations are actually more enlightening regarding the nature of language and UG than detailed investigations of individual languages (the latter are of course a prerequisite for the former; the practitioners of the latter are, however, often a reviewing stop sign for typological works since they often require the same kind of detailed investigation for every individual language considered in a typological work that would be found in a work devoted to just one language). The talk presented at the workshop was intended to reflect all this; while typology was at its center, it also addressed a number of prosodic, semantic, morphological, language acquisition, and language change issues (only the last one (and to some extent language acquisition) is reflected in this paper, which elaborates on a small part of the workshop talk). This was also reflected in the title of the workshop talk, "On the nature of language and linguistics: it's all just linguistics" (the handout for the talk can be found at: https://boskovic.linguistics.uconn.edu/wp-content/uploads/sites/2801/2020/08/FormalistFunctionalistShort.pdf).

by looking at the function of language: the complexity of language is way too great to be able to handle it comprehensively only with one of these points of view. There is no competition between the two approaches in general, but there is with respect to particular analyses of particular phenomena, since we don't a priori know in which domain a particular phenomenon belongs to.

At any rate, the tone of the final roundtable became much more productive. But this is what I actually found to be the most inspiring part of the whole workshop, in fact so inspiring that it reminded me of the early days of minimalism. Minimalism lead to an Occam-razor based re-evaluation of all the basic mechanisms which were taken for granted in the Government and Binding (GB) era: levels of representation, the way syntax interacts with semantics and phonology, basic structure building, word order…. everything that was taken for granted was questioned. The kind of questions that were raised at the conference, which many formalists would have perceived as hostility and/or unjustified attacks, actually reminded me of those minimalist days. They were questioning even the most basic things, exactly the way we were doing it in the early days of minimalism. True, there was occasionally a bit of an underlying tone (due to the skepticism regarding the viability of the overall formalist approach) which we did not have when we were questioning GB (more precisely, ourselves, since we were all GB practitioners), but it was worth ignoring that to appreciate the questioning itself. And it was productive, it got me to think of certain issues in a different, simpler way.

In one part of the talk, I discussed Marcin Dadan's dissertation. I tried to stay as neutral as possible theoretically. The main point was that diachronic change often involves loss of movement, which Dadan tied to the loss of specifiers. Dadan deduces this from the labeling framework of Chomsky (2013), by comparing different aspects of Chomsky's labeling algorithm. I did not go into any of the technical details of Dadan's analysis (so there was no discussion of labeling), but I still got a question, can we do all this without using the technical term specifier. The question was reminiscent of the early minimalist days. This is exactly the kind of a question we were asking regarding the concepts that were taken for granted in the GB theory. The question did get me to think about the relevant issues differently, which I would not have done otherwise, in fact in a more general way which establishes a connection with other phenomena that otherwise could not be related, i.e. which could not all be related through the notion of specifier. This paper is a result of that. While the paper is still rather "technical",[2] the point is that it grew out of what I would consider a fruitful "formalists vs functionalists" (labels which I am not fond of) interaction—in several places it also appeals to functional explanations and in fact leaves open whether the main point should ultimately have a formalist or a non-formalist explanation (with a more general formalism vs functionalism discussion in the last section of the paper, whose point is that the two approaches are not as incompatible as they are standardly assumed to be and can in fact inform each other). If it turns out to be the latter, the paper can be looked at as an example of how a theoretical investigation of formal properties of language can elucidate connections with broader cognitive properties.

1. Introduction: Asymmetricity

The focus of this paper will be on the notion of asymmetricity in language. There are many cases of asymmetric relations in language, many of which have been pointed out in Kayne's work (see e.g. Kayne 1994, 2010). Linear order is obviously asymmetric—if X precedes Y, it cannot be the case that Y precedes X—but this is certainly not the only asymmetric relation. Kayne (2010) presents a more general case that our language faculty (FL) has the property of being asymmetric, though most of the cases he provides still concern word order (i.e. correlations between various syntactic phenomena and word order). Much of his argumentation concerns the lack of what we

[2]An attempt will be made to confine some of the most "technical" aspects that are not crucial to the central points of the paper to the footnotes, a point which more "technically" oriented readers should bear in mind.

would expect to find if FL were symmetric in the domain of word order. Thus, he points out that there is no pair of languages x,y where y is the mirror image of x such that for any sentence of x, the corresponding sentence of y would be its mirror image in word order.

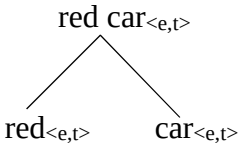
But there are also asymmetric relations outside of word order, which is what this paper will be concerned with. For one thing, the notion of the head of a phrase, more precisely, the unique head of a phrase, is inherently asymmetric: it says, informally, that one element in a phrase is more important than others. One can easily imagine FL, and the concept of structure, without the notion of the head of a phrase. In fact, we do not need imagination for that. Until the rise of the X-bar theory, the sentence was assumed to be S, with its immediate daughters being NP and VP—S simply did not have a head; we needed the X-bar theory to force headhood on it.[3]

We find abstractly similar situations in semantics, with the lack of the counterpart of headhood from the X-bar theory approach to structural relations. Consider e.g. Heim and Kratzer's (1998) Predicate Modification rule:

(1) For any branching node α whose daughters are β and γ, if both β and γ are of type <σ, t>, then [[α]]=[λxσ. [[β]] (x) and [[γ]](x)], where σ is any type.

To informally illustrate the working of (1), in (2) *red* and *car* are β and γ from (1); they are both of type <e,t>; the object we get by combining β and γ here, α from (1), is also of type <e,t>.

(2)



Looking at the semantic composition in terms of types, there is really no head for the phrase in (2): "red" and "car" as well as the phrase "red car" are all of type <e,t>.

The point of the above discussion is that the notion of the (unique) head of a phrase is inherently asymmetric—it in fact represents a case of asymmetry outside of word order. It seems real, though one can certainly imagine structures without it. However, it does not seem to be the case that there are no symmetric relations in FL. In fact, even the notion of c-command, which Kayne (1994) uses to determine word order, which is by its very nature asymmetric, is not inherently asymmetric (as Kayne 1994 himself notes)—it is not the case that there cannot be two nodes/constituents such that they c-command each other. True, one can impose asymmetricity on it by brute force (i.e. definitionally), which is what Kayne (1994) in fact does, but the point is that the notion itself is not inherently asymmetric.

In some cases, there has been a debate whether a particular mechanism is asymmetric or not although the debate was actually never framed in such a way, hence the relevance of the broader issue under consideration here for the mechanisms in question was never explicitly noted. Consider e.g. Case. Under the GB-style Case assignment implementation of Case licensing, as well as under the current Case valuation approach to it (see Chomsky 2000, 2001), Case licensing is asymmetric (informally, I do something to you, and you don't do that to me[4]); under the early minimalist approach in terms of Case checking, it was in fact symmetric (informally, we do it to

[3] There have also been post X-bar theory proposals for structures without a head, where it was assumed that such structures can be generated but cannot survive—with movement forced to destroy such symmetric (i.e. lacking a head) structures—see in this respect especially Moro (2000) (see also Ott 2012).

[4] More technically, in GB the verb assigns case to its nominal object (instrumental in the case given in (i) in fn 5), and in Chomsky (2000, 2001), the verb values the Case of its object (giving it a particular Case specification)—the nominal object does not do anything of that sort to the verb.

each other), which led to the so called Inverse Case Filter (see Bošković 1997), a requirement that traditional Case assigners check (i.e. assign) their Case.[5] Case licensing is then another example of an asymmetric relation outside of word order, but only under the Case-valuation/assignment view, not under the Case-checking view.

Without outright denying that symmetric relations can at all exist, but taking the kind of considerations that Kayne (and others; see also the above discussion) have brought up seriously, takes us to the position that FL favors asymmetric relations, i.e. it leads us to (3), where MAR is a preference principle (in a sense to be made clear below), and the domain where MAR holds is the computational system (informally syntax), including spell-out itself.[6]

(3) Maximize Asymmetric Relations (MAR)

This paper will argue for (3), based on phenomena independent of word order. Arguing for asymmetric relations is of course not new.[7] What is new is the kind of phenomena that will be

[5] Under Case checking, there is no inherent difference between e.g. a verb and a nominal element regarding Case in a Case-licensing relation—they check Case against each other. Thus, under this approach, the verb and its object NP both have the instrumental Case feature in (i) and they check each other's Case feature.

(i) On je ovladao Andorom.
he is conquered Andorra$_{INSTR}$
'He conquered Andorra.' (Serbo-Croatian)

Under the Case checking approach we would expect that two traditional Case assigners can check Case against each other—Bošković (2006) in fact reports a rather clear instance of that sort, illustrated in (ii), where a verb and a preposition check Case against each other in Serbo-Croatian (SC). (Higher numerals, which do not show Case distinctions in SC, assign genitive to the following noun, which must be in genitive; (iia) is ruled out because the instrumental Case feature of the verb is not checked; in (iib) it is checked by the preposition "with", which on its own also licenses instrumental Case on the following nominal (iic). The problem with (iid) is, then, that if the verb and the preposition check Case against each other, the instrumental of the noun remains unchecked).

(ii) a. *On je ovladao pet zemalja.
he is conquered five countries$_{GEN}$
b. On je ovladao sa pet zemalja.
he is conquered with five countries$_{GEN}$
'He conquered five countries.'
c. sa zemljom
with country$_{INSTR}$
d. *On je ovladao sa zemljom.
he is conquered with country$_{INSTR}$ (Bošković 2006)

While the Case checking approach nicely accommodates the data in (ii), if this approach were right, we would expect to find such cases all over the place; this, however, is the only example of that sort that I am aware of (there is also the issue of morphological realization—thus, the Case feature is morphologically realized on the NP but not on the verb in (i)—the two are not symmetric in this respect). The reader is also referred to Bošković (2011a) for arguments against the Inverse Case Filter, which can be interpreted as arguments for an asymmetric approach to Case, like Case valuation. (Since in the Case-checking approach there is really no formal difference between Case Filter and Inverse Case—the two are just descriptive terms—arguments against Inverse Case Filter are also arguments against the Case-checking approach.)

[6] Assuming that the Predicate Modification rule applies in the semantics, it would not be relevant to MAR.

[7] For a position similar to (3), see Di Sciullo (2015, 2017). ((3) is argued for on rather different and broader grounds here though; furthermore, Di Sciullo actually does not treat MAR as a preference principle in the sense argued for in this paper; in this particular respect Di Sciullo's position is more similar to Moro 2000 than to the current work (it is in fact even stronger than Moro's approach in that for Di Sciullo, antisymmetry holds at each step of the derivation). Still, Di Sciullo 2015 is an important predecessor of this work).

looked at from this perspective in this paper; in fact, a number of superficially rather different phenomena, which come outside of the domain of word order, will be brought together under this perspective here (the discussion will also shed new light on some of these phenomena). It should be obvious that the position taken here, MAR, is weaker than Kayne's (2010) position that FL is fully asymmetric; however, the discussion here will apply to a broader domain, going considerably beyond issues regarding word order, which is what Kayne was concerned with. The discussion in the paper will thus expand the domain of asymmetricity. As a result, I will also refer to (3) below as Generalized Asymmetry. However, given the nature of the paper, the discussion will be to some extent speculative and programmatic—I will not be able to examine the relevant phenomena comprehensively but will only discuss the aspects of these phenomena that are relevant to our main concern, i.e. (3). I will also not concern myself here with the issue of what (3) could follow from; however, the issue of whether (3) can be traced back to FL external factors will be briefly addressed in section 6 (as noted there, Kayne 2010 suggests a more general connection with parsing and language production; see also Di Sciullo 2011 for a broader biolinguistic perspective).

In addition to providing a unifying perspective on a number of phenomena, we will see that MAR has additional consequences in that it favors certain approaches to particular theoretical mechanisms over their alternatives. In fact, we have already seen this with respect to Case licensing, where Case licensing reflects the spirit of MAR under the Case valuation approach, but not under the Case checking approach. (The notion of the head of a phrase also reflects the spirit of MAR.) We will also see that MAR has relevance for several aspects of the Phase Theory. Much of the discussion below will concern specifiers, which we will see are particularly relevant to MAR. I will thus start the discussion below by examining a rather interesting issue concerning specifiers in language change, noted by Dadan (2019), which will lead us to examine the nature of specifiers more broadly. Regarding the latter, among other things, MAR will be shown to provide evidence against the implementation of the EPP as a formal requirement to have a Spec (and more generally, against formal requirements to have a Spec). This will in turn lead to a more general discussion of the formalist vs functionalist perspective to language in the conclusion of the paper, with an overarching view where both of these approaches have a place.

2. Specifiers diachronically and synchronically:

Diachronic change often involves loss of movement (see e.g. Roberts 1993, 2007, van Gelderen 2009, 2011). Dadan (2019) observes that this is in fact the general direction of diachronic change. Dadan gives a number of cases illustrating this; I will only give one illustration here. There are many examples of this kind of change regarding wh-dependencies, where Dadan makes a very interesting observation that the general direction of the diachronic change is from wh-movement to wh-in-situ, not the other way round. Thus, there is a loss of obligatory *wh*-movement from **Old Japanese to modern Japanese** (Ogawa 1976, Whitman 2001, Watanabe 2002, Kuroda 2007, Aldridge 2009, 2018), from **archaic to modern Chinese** (Aldridge 2010, 2011), from **Vedic Sanskrit to modern Indic languages** (Hale 1987, Fortson 2004), or from **Latin**, which was actually a multiple wh-fronting language (Spevak 2010, Danckaert 2012, Ledgeway 2012) **to modern Romance**, wh-in-situ being possible as an option in modern Romance[8] but it wasn't possible at all in Latin (see Dadan 2019 and references therein). There is also an on-going change in **Navarro-Labourdin Basque** (Duguine and Irurtzun 2014). Dadan observes that what the loss of wh-movement leads to is the loss of a specifier. (Another case of this sort is the loss of V-2, which also involves movement to SpecCP, as in e.g. Old Romance (Wolfe 2018) and English (Roberts 1997); see Dadan's work for other cases, one of which is noted below (11) regarding the OV to VO word order change).

[8] With the exception of Romanian, a Bulgarian-style multiple wh-fronting language which moves all wh-phrases to interrogative SpecCP (see Rudin 1988).

There is another way to lose a Spec, without the loss of movement itself. Bošković (2001) observes different behavior of the Q/focus marker *li* in Serbo-Croatian (SC) and Bulgarian, which can be captured if the Q/focus marker *li* has lost its ability to support a specifier in SC. In particular, Q/focus marker *li* in SC cannot host unambiguously phrasal elements (4a-b) or license sluicing (4c), which requires a Spec-head relation (see Lobeck 1990 and Saito and Murasugi 1990). On the other hand, both of these are possible in Bulgarian (5).

(4) a. *Novu kuću li prodaje?
new house LI sells?
'Is he selling the new house?'
b. Novu li kuću prodaje?
c. *Novu li ~~kuću prodaje~~ ?

(5) a. Novata kŭšta li prodade?
new-the house LI sold
'Did he sell the new house?'
b. Novata kŭšta li ~~prodade~~? (Bošković 2001)

What is going on here is that movement to *li*, which is an enclitic hence it needs something in front of it to support it prosodically, still must take place in SC, but it takes place through head-adjunction to *li*, hence the one-word restriction on the host of *li* and *li*'s inability to license sluicing, which is licensed through a Spec-head agreement relation (see Lobeck 1990 and Saito and Murasugi 1990). In Bulgarian, both phrasal elements in front of *li* and sluicing are possible, indicating that the two indeed go hand-in-hand. SC *li* has thus lost the ability to take a specifier. (In fact, this usage of *li* is archaic in SC—it appears that the first step in the loss of the construction in question is in fact the loss of the Spec).

Another way of losing specifiers is to reanalyze them as heads. This is especially prolific in the domain of complementizers, where phrases in SpecCP get reanalyzed as complementizer heads. Here are some illustrations noted by Dadan (2019) (there are many cases of this sort, spec-to-a-head change is in fact quite common even outside of the domain of complementizers, see especially van Gelderen 2004).[9]

(6) **Georgian** *wh*-phrase *ray* 'what'>complementizer *raytamca* (Harris & Campbell 1995; this is quite frequent crosslinguistically); **Russian** *čto* 'what'>čem 'than' (comparison complementizer, Willis 2007); **English** *how*>subordinating complementizer (Huddleston & Pullum 2002, also many Slavic languages e.g. **Polish**, **Slovak** *jak*, and **Breton** *penaos*); **German** complementizer *dass* from relative pronouns in SpecCP (Axel-Tober 2017; also common in e.g. Slavic, Meyer 2017; and Greek, Roberts and Roussou 2003); **French** *par ce que* 'by this that'>*parce que* 'because' (van Gelderen 2004); **Early Germanic** *hwæt* reanalyzed as an exclamative C-head (Walkden 2014).

Another case of this is the emergence of agreeing complementizers from pronouns in Welsh, e.g. complementizer *mi* derives from a 1SG subject pronoun, and the particle *fe* from a masculine 3SG subject pronoun (see Willis 2007). The former is illustrated by (7). What facilitated this change was pronoun doubling, as in (8), where a pronoun occurs both in its base position and in the left periphery of the clause—the latter then got reanalyzed as an agreeing complementizer, as in (7).

(7) Mi welais I 'r gêm
PRT see.PAST.1SG I the game

[9] Wang (2019) argues that there is an intermediate stage in the Spec-to-a-head change, where the relevant element is base-generated adjoined to another head, before it projects a phrase on its own.

(8) Mi arhosais (,) fi
1SG.IND wait.PAST.1SG 1SG.IND
'I waited, me.' (Willis 2007: 459)

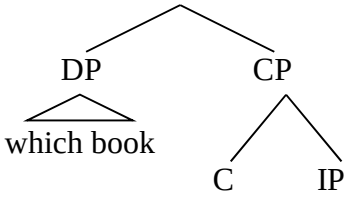
What we see in all these cases is the loss of a Spec. Dadan (2019) deduces this from the labeling framework of Chomsky (2013), arguing that the way structure building works there favors head-complement relations over traditional Spec-head relations, which require an additional step to label the object in question (agreement or movement (see fn 21 regarding Chomsky's 2013 labeling approach); for another labeling-based approach that applies to the Spec-to-head reanalysis in particular, see van Gelderen 2015).[10] I will, however, pursue here an alternative, broader way of explaining the preference for the loss of specifiers, which in fact will not appeal to the notion of *specifier* per se but will provide a more general explanation that will establish a connection with other phenomena that all this otherwise cannot be related to (some were in fact already mentioned in section 1 regarding the notion of the head of a phrase and a particular approach to Case licensing).

The head-complement relation involves merger of two elements that are not equal in their phrase structure status, one is a head and the other one is a phrase. This is not the case with the traditional Spec-head relation. In the Bare Phrase Structure system (Chomsky 1995), what we have in that case is the merger of two phrases, at the point of the merger itself. Consider (9).

(9) Which book did John buy?

The relevant step of the derivation before wh-movement takes place first involves merger of C, a head, and IP, a phrase, which yields a phrase, CP. The wh-phrase then merges with this object.

(10)



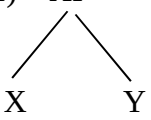
What we then have with wh-movement is a merger of two phrases, *which book* and the CP in (10). This in fact holds quite generally: traditional Specs involve a merger of two phrases in the Bare Phrase Structure system. The suggestion, discussed in section 1, is that syntax quite generally prefers asymmetric relations (cf. (3)), this is why it prefers head-complement over Spec-head relations: the former involves an asymmetric merger, i.e. it involves merger of a head and a phrase; the latter involves a symmetric merger, i.e. it involves merger of two phrases. This is then the reason why the diachronic change in the case at hand (i.e. wh-dependency) involves the loss of wh-movement, not its gain. By eliminating a Spec, the former eliminates a case of a phrase-phrase merger. On the other hand, the latter would involve creation of a Spec, hence gain of a Spec, which would mean an additional phrase-phrase merger.[11]

There is an immediate connection here with another proposal, namely Kayne's (1994) antisymmetry of syntax, which is the proposal that word order is read off asymmetric c-command

[10] But see fn 32 for an issue that arises with the labeling framework regarding MAR.

[11] It should be noted that Kayne (2010) simply bans merger of two phrases: "The merger of two phrases is unavailable" (see also Narita 2012). This illustrates the general difference between Kayne (2010) and the position taken here, discussed in section 1, where FL is taken to prefer asymmetric relations (this point will be made even more clearly below).

relations, where, roughly, if X asymmetrically c-commands Y, X precedes Y, and everything dominated by X precedes everything dominated by Y.[12] In Chomsky's (1995) reinterpretation, this proposal led to elimination of word order from the syntax—syntax is all about structural relations like dominance and c-command, word order is then imposed in PF due to the nature of the PF interface, which interacts with our articulatory-perceptual system, which by its very nature requires word order. In particular, word order is imposed by linearization of structural relations, where asymmetric c-command plays a crucial role (this is referred to as LCA). At any rate, the LCA rules out all symmetric structures (symmetric in a sense to be made more precise below). Under Chomsky's version of the LCA, they can be created but they have to be eliminated before spell-out.[13] Thus, in the Bare Phrase Structure system, a non-branching element is both a head and a phrase. If such an element is merged as a traditional complement, as in (11), we get a structure that is too symmetric: a problem which is resolved by moving Y in (11) (so that Y does not have to be linearized in the original position, given that it is not pronounced in that position). In a sense, then, the movement here is driven by MAR.[14]

(11)


Both the diachronic tendency to lose specifiers and Kayne's LCA can then be looked at as the preference for asymmetric relations, and therefore unified from that perspective.

There is in fact a case, noted briefly above, where the two are quite clearly brought together. Kiparsky (1996) observes that the OV-to-VO word order change is way more common than the VO-to-OV word order change (see Biberauer and Roberts 2006, Roberts 1997, 2007, and Dadan 2019). From the perspective of Kayne (1994), the OV word order is derived from the VO word order, with object movement (see e.g. Zwart 1997, who analyzes it in terms of object shift; regarding the change itself, see Kiparsky 1996, Roberts 1997, 2007, Danckaert 2012, Dadan 2019, among others). The OV-to-VO change then in fact involves a loss of movement and results in the loss of a Spec. The relationship between the OV and VO word order is then the same as the relationship between wh-fronting and wh-in-situ, with the same direction of diachronic change.

Also relevant here are several cases of diachronic change noted in Di Sciullo (2011) (see also Di Sciullo, Nicolis, and Somesfalean 2020), of the form depicted in (12)-(13), where the situation in which αP either precedes or follows head X leads to a diachronic change where αP only follows head X. One case of this sort concerns genitive theme complements of nouns in Greek: in Classical Greek, they could either precede or follow the head noun, while in Modern Greek they must follow the head noun (see Alexiadou 2002).

(12) a. αP X
b. X αP
(13) X αP

[12] This leads to a universal Spec-head-complement base order; any departures from this order then must result from movement.

[13] Kayne actually argues that the LCA holds throughout syntax, which means that symmetric structures could not even be created (note, however, that Kayne does not assume Bare Phrase Structure).

[14] As noted in fn 3, Moro (2000) and Ott (2012) argue that XP and YP can also be base-generated as sisters (with neither of them projecting). In that case one of them has to move away for the same reason movement has to take place in (11), namely because the base-generated structure in question is too symmetrical.

From the current perspective (and parallel to the discussion of the OV-VO variation above), (12) would be interpreted as involving optional movement of αP (in particular, the movement would take place in (12a), but not (12b) (see also Alexiadou 2002), with the movement getting lost in the stage depicted in (13). In fact, as discussed in section 3, this may be what is happening with wh-movement in Modern Romance: while Latin was an obligatory wh-fronting language (so in stage (12a) regarding wh-movement), Modern Romance languages have optional wh-movement (see below for a formal implementation of this change), abstractly showing the stage depicted in (12a-b) in this respect. As discussed in section 3 (see Dadan 2019), there are reasons to believe that Modern Romance languages are on their way to becoming wh-in-situ languages: this would then also represent a change from (12) to (13).[15]

A question then arises why all specifiers don't get lost. That would essentially lead to the loss of movement, so the question is actually broader: why do we have movement in the first place. Chomsky's (2000:120-121) position here is that this has to do with "externally imposed legibility conditions", i.e. it is due to "conditions imposed by the external systems". What this means is that the reason for it is essentially functional, or more broadly non-syntactic: to be able to express notions that go beyond the basic argument structure (which is what we would have without specifiers): more complex semantic notions involving issues like scope/scopal ambiguities, pragmatic notions concerning things like topic/focus interpretation, specificity…,[16] in fact even argument structure that goes beyond a simple predicate with one internal argument requires a specifier (see section 4). At any rate, as noted by Dadan (2019), from this perspective, cases of e.g. gaining specifiers may be expected to be non-syntax-driven, i.e. interface-driven and/or attributed to extra-syntactic factors, e.g. prosody or pragmatics/semantics (see Dadan 2019 for a more detailed discussion). Note, however, that the preference nature of MAR does leave room for a syntactically-driven gain of specifiers.

Consider in this respect the change from a *pro*-drop to a non-*pro*-drop language that occurred in e.g. French (see Adams 1987). Romance-style *pro*-drop is licensed by rich verbal morphology; what is relevant here is that there was a change in the richness of verbal morphology, which led to a problem with the licensing conditions on *pro* (i.e. the phonologically null pronominal subject). Now, while the exact analysis of traditional *pro*-drop is certainly controversial (see Roberts and Holmberg 2010 for an overview) a number of authors (e.g. Borer 1986, Alexiadou and Anagnostopoulou 1998, Barbosa 1995) have argued that traditional *pro*-drop

[15] It should be noted that Di Sciullo (2011) also argues that asymmetry plays a role in diachrony, though she uses asymmetry in this respect in a very different sense from the one used here. For Di Sciullo, it is a matter of eliminating choice, where two possible word orders are reduced to one of these two word orders—this is what Di Sciullo refers to as asymmetry in diachronic change. This particular sense of asymmetry would not extend to many relevant cases of the loss of specifiers (some traditional grammaticalization cases (cf. (6)) do not even involve a change in word order, cf. also the case of *li* in (4)-(5)). It is also not concerned with the issue of which of the two word orders survives.

[16] Chomsky (2000) in fact associates these notions with specifiers. Some of these may have led to the development of formal requirements (which would then force movement, as discussed below; note that non-syntactic factors could ultimately be behind crosslinguistic differences in this domain, e.g. it is possible that what is behind the different syntactic behavior of wh-phrases in Bulgarian and Japanese is that they are subject to different interpretation, see section 6). There could actually also be prosodic reasons for movement, e.g. to support an enclitic head like *li* in (5a) (for a much broader proposal along these lines where prosodic factors motivate movement, see Richards 2010, 2016). The notion of a canonical order may also be relevant here. E.g., shift from OV to VO may be less likely to happen if SOV is the canonical order in the language, which means that it would be used in a discourse-neutral way to answer a question like 'What happened?' (note that canonical word order does not entail lack of movement: to note just one case, following Dryer (1992) Kayne (2010) observes that there are languages where S-O-Neg-V is the canonical word order (the order would have to involve movement even if Kayne's 1994 approach is not adopted).

does not actually involve a null argument in a specifier (i.e. it does not involve *pro*); rather, verbal morphology itself is the argument. Under this analysis, losing *pro*-drop actually involves a gain of a specifier (lexical subjects are specifiers), but in a manner that is fully consistent with MAR as a preference principle. In fact, under this analysis of *pro*-drop, the loss of *pro*-drop in French can be taken as a confirmation of the status of MAR as a preference principle (see sections 1 and 4).

It should also be noted that Dadan (2019) argues that a pattern similar to the diachronic tendency to lose specifiers is also found in language acquisition. More precisely, he argues that many cases of errors in child language acquisition actually arise due to the avoidance of Specs. In other words, the diachronic tendency to lose Specs is reflected in language acquisition as a tendency to analyze structures in a way which would avoid Specs.[17] This is not at all surprising under the approach to the issue under consideration discussed above. It seems plausible that children are poor in those extra semantic/pragmatic notions which require (hence justify) specifiers, hence the MAR strategy is even more strongly at work in child language.[18]

A number of other issues may also be relevant here. Consider the semantics of multiple wh-questions. While this is certainly a hotly debated issue, a number of authors have argued that the most transparent and simplest syntax-semantics mapping in this domain is provided by multiple wh-fronting languages (see e.g. Pesetsky 1987), where all wh-phrases front overtly (creating an operator-variable relation), as in Bulgarian (14), which is analyzed in terms of multiple specifiers of CP, with each wh-phrase located in a separate specifier (see Koizumi 1994, Richards 2001).

(14) Kogo kakvo kak e pital?
whom what how is asked
'How did he ask who what?'

Given this, one might expect the multiple wh-fronting strategy to be quite common. However, very few languages actually employ it (see Bošković 2012 for a list). This may not be surprising in light of the discussion above: the dispreference for specifiers is particularly relevant here, since constructions like (14) involve multiple specifiers of CP.

In fact, the issue in question seems to be quite general. In Chomsky's (1995) bare phrase structure, there is nothing special about multiple Spec constructions, in fact one would expect them to be quite common. Curiously, an obvious point has never been made in this respect before: such cases are in fact quite rare crosslinguistically.[19] From the current perspective, all this may be due to the general dispreference for specifiers. Recall that creation of a traditional specifier involves merger of two phrases: with multiple specifiers, creation of each specifier involves merger of two phrases—multiple Spec constructions are thus particularly offensive to the preference for asymmetric relations. As discussed above, there is pragmatic/semantic/prosodic pressure not to lose all specifiers; this pressure is weaker regarding multiple Spec constructions since in many cases creation of a single Spec suffices to express the relevant pragmatic/semantic notions (or at least decreases the need for another Spec), or do the relevant prosodic job (support an enclitic).

[17] On the relationship between language acquisition and language change, see Lightfoot (1979), van Gelderen (2011), Roberts (2007), among others.

[18] It may be worth noting here that Uriagereka (2012) argues that all Specs are islands, disallowing movement out of them. If this is correct (the issue is controversial—thus, there is a controversy regarding whether extraction is possible out of subjects in SpecvP—Uriagereka argues, contra Takahashi 1994 and Stepanov 2001, that it isn't)), it is possible that the avoid-the-Spec strategy results in islandhood: Spec-creation creates a dispreferred configuration from which extraction is not possible.

[19] The rarity of the multiple wh-fronting strategy is a particularly drastic case of this if this strategy indeed yields the most transparent and simplest syntax-semantics mapping for multiple questions, hence should otherwise be favored on these grounds.

3. On the Phase-Impenetrability Condition

All of this may also help us gain a new perspective on the Phase-Impenetrability Condition (PIC), in fact deduce it from generalized asymmetry.

Phases define locality domains in the syntax and determine the points of spell-out, i.e. when the syntactic structure is sent to the interfaces (on both the meaning and the sound side). Syntactic dependencies are not unbounded—they are subject to locality effects. Certain phrases count as phases (more on this below)—those phrases determine the locality domains for syntactic dependencies. Under the standard approach to phases/phase-based locality effects, the Spec of a phase is accessible for movement outside of the phase; the complement of a phase, which is what is sent to spell-out, is not (this is what is referred to as the PIC). In other words, in a phase-based derivation, Spec of phase XP is essentially in a different locality domain from the rest of XP. This can actually be looked at as a way of resolving the Spec conundrum discussed above: Spec is separated from the rest of the structure into a different locality domain, reducing the problem that Specs raise for the asymmetric nature of syntax if such burden is actually computed domain by domain, as is natural in the derivation by phase.

The above suggestion implies that when the PIC pushes a Spec into another domain, it is not really a Spec in the new domain, which essentially means that the exact same full structure is not present in the new domain, so that when the relevant element is pushed into another domain, it has a different status. In other words, the PIC separates a Spec so that it is not in a Spec configuration any more. Interestingly, a number of authors have independently made proposals that accomplish exactly that, in particular, Epstein (2007, 2009), Chomsky (2008), Goto (2013), Narita (2011, 2012), and Takita, Goto, and Shibata (2016). Consider the last work. Under the multiple spell-out hypothesis (Uriagereka 1999, Epstein 1999, Chomsky 2000, among many others), syntax interacts with the interfaces during the syntactic derivation, by repeatedly sending chunks of structure to the interfaces as the syntactic derivation unfolds. Phases determine the points of spell-out, i.e. when the syntactic structure is sent to the interfaces. Under standard assumptions, spell-out occurs at the phasal level, with the phasal complement being what undergoes spell-out.[20] Takita et al (2016) suggest that spell-out essentially removes the phasal complement, changing the syntactic object {X, YP} (where X is the phasal head and YP its complement) into a single head X (for a trivial reason, since YP is no longer there). They present a number of arguments for this view (for relevant discussion, see also Goto 2013, Narita 2011, 2012, Epstein 2007, 2009), one of their concerns being a problem that arises in Chomsky's (2013) labeling system with successive-cyclic movement, as in e.g. (15) (CP is a phase, which means that movement out of CP must proceed via SpecCP, otherwise the wh-phrase will not be accessible outside of the embedded CP due to the PIC, as discussed above), where the structure cannot be labeled after *which book* merges in the position of t' (which involves merger of two phrases), the derivational point shown in (16), due to the lack of agreement/feature-sharing between the relevant elements.[21] To deal with this,

[20] Bošković (2016b) argues that what undergoes spell-out is actually a full phase, with successive-cyclic movement targeting the phrase right above the phase. The discussion in the text can be easily adapted to that approach.

[21] Chomsky's (2013) labeling approach assumes that labeling of an object that results from a merger of two elements X and Y is not an automatic result of the application of the Merge operation: unlabeled objects are in fact allowed during the syntactic derivation but not in the final representations which are sent to the interfaces, because the interfaces themselves require labels. Chomsky then proposes an algorithm for determining labeling of objects formed by merger, the gist of which is that when a head and a phrase merge, the head determines the label of the object in question, while in the case of a merger of two phrases the elements in question need to undergo agreement to enable labeling. Crucially, as argued extensively in Bošković (1997, 2002b, 2007, 2008b), Chomsky assumes that successive-cyclic movement, such as the movement to the embedded clause edge in (15), does not involve agreement, which means that labeling

Chomsky essentially stipulates that traces are invisible to labeling, so that the structure is labeled (as CP) after *which book* moves away (see (15)).

(15) Which book$_i$ do you think [$_{CP}$ t'$_i$ that John bought t$_i$]
(16) [$_{DP}$ which book]$_i$ [$_{CP}$ that [$_{IP}$ John bought t$_i$]

What Takita et al's proposal regarding spell-out does here is change the syntactic object {C(*that*), IP} into a single head C(*that*) (after the IP is sent to spell-out). The label of the syntactic object that corresponds to the embedded clause of (15) at the point when *which book* is present in that part of the structure can then be determined straightforwardly even before the wh-phrase moves away given that this syntactic object now consists of a head (C) and a phrase (the wh-phrase), eliminating the need for labeling through traces (i.e. the assumption that traces are invisible for labeling, which Takita, Goto, and Shibata 2016 show is problematic; note that the head-phrase configuration can be labeled in Chomsky 2013, with the head providing the label, see fn 21).

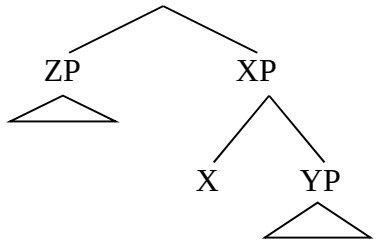
The most obvious argument for the proposal in question, however, concerns the standard assumption that only the edge of a phase is accessible from outside of the phase. The reason why this is the case is then rather straightforward: only the edge is actually there. To see this more clearly, what Takita, Goto, and Shibata (2016) argue is that when spell-out applies to (17) (where XP is a phase and YP is what is spelled out), it essentially changes (17) to (18) (since by removing YP from (17), it changes the syntactic object {X, YP}, which is XP in (17), into a single head X, as shown in (18)). The other authors cited above make similar proposals. Thus, Narita (2011, 2012) argues that spell-out removes a constituent from the derivational workspace so that what remains after spell-out applies to (17) (it applies to YP in (17)) is (18) (Chomsky 2008 in fact also suggests that the PIC effect arises because what is spelled out is eliminated). Note that in the Bare Phrase Structure system, discussed in more detail in section 4, where there are no vacuous projections and the phrasal status of an element is determined contextually, removing YP from (17) changes the phrasal status of the sister of ZP—the sister of ZP is no longer a phrase, but a head, as shown in (18).[22] The idea behind the proposals in question is thus rather simple and appealing: if something

through agreement is not possible here. As noted directly in the text, Chomsky assumes that in such a case, one of the relevant elements needs to move away so that it can be ignored for the purposes of labeling.

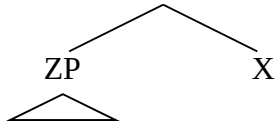
[22]See Narita (2011, 2012) for discussion of how the information that X was merged with YP is encoded and accessed in the interfaces under this approach (Narita argues the information that YP has undergone Merge with X in (17), including the relevant c-command relations, is also transferred, which is important in the recombination of separately transferred bits of structures (see Boeckx and Grohmann 2007), e.g. for the purpose of linearization. The alternative is to treat linearization in the relevant respect like Chomsky (2013) treats labeling, which is rather natural given that Chomsky (2013) treats labeling just like linearization in that labels are needed only at the interfaces hence merger of syntactic objects proceeds without labeling. Chomsky argues that labeling is done at the phasal level, for the whole phase, prior to transfer/spell-out. (Under the Takita et al analysis, if at this point labeling fails due to the lack of feature-sharing, it can reapply after the relevant transfer operation, which is in fact what Chomsky 2013 also assumes.) Given the conceptual similarity between linearization and labeling, it seems natural to assume that the same would hold for linearization. The effect of this is that the c-command relations relevant to linearization would be read off at the derivational point in (17) (note that the issue of linearization matters only if we are dealing here with the final landing site of ZP and if XP is phase; it does not arise e.g. with successive-cyclic movement). There are other rather interesting options here, but exploring them in detail would take us beyond the scope of this paper. E.g., Kayne (2010) suggests linear order is determined by probe-goal relations, where the probe-goal search shares the directionality of parsing and production, hence the probe precedes the goal (a head then precedes the complement since it probes the complement (this is generalized over selector-selectee dependencies) or an element within the complement). What is of interest here is that in Bošković's (2007) system, discussed in section 5.2, all movement is moving-element driven, with the moving element serving as a probe from its final landing site (the moving element has a uK feature which

is not accessible it is really not there (in fact, more generally, this is the best and simplest way of dealing with the kind of effect where something (YP in this case) behaves as if it is not there—it really is not there). What matters for us is that this changes the phrase-phrase merger from (17) into a head-phrase merger in (18).

(17)



(18)



The above discussion (i.e. MAR), then, gives us a new perspective on the PIC. A Spec involves a symmetrical, phrase-phrase merger. The PIC in effect reintroduces asymmetry into the merger (compare (17) and (18)). The above then amounts to a deduction of the PIC—it is seen as a mechanism for maximizing asymmetry of syntax.[23]

It should be also noted that under the above approach to the PIC, the PIC is not a condition that explicitly states what is (in)accessible outside of phase α: all we have is that the complement of the phase head is sent to spell-out, which removes it from the derivational workspace. This is all—(in)accessibility trivially follows from that (whatever is present in the derivational workspace is accessible, what is not present is trivially not accessible[24]).

---

forces it to move to serve as a probe). What this means is that if what we see in (17)-(18) is the final landing site of ZP (which is the only relevant case since otherwise ZP would not be pronounced in this position, making linearization irrelevant), ZP will function as a probe in Bošković' system, which means that under Kayne's (2010) assumption regarding probe-goal relations, ZP would precede XP in (17) (in fact, nothing would change in (18)). Another assumption from Kayne (2010) opens up another possibility. Kayne suggests that if a head H merges with two complements, they must be on the opposite side of H. Given the bottom-up derivation, (17) first determines that X must precede YP. ZP then must be on the opposite side of X even in (18). Another possibility is made available by a proposal from Narita (2011), the gist of which is that in {α, β}, if there is a lower copy of α within β, α precedes β. There are thus a number of very interesting options here. However, teasing them apart, which would require their precise formalization, would take us beyond the limits of this paper, hence I leave this potentially rather rewarding endeavor for future research.

[23]As observed by Takita et al (2016), under their proposal labeling considerations cannot be the driving force behind successive-cyclic movement as in Chomsky (2013) (Chomsky's analysis in this respect was based on the stipulation that traces are ignored for labeling, which, as noted above, is now eliminable). They also observe that labeling anyway cannot be taken to be the driving force of movement in general, e.g. no labeling problem arises in the base position of the direct object in *Mary$_i$, John likes t$_i$*. Takita at al (2016) adopt Bošković's (2007) proposal, discussed in section 5.2, that movement is driven by an uninterpretable feature (uK) of the moving element, which applies both to the initial step of movement and successive-cyclic movement (one could, however, try to incorporate this into the labeling system under Bošković's in press assumption that a uK blocks labeling—a moving element would then always have a uK, which would block labeling (with feature-sharing eliminating the uK)).

[24] Regarding actual (in)accessibility, nothing changes from Chomsky's (2000) "definitional" approach to the PIC, which simply states, as a matter of definition, that what is accessible from the outside is the Spec

There is a similarity between the diachronic loss of specifiers and the PIC that should be noted: while the two are superficially very different, like the former, the PIC also leads to the loss of a specifier, i.e. undoing of a phrase-phrase merger situation. One could in fact look at the PIC as a derivational manifestation of the diachronic pressure to lose specifiers.

A question, however, arises in this respect whether the avoidance to create a specifier can be the motivation for the diachronic loss of wh-movement if the relevant specifiers are anyway reanalyzed derivationally during transfer/spell-out. Two issues are relevant in this respect. For Chomsky (2000, 2001), CP, but not lower clausal projection (such as IP), is a phase. However, wh-movement does not seem to target SpecCP, i.e. a phasal edge, in all languages (see in this respect section 5). In fact, in some of the cases where wh-movement got lost, it appears that the landing site of obligatory wh-fronting was actually lower than SpecCP (i.e. it did not target a phase edge, which means that the relevant Spec would not have been PIC-reanalyzed), see in this respect Aldridge (2018), who argues that wh-movement in Old Japanese targeted SpecIP[25] (see also Watanabe 2002 and Aldridge 2009, as well as Aldridge 2010, 2018 regarding Archaic Chinese).

Second, even in the cases where wh-movement does create a Spec of a phase (i.e. where it lands in SpecCP), in the case of a derivational loss of a specifier that is accomplished through the PIC the specifier is still first created (cf. the discussion of (15)-(18) above). With the diachronic loss of wh-movement, the specifier is never created. This means that the loss of wh-movement is a stronger way of satisfying MAR than the PIC; the avoidance to create a specifier can then still be the motivation for the loss of wh-movement in spite of the role of the PIC described above.

At any rate, languages do seem to vary in the landing site of wh-movement (see section 5).[26] If this is indeed true, and if in some cases wh-movement does target a phasal edge while in others it does not, we may expect that the latter would be more likely to be lost given that the PIC would relieve the Spec pressure to some extent in the former case. While the prediction still remains to be verified, it should be noted that some of the cases where wh-movement got lost indeed seem to have involved wh-movement of the latter kind. Thus, Aldridge (2009, 2010, 2018) argues that this was the case in Archaic Chinese and Old Japanese. While we cannot go into a detailed discussion of the position of wh-phrases in Archaic Chinese and Old Japanese here, the position of the wh-phrases in (19) (wh-fronting was obligatory in Archaic Chinese) rather clearly indicates that the landing site of the wh-fronting is lower than in English, which is exactly what Aldridge argues.[27]

(19) a. Gong **shei$_i$** yu xiang t_i?
you who want appoint
'Who do you want to appoint?'
b. Wo jiang **he$_i$** qiu t_i?
I will what ask.for
'What will I ask for?' (Archaic Chinese, Aldridge 2010:10)

---

and the head of a phase. The same elements are still (in)accessible, it's just that the status of one of these elements has been changed derivationally from a Spec to a complement.

[25]She argues SpecIP in this case becomes a wh/focus position through Chomsky's (2008) C-T inheritance.

[26]To mention just one case, while in many languages with focal particles wh-phrases move in front of the focal particle, which means to the Spec of the particle, Yuan (2017) shows that in Kikuyu they move to the complement of the focal particle (as a result of which they follow it; see also Tuller 1992 for relevant crosslinguistic variation within Chadic languages).

[27]See Aldridge's work for arguments that this is indeed the case, in Archaic Chinese as well as Old Japanese. Regarding the latter, while wh-phrases had to be fronted in Old Japanese, they were very often not clause initial (note that Aldridge shows that it is not the case that only topics could precede fronted wh-phrases in Old Japanese). This is very different from Latin, to be discussed below, where wh-phrases were typically clause initial (see the data in Brown, Joseph, and Wallace 2009, Danckaert 2012, Davin and Stephens 2006, Ledgeway 2012, Spevak 2010).

Also relevant is Romance. As noted above, Latin was an obligatory wh-fronting language. Modern Romance languages, on the other hand, allow the wh-in-situ option (Latin did not allow it at all), in fact quite productively. As noted by Dadan (2019), the option is actually on the rise in terms of the frequency of occurrence. Thus, Coveney (1995) explicitly argues that this is the case for French. In fact, although French has more than one overt movement strategy, Coveney (1995) has found that the wh-in-situ option is employed in nearly 40% of questions in the speech of Parisian middle class (see also Lopes-Rossi 1996 for Brazilian Portuguese). The wh-in-situ option is even more prevalent in child speech (in fact, even older children, who have fully mastered the wh-option, use wh-in-situ more frequently than wh-movement, and crucially more frequently than their parents, using it even in the contexts where wh-in-situ is not allowed in adult French, see Zuckerman 2001, Oiry 2011). Based on all this, Dadan (2019) argues that we are witnessing here a change in progress, with French being in the process of becoming a wh-in-situ language.

What is important for us is that Ledgeway (2012) suggests that there was a change in the landing site of wh-fronting from Latin to Modern Romance; in particular, the landing site of wh-fronting in Latin was the highest clausal projection (see also Danckaert 2012), which is not the case in Modern Romance, where Ledgeway (2012) assumes that wh-fronting lands in Rizzi's (1997) FocP. What this means is that in Latin, wh-fronting targeted a phasal projection, which is not the case in Modern Romance.[28] [29] In fact, Danckaert (2012: 245-250) suggests that Latin was a multiple wh-fronting language of Bulgarian type (cf. (14)), with fronted wh-phrases clustering together in a superiority-obeying manner clause initially (cf. Rudin 1988 on two types of multiple wh-fronting languages). In the literature on multiple wh-fronting, this is generally taken as a diagnostic of wh-fronting to the highest clausal projection (see e.g. Rudin 1988, Bošković 2002a, Richards 2001). This means that there was a change in the landing site of wh-fronting, from a phasal to a non-phasal projection, before the development of wh-in-situ in Romance.

At any rate, while it is impossible to be conclusive in this respect, there is suggestive evidence (especially from (Archaic) Chinese and (Old) Japanese) that wh-fronting is indeed more likely to be lost if it targets a non-phasal projection, which can be captured under the suggestion

[28] Rizzi (1997) proposes splitting the traditional CP into the projections shown in (i) (all these projections dominate IP under Rizzi's proposal)

(i) [$_{\text{ForceP}}$ [$_{\text{TopP}}$ [$_{\text{FocP}}$ [$_{\text{TopP}}$ [$_{\text{FinP}}$

Adapting Rizzi's split CP structure (developed mostly on the basis of Italian), where ForceP, not FocP, is the highest clausal projection, to Chomsky's (2000, 2001) phasal system or the contextual, the highest-phrase-in-the-clause-is-a-phase system (e.g. Bošković 2014, 2015, Wurmbrand 2013), FocP would not be a phasal projection in either of these approaches to phases (ForceP would be). In fact, I am not aware of any approach to phases where that would be the case. Note that more generally, Ledgeway (2012) suggests that Latin did not have split (or fully split) CP, which modern Romance languages do have (for independent evidence that there is crosslinguistic variation in this respect, see e.g. Bošković 2016a, Erlewine 2016; note that Rizzi 1997 himself suggests that CP is not always split, in fact even within a single language, and that the works assuming split CP for particular languages argue that there can be variation in the landing site of wh-movement across languages even within split CP, see e.g. Haegeman 2000 and Roberts 2004).

[29] Note also that in Spanish (which is also developing wh-in-situ), a traditional complementizer can in fact precede a fronted wh-phrase (see e.g. Uriagereka 1988, Rizzi 2001, Villa-García 2015), as illustrated by (i), which confirms that the wh-fronting here does not target the highest clausal projection (see also Bošković 2002a, Reglero 2007, Reglero and Ticio 2013, Figueiredo Silva and Grolla 2016, among others, for arguments that wh-in-situ in modern Romance languages is true wh-in-situ; it should not be analyzed as involving wh-movement that is followed by remnant fronting).

(i) Julio preguntó que qué$_i$ íbamos a comprar t$_i$
Julio asked that what (we) were.going to buy

made in this section since the PIC would relieve the Spec pressure at least to some extent with phasal Specs, as discussed above.[30]

Another issue is potentially relevant here. Rizzi (1997) suggests that even in languages where wh-movement targets FocP, a non-phasal projection as discussed above, wh-movement in relative clauses still targets the highest phrase in split CP (his ForceP, see fn 28), which is a phasal projection (see fn 28) . Interestingly, it appears that wh-movement is harder to lose in relative clauses; thus, modern Romance languages still require it in relative clauses; in fact, even Chinese still has wh-movement in relative clauses (see e.g. Huang 1982). This could be another case where a phasal Spec is more resistant to a loss than a non-phasal Spec, though obviously a much more careful investigation is needed before this conclusion can be endorsed.[31]

Another point is worth noting here. One of the tenets of the minimalist program is that language (i.e. Universal Grammar) is characterized by optimal, computationally efficient design. Phases and multiple spell-out are taken to contribute to the efficient design, i.e. they are efficient design mechanisms. Early research within the generative paradigm has already noticed 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 multiple spell-out, accomplishes this by limiting the number of syntactic objects/the amount of syntactic structure that 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). Phases determine the transfer points, the PIC playing a crucial role here.

Phases and multiple spell-out not only limit the amount of structure that the derivation is working on, but they also maximize the MAR effect by eliminating Specs (by changing their status, as discussed above). From this perspective, the more phasal points we have, the better (for both concerns under consideration). There are a number of different approaches to phases; I will leave it to the reader to compare them from the perspective of these concerns (see e.g. Chomsky 2000, 2001, Bošković 2014, Epstein and Seely 2002, and Müller 2010).

The above approach to the PIC also has a bearing on the proper formulation of the PIC. Following the original multiple-spell out proposal by Uriagereka (1999), Bošković (2015) argues for an approach to the PIC where both the Spec and the complement of phase XP are accessible from the outside (though nothing that is dominated by these elements is). This conception of the PIC would not follow from the maximize-asymmetry-approach to the PIC: complements do not raise a problem for the asymmetry of syntax; furthermore, this approach does not sever the Spec from the rest of the structure, by placing it in a different domain. As a result, if the maximize-asymmetry approach to the PIC is on the right track, the conception of the PIC where only the Spec is accessible from the outside is to be preferred.

The above discussion has thus unified the diachronic tendency to lose Specs, the avoidance of Specs in language acquisition, the LCA, and the PIC: all of these are there because of the asymmetric nature of syntax. Superficially, we are dealing with very different mechanisms but abstractly they all have something in common, namely MAR. The diachronic loss of specifiers (which is essentially reflected in language acquisition), the LCA, and the PIC are all different strategies for dealing with a symmetric merger situation: with the first one, one of the relevant elements is lost, with the second one, movement of one of the elements is forced, and the third one changes the status of one of the relevant elements—crucially, they all target and change a

[30] All this may also be expected to have a reflex in language acquisition, an issue which will have to be left for future research.

[31] It is also possible that a non-syntactic issue, namely the creation of a predicate for the 'head' of the relative clause, is involved in the case of relative clauses.

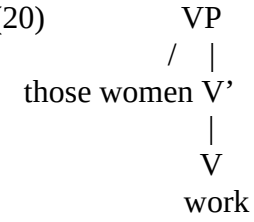
symmetric merger situation. We have also seen that that bringing the PIC into the diachronic loss of specifiers makes a prediction that non-phasal specifiers may be more likely to get lost than phasal specifiers. The current discussion also has consequences for the PIC and phases: it favors one particular approach to the PIC and favors approaches to phases that maximize phasal points. (Recall that MAR also favors a particular way of implementing Case licensing.)

4. Bare Phrase Structure

In this section I will discuss basic structure building from the perspective of MAR. We have seen in section 1 that the notion of the head of a phrase itself reflects the spirit of MAR. In this section we will see that the MAR perspective also provides an argument for Chomsky's (1995) conception of Bare Phrase Structure (BPS), which also favors it over GB-style structure building. (Recall also that the MAR perspective explains the rarity of multiple Spec constructions, which BPS otherwise freely allows.)

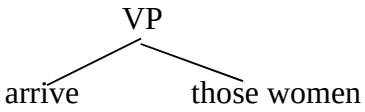
Chomsky (1995) proposes a relational definition of Specs and complements where the first element merged with a head is a complement, everything else is a Spec. This in itself favors complements over Specs (capturing the MAR intuition); in fact, there cannot be a Spec unless there is a complement. GB structure building was not like that, it was perfectly fine to have a Spec without a complement, as in (20) (under the Predicate Internal Subject hypothesis).

(20)



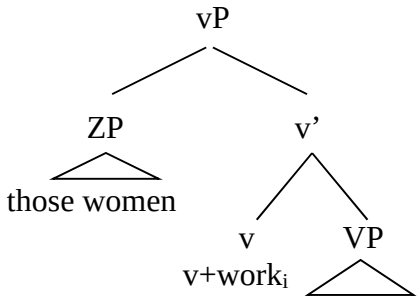
This is not possible in BPS. Attempting something of this sort would only give us a structure that is appropriate for an ergative verb, where the sole argument is base-generated as an object, i.e. a complement (see (21), where VP is used for ease of exposition; the same holds for the bar-level in (22)). The reason for this is simple: there cannot be a Spec unless there is a complement in BPS, which, as noted above, captures the MAR intuition by favoring complements.

(21)



In fact, this is what gave rise to vP as a projection that introduces the external argument: if the external argument is going to be a Spec, the head that introduces it must take a complement, otherwise it could not take a Spec (vP is then there essentially due to MAR concerns).

(22)



The intuition behind all of this is that Specs are created when there is no more space within a phrase, they are sort of last resort in structure building: first comes the complement, whose merger into the structure is asymmetric; if needed, we then get a Spec. The "last resort" character of Specs (they are there only when there is no more space within a phrase) was not present in the GB structure building, which does not favor complements over Specs; hence, to the extent that it is real, MAR can be taken to favor BPS.

To complete the discussion of base argument structure building, compare simple transitive and ditransitive constructions in (23)-(24) (where only the traditional VP structure is presented, before any verb movement).

(23) [bought a book]
(24) [Mary give a book]

A single internal argument can be merged as a complement (*a book* in (23)); this is not possible with the second internal argument in (24) (*Mary*), where creation of a specifier is then forced by semantic reasons (the creation of the Spec in (24) then does not violate the MAR preference). As noted above, under standard assumptions, external arguments are Specs, but in that case the creation of a Spec is also unavoidable, given that the relevant head, v, also needs to take a complement.

A number of things then get unified from the MAR perspective: the diachronic loss of Specs and their avoidance in language acquisition, the LCA, the Phase Impenetrability Condition, and the no-Spec-without-complement aspect of Bare Phrase Structure (regarding structure building, the latter is in fact brought together with the notion of the head of a phrase and the rarity of multiple-Spec constructions).[32] We have seen that the Bare Phrase Structure system is in fact characterized by an avoid-a-Spec-if-you-can property, which is exactly the spirit of MAR, in fact MAR as a preference principle, as argued here.

5. On intermediate movement effects

[32] It is worth noting that there is an aspect of Chomsky's (2013) labeling framework (see fn 21) that goes against the spirit of MAR. We have seen in section 1 that the notion of the head of a phrase expresses MAR—it is inherently asymmetric in that it makes one element in a phrase, or one element in any merger, more important than the other(s). While in BPS all structure building is asymmetric in that one element in a merger always projects (labeling the resulting structure, thus functioning as the head of the resulting structure) this is not the case in Chomsky (2013). E.g., in (i), when *Mary* and IP merge what labels the resulting structure in Chomsky (2013) is prominent features they share, namely φ-features. The two elements thus contribute equally to the structure building here. Similarly, the merger of *what* and the CP in (ii) (*I wonder what she bought*) is labeled by the shared feature, Q, with the two elements again contributing equally to structure building. This is all in contrast to the BPS system, where only one element projects in each merger, labeling the resulting structure.

(i) [<φ, φ> [$_{DP}$ Mary] [$_{IP}$ left]]
(ii) …. [<Q,Q> [$_{DP}$ what] [$_{CP}$ she bought]]

The structures in (i) and (ii) raise questions, e.g., there is the issue of how < φ, φ> in (i) is interpreted in the semantics (note that Chomsky 2013 actually argues that the semantics, not syntax, needs labels. The same issue may arise with object shift, since when object shift takes place, the resulting structure would presumably be also labeled as <φ, φ>). Putting these issues aside (it is worth noting here that the works in the labeling framework often adopt traditional labels like IP and CP for (i)/(ii) for ease of exposition, though the issue is whether this is really just for expository reasons), the point made here is that the BPS structure building is more in line with MAR than structure building in the labeling framework, so MAR can actually be taken to favor the BPS system over both GB structure building and the labeling-framework structure building, though for different reasons.

This section is somewhat speculative and open ended. Its goal is to note one particular consequence of the above discussion, i.e. MAR, which due to the scope of this work (as well as the controversial nature of the issues under discussion) cannot be discussed in any real detail here.

5. 1. Intermediate steps of movement

Above, we have seen that there is a deep dislike for specifiers. What is behind it is the general preference for asymmetric relations. In light of the above discussion, where we have seen that there is a tendency to lose specifiers diachronically and change their status derivationally due to the preference for asymmetric relations, we would not expect to have free, superfluous specifiers. As noted above, the existence of Specs is related to the broader question why we have movement in the first place (Chomsky's 2000 answer is that this is due to the needs of the external systems); most of the time they are used to express various semantic and pragmatic notions (see also fn 16). There can also be prosodic reasons for them, e.g. to support an enclitic head. But there are other considerations too. Consider successive-cyclic movement, in particular, consider (25), focusing on one intermediate step, namely, movement to the intermediate SpecCP.

(25) Which book do you think [*t* that John bought]?

When *which book* moves to merge to the position indicated by *t* in (25) we get a merger of two phrases. In this case, there are no non-syntactic reasons of the kind discussed above that would motivate creating the dispreferred phrase-phrase merger. The reason why the spec in question is created is syntactic, namely due to syntactic locality. Since CP is a phase, *which book* would not be able to move out of the CP without moving through its edge.[33] The Maximize Asymmetric Relations (MAR) is a preference principle, it says that such relations should be maximized as much as possible—here it is simply not possible. Under this approach, we would then expect successive-cyclic movement to occur only when it is really necessary, namely, when it is forced by the PIC, which means that successive-cyclic movement should proceed only through phasal edges. In other words, there should be no free successive-cyclic movement. For arguments that this is indeed the case, the reader is referred to Kang (2014). The position will not be defended here, the issue is too controversial and involves a number of constructions—anything even remotely approaching a conclusive discussion of the issue would go way beyond the scope of this paper, whose goal regarding this particular issue is simply to point out one consequence of MAR and, additionally, to discuss a case (referred to below as the *who left* effect) that was not considered before from this perspective. Regarding arguments for potential free successive-cyclic movement (i.e. successive-cyclic movement that does not proceed via phasal edges) in the literature, such arguments should either be reanalyzed in a way that does not involve successive-cyclic movement, as is done for a number of such cases in Epstein and Seely (2002, 2006), or there should be more phasal boundaries than is standardly assumed so that the movements in question actually target phasal edges (in this respect, see e.g., the claim from Bošković (2014, 2015) and Wurmbrand (2013) that the highest clausal projection is a phase, which means that even IPs that are not dominated by CP, as in the case of raising and ECM infinitives (under standard assumptions), are phases; note also that under Bošković's 2014 approach to phases, on which all lexical heads project phasal domains, even passive and ergative verbs, as well as nouns, prepositions, and adjectives, project phasal domains). At any rate, given that intermediate movements involve creation of specifiers, given the above discussion we would expect that there would be no superfluous intermediate movement steps.[34]

[33] Given that what is sent to spell-out is no longer accessible to syntactic computation, a moving element needs to move to the phasal edge, and out of the phasal complement before the complement is sent to spell-put. Successive-cyclic movement then must target phasal edges.

[34] Superfluous here should be taken rather broadly. In fact, the discussion we are about to get into indicates that 'superfluous' should not be only defined in terms of phases.

This may also help us address the *who left* effect (and more generally, Bošković's 2008b claim that feature-checking movement cannot feed another feature-checking movement)—local subject questions of this sort in fact provide a rather dramatic illustration of the ban on superfluous intermediate steps, which goes beyond phasal considerations. Consider the following paradigm (for discussion of the paradigm see Bošković 2016a, Messick 2020 and references therein).

(26) Who left?
(27) a. *Who bought what the hell?
b. What the hell did John buy?
c. Who the hell bought that house?
(28) Who loves everyone? (who>everyone; *everyone>who)
(29) Someone loves everyone. (someone >everyone; everyone>someone)
(30) Someone bought a car. Who?

Questions like (26) are sometimes assumed not to involve wh-movement at all (see e.g. Carstens, Hornstein, and Seely 2016, Chomsky 1986). There is, however, evidence that the wh-phrase in (26) does not remain in SpecIP. Very briefly, if we take (27a-b) to indicate that *the hell* can only modify wh-phrases in SpecCP, (27c) provides evidence that *who* in (26) does not stay in SpecIP (see Ginzburg and Sag 2000; Pesetsky and Torrego 2001). Furthermore, *everyone* can take scope over the subject in (29) but not in (28). This is unexpected if the subject in (28) could stay in SpecIP (see Mizuguchi 2014). Finally, if sluicing involves wh-movement followed by IP deletion, as is standardly assumed, the wh-phrase in (30) cannot be located in SpecIP (see Agbayani 2000, Messick 2020; the latter also shows that (30) is not a case of pseudoslucing, i.e. ellipsis of an underlying cleft). (27)-(30) thus provide evidence that *who* does not stay in SpecIP in (26).

Furthermore, in a number of languages that allow both the SV and the VS order, where in the latter the subject does not move to SpecIP, the two orders are associated with different subject-agreement morphology. What we get in *who left* in such languages is the morphology associated with the VS order (e.g in some dialects of Italian). This indicates not only that the subject in subject questions does not remain in SpecIP, but that wh-movement to SpecCP cannot even proceed through SpecIP, otherwise we would get the morphology associated with the SV word order. The same point can be made regarding languages where the usual subject agreement morphology that is associated with subjects being in SpecIP has to be dropped in *who left* (e.g. Kinande, Kaqchikel).

Consider also British English *do*-ellipsis, where *do* co-occurs with a modal. It has been noted that A-movement out of a *do*-ellipsis site is allowed, while A'-movement is not, as (31) shows (e.g. Baltin 2007, Haddican 2007, Bošković 2014, den Dikken and Griffiths 2018, Messick 2020).

(31) a. John might seem to enjoy that, and Pete$_i$ might (do) ~~seem t_i to enjoy that~~ too.
b. I know who$_i$ John will kiss and who Pete will (*do) ~~kiss t_i~~

Importantly, such ellipsis is also disallowed with subject questions (see den Dikken and Griffiths 2018, Messick 2020): if wh-movement in subject questions were to proceed via SpecIP, (32b) would involve only A-movement out of *do*-ellipsis, just like (31a), hence would be expected to pattern with (31a) rather than (31b).

(32) A: Sue wouldn't kiss Peter last night
B: Well, who$_i$ would (*do) ~~t_i kiss him~~ (den Dikken and Griffiths 2018)

Another, new, argument to this effect concerns the well-known fact (see e.g. Bresnan 1971, Selkirk 1972, Kaisse 1983) that auxiliary contraction is not possible when the auxiliary is followed by a

wh-trace (in work in preparation I show that this holds when the auxiliary and the wh-trace are located in the same phase).

(33) a. I know where$_i$ John is t_i (tonight).
b. *I know where$_i$ John's t_i (tonight).

The fact that auxiliary contraction is allowed in (34) then indicates that wh-movement in (34) does not proceed via SpecIP, leaving a wh-trace in that position.[35]

(34) Who's leaving tonight?

The following West Ulster English (WUE) data, noted by McCloskey (2000), provide a rather strong confirmation that local subject questions do not involve wh-movement via SpecIP.

(35) Who$_i$ was arrested all t_i in Duke Street?
(36) *They$_i$ were arrested all t_i last night.
(37) What$_i$ did he say all t_i that he wanted?

In contrast to Standard English, WUE allows quantifier (Q)-float under wh-movement, as shown by (37); such Q-float is also possible in (35). Still, just like standard English, WUE disallows (36). (36) indicates that a subject in SpecIP cannot float a quantifier in the postverbal position in passives. This rules out the derivation where *who* in (35) moves to SpecCP via SpecIP. If that were the case, the quantifier in (35) would be floated under movement to SpecIP, which (36) shows is not possible. (This also rules out the derivation where *who* in (35) stays in SpecIP). These data then provide evidence that *who* does not even pass through SpecIP in (26), which is in fact what McCloskey (2000) concludes. How come?[36]

Rizzi (2006) argues that SpecIP is actually a criterial position (like e.g. Spec of +whC, Spec of FocP…); what this means is that movement to this position (i.e. being in this position) leads to a certain interpretation (the same has been argued for object shift, see e.g. Diesing 1996).[37] Under this approach, non-syntactic reasons are then (at least partially) behind creation of SpecIP, which would in essence mean that this movement is not taking place for a strictly formal reason. But this non-formal, interpretation-related reason, which fits well with the above discussion regarding why we have movement, could apply only if the element actually stays (and is interpreted) in that position; if the element has to move away for other reasons, this non-syntactic reason would not apply. Given that IP that is dominated by CP is not a phase, phases/PIC would also not require

[35]It is occasionally suggested that subject questions exceptionally do not involve inversion due to the lack of *do*-support. The conclusion is erroneous: *do*-support is a last resort mechanism that takes place to support a stranded tense affix when a phonologically realized element intervenes between the affix and the verb (an account that goes back to Chomsky 1957, see also Lasnik 1995b, Halle and Marantz 1993, Bobaljik 1995, among many others). There is no phonologically realized intervener in *Who walked* (cf. Who C+**T**(**ed**) **walk;** where C+T indicates T-to-C movement), just as in *Mary walked* (cf. Mary **T**(**ed**) **walk**), and in contrast to *What did Mary buy* (cf. What C+**T**(**ed**) Mary **buy**). Only the last case then triggers *do*-support.

[36] The question is particularly interesting in light of the fact that English is not a pro-drop language and requires SV word order, which is standardly taken to indicate that English requires filling the SpecIP position, the standard assumption being that this is a formal requirement.

[37] I refer the reader to Rizzi (2006) for discussion of the Subject Criterion, i.e. the interpretation that is associated with the position in question (including the status of traditional expletives under this approach, though see Moro 1997 for a semantically-contentful-element approach to expletives). Rizzi in fact explicitly considers the traditional EPP to be a manifestation of the Subject Criterion, comparing it in this respect to the situation found with e.g. TopP and FocP.

movement to that position. Given that intermediate movements take place only when forced by phase/PIC reasons, then movement to SpecCP would not even proceed via SpecIP, which captures the *who left* effect. The reason why there is no movement through SpecIP in *who left* is then the same as the reason why specifiers are lost diachronically, and in fact, more abstractly, it is the same reason as the one behind the LCA and the PIC: MAR, or the general asymmetric nature of language, which disfavors Specs.

There is actually a more general freezing effect associated with criterial positions in Rizzi's sense: as discussed in Rizzi (2006), once XP moves to a criterial position, it gets frozen there—movement from a criterial to a criterial position is not possible. Bošković (2008b) generalizes this effect formally in terms of feature checking, where a feature-checking movement cannot feed another feature-checking movement. It should, however, be noted that the above discussion most naturally fits with Chomsky's (2008) position regarding movement to criterial positions: Chomsky (2008) suggests that such movement is not formally (i.e. feature-checking) driven, what licenses movement to positions like SpecTopP, SpecFocP…, i.e. what in effect then licenses Spec creation in such cases, is getting a certain interpretation, which fits well with the above discussion regarding "licensing" of specifiers.[38] The more general criterial freezing effect can then be captured as discussed above: if α simply moves through a criterial position X on its way to a higher criterial position, the interpretation associated with it would be lost, since α would not be interpreted in that (X) position (any kind of forced reconstruction would raise the same problem regarding the higher criterial position).[39]

In conclusion, given that intermediate movements involve Spec creation, given MAR, we would expect that there would be no superfluous intermediate movement steps. The *who left* effect represents a rather dramatic confirmation of the ban on superfluous intermediate steps. Given that intermediate movement (to SpecIP here) is banned even in this case, it appears that the null hypothesis should indeed be that intermediate movements take place only through phasal edges (i.e. when they are forced by phases/PIC), which raises a number of interesting issues that were noted in the beginning of this section. The above discussion (i.e. the *who left* effect) has also

[38] In this respect, the reader is referred to the discussion of Turkish in Sener (2010), who argues that Turkish has fully transparent syntax-semantics mapping in that all movements in Turkish are interpretation-driven, i.e. he argues there are no purely formal driven movements in Turkish in the sense that the movement would not be required to get a particular semantic interpretation (a proviso needs to be added here for successive-cyclic movement via phasal edges, discussed in the following section). This is a very interesting and intriguing conclusion, which would make Turkish a "perfect" language in the relevant respect, but which also raises the question why is it that all languages are not like Turkish? To illustrate, while there are clear cases where head-movement has semantic effects (see Dékány 2018, Lambova 2004, Roberts 1991, Gribanova 2017*;* this means that head-movement cannot be simply pushed into PF*;* another reason is that head-movement affects phasal domains for syntactic movement, see Bošković 2013, 2015, den Dikken 2007, Gallego and Uriagereka 2007), V-to-T movement, present in many languages (but not in Turkish), does not ever seem to have it. Another case is Japanese-style long-distance scrambling, which is almost fully devoid of semantic effects (see Saito 1989, 1992; this has actually led to treating Japanese-style scrambling as not involving movement at all, see e.g. Bošković and Takahashi 1998).

[39] Additional assumptions are needed if the movements in question are treated in terms of feature checking (hence driven by a formal requirement) since a feature can be checked on the way to a higher position. Bošković (2008a), who gives such a formally-driven movement analysis, in fact adopts an additional assumption, in terms of Chomsky's (2000, 2001) Activation Condition, where movement of α to a criterial position deactivates α for further movement. As discussed below, another issue arises under the formally-driven movement approach if movement to e.g. SpecTopP is taken to be driven by a requirement for Top to have a Spec. Given MAR, it would be strange to have a formal requirement that would be directly in conflict with MAR in this manner, as discussed in more detail below.

provided evidence against the traditional view of the EPP effect as a formal requirement to fill a particular Spec position, namely SpecIP. I will address this issue more in the following section.

5.2. More on the EPP and MAR

In this section I will address the EPP effect from the perspective of MAR, also generalizing the discussion to the broader issue of what drives movement. (I will not be able to provide a comprehensive discussion of the EPP effect, I will merely point out the relevance of the above discussion for it. Furthermore, since I will not be concerned here with the question of whether the traditional IP should be split (and how it should be treated), I will interchangeably use the terms I(nfl) and T(ense) for the head associated with the EPP effect, depending on what the accounts discussed below assume in this respect. No deeper significance should be attached to this.)

There are two broad approaches to the traditional EPP effect (which is the requirement to fill the SpecIP position) that I will be concerned with here: (a) movement to SpecIP is driven by an inadequacy of the target (I), which requires a Spec; (b) the movement is triggered by a problem in the structure prior to the movement to SpecIP, i.e. a problem which arises when *who* is located in its base-generated position (SpecvP in (26) (see e.g. Bošković 2007, Epstein and Seely 2006, Chomsky 2013). We have seen that SpecIP is never filled in (26); this provides evidence against (a), i.e. against an approach to the traditional EPP effect that would simply require creation of SpecIP for formal reasons—that position is simply not created in (26).[40] On the other hand, such constructions can be captured under approaches along the lines of (b), where the traditional EPP effect is tied to the moving element itself, since such approaches do not per se require SpecIP to be filled. Thus, there are Case-driven approaches to the traditional EPP effect; e.g., in Bošković (2007) the Case requirement is formulated in such a way that a nominative DP simply needs to c-command Infl for its Case to be licensed (i.e. the DP must be a probe here[41]): it undergoes the shortest movement possible to achieve this (due to the more general economy of derivation requirement that every requirement be satisfied through the shortest movement possible). In (26), *who* independently needs to move to SpecCP: since in this position *who* also c-commands Infl, there is then no need to move to SpecIP at all (under this approach to the EPP effect), hence such movement is then not allowed, given the above discussion.[42] As pointed out by Messick (2020), the same actually holds under Chomsky's (2013) labeling approach to the EPP effect, which is abstractly similar to Bošković's: it is something about the base-generated position of the subject that forces its movement—as in Bošković (2007), in Chomsky (2013) there is no requirement to

[40] Chomsky (2008) proposes an account that relies on a version of the traditional EPP along the lines of (a) where there is still no movement from SpecIP to SpecCP in (26). In particular, he proposes a parallel movement account on which *who* moves in parallel from SpecvP to SpecIP and from SpecvP to SpecCP. While this account could handle (35)-(36) if we assume that the latter movement suffices to license Q-float here (see Hiraiwa 2005), it still faces a problem with languages that have different morphology depending on whether or not subject moves to SpecIP in non-questions, where (26) shows the latter morphology. This is surprising since under the parallel movement account there can be movement to SpecIP in the counterpart of (26) in such languages.

[41] For independent evidence for this approach to Case, see Villa-García (2015), Stjepanović (2011), Saito (2016), Aldridge (2018), Dadan (in press), among others. Particularly strong are the arguments given by Villa-García, who gives examples where a DP that is base-generated in the left periphery does not get default case but it gets its case from lower down. Since the relevant functional head does not c-command the DP at any point of the derivation, Villa-García concludes that it must be the case that the DP probes down to be Case-licensed, as in Bošković's (2007) approach.

[42] Since we are dealing here with a matrix question, which in English involves Infl-to-C movement (inversion), Infl would actually move to C; however, such movement would not take place if (26) is embedded under a verb like *ask*, since inversion does not take place in indirect questions.

create SpecIP. The independently required movement of the subject to SpecCP in (26) resolves the issue in question, so that there is no need for movement to SpecIP, i.e. to create SpecIP.[43]

Even putting aside the issue that (26) raises for the traditional EPP approach, which requires filled SpecIP, there is a more serious conceptual issue here, raised by MAR. Given MAR, which disfavors Specs, it would be rather strange to have a condition which requires a Spec, which is exactly what the traditional EPP is.

The issue is in fact more general, it goes beyond the traditional EPP—it concerns the more general question of whether movement is driven by a property of the moving element or by a property of the target.[44] In Bošković (2007), movement in general is in fact never driven by an inadequacy of the target, but by an inadequacy of the moving element.[45] Consider e.g. successive-cyclic movement. The crucial ingredients of Bošković's account of successive-cyclic movement are that there is no feature-checking/agreement in the intermediate positions of successive-cyclic movement (thus, there is no feature-checking between the wh-phrase and the complementizer *that* in the embedded clause of (25); the movement to the embedded SpecCP actually has nothing to do with the complementizer *that*) and that for each step of successive-cyclic movement, in fact *any* movement, it is something about the base position of the movement that drives it,[46] in a sense that something would go wrong in the base position of the movement if it does not take place–there is nothing about the final target, or anything in the higher structure, that motivates it in this sense (thus, if the wh-phrase does not move from the embedded SpecCP in (25), a problem will arise in exactly this part of the structure; nothing would go wrong anywhere else). Note that all these are also the crucial ingredients of Chomsky's (2013) treatment of successive-cyclic wh-movement.

An alternative to the moving element driven movement is a system like Chomsky (2000, 2001). In Chomsky (2000, 2001), X and Y undergo an Agree relation in (38), with X probing Y to value its unvalued F feature. X may or not have the EPP property, which is simply a formal requirement to have a Spec.[47] If it has it, the Agree relation is followed by movement of Y to SpecXP.

(38) X Y
unF val F
(EPP-I need a Spec)

Now, consider the nature of movement driven by a property of the target vs movement driven by a property of the moving element. In the former, movement is driven directly by a formal

[43]Chomsky (2015) proposes a different labeling account of the EPP, where SpecTP always needs to be there (hence the *who left* problem still arises) though there is no explicit requirement to this effect. I will discuss it below.

[44] There are also approaches that allow both, see Lasnik (1995a) and Zyman (2018). See also Nunes (2014, 2019) for an approach that combines Bošković (2007) (where movement is moving-element driven) and Chomsky (2000) (where movement is target-driven) in a way that would still allow us to maintain the conclusions reached below; in particular, Nunes argues that in some cases the property that drives movement (an uninterpretable feature (uK) for Bošković) originates on the phase head but is passed on to the moving element, so that it is still a uK of the moving element that drives the movement.

[45]Bošković (2007, 2011b) discusses cases which are argued to provide support for the base rather than the target driven movement, like quantifier raising. (There is nothing about the target of QR that would require it, i.e. nothing would go wrong with the target of QR if QR does not take place; it is the moving element that needs it.) Another argument to this effect regarding obligatory wh-fronting is given later in this section.

[46] The base position here does not refer to the base-generated position of the moving element, but the tail of any movement step.

[47]The requirement is more general than the traditional EPP—it is applicable to all heads, not just Infl. (It is basically the counterpart of the strength property of Chomsky 1993.)

requirement to create a Spec. This is not the case with the latter: it is not the case that a moving element has a direct requirement to be a Spec. True, satisfying the relevant requirement will force movement, which will end up creating a Spec—but this is so only indirectly, there is no direct requirement to create a Spec.

For the sake of concreteness, consider in this respect successive-cyclic movement in (39), which for ease of exposition shows only one step of successive-cyclic movement (t'_i). Under a moving element driven approach like Bošković (2007), *which book* moves to the edge of the embedded clause to escape being sent to spell-out, movement is not driven by a property of the target head, *that*. On the other hand, consider a purely target-driven approach like Chomsky (2000, 2001): there, *that* is optionally given the property I-need-a-Spec to drive movement to the Spec of *that* (with the further proviso that *that* can be given the I-need-a-Spec property only when this is needed to make successive-cyclic movement possible, a clear instance of look ahead).

(39) Which book$_i$ do you think [t'_i that John bought t_i]?

It should be obvious from the above that the moving element driven system conforms better with the spirit of MAR than the target-driven system, which relies on a requirement to have a Spec, in a direct conflict with MAR (if a head which takes a complement, and the relevant head always does in the BPS system (see section 4)), has an EPP requirement, the requirement directly forces merger with a phrase, i.e. a phrase-phrase merger). In other words, it would be strange to have a formal requirement that would be directly in conflict with MAR in this manner (to put it more bluntly, to require specifiers, in fact all over the place, as in Chomsky's 2000, 2001 target-driven system, would be rather strange in a system which really dislikes Specs).

In this respect, it is worth noting here Chomsky's (2015) approach to the traditional EPP, briefly noted in fn 43, where the traditional EPP effect is tied to an inadequacy of the target but is stated differently, without an explicit requirement to take a Spec. The account actually ties the traditional EPP effect to the head-complement relation: In Chomsky's (2013, 2015) labeling system, when a head and a phrase merge the head projects, labeling the resulting object. However, Chomsky (2015) suggests that T is too weak to label itself (this is a departure from Chomsky 2013), this is why another merger with the object that is created by the T-complement merger is needed. In this account, there is actually no requirement to have a Spec (i.e. for T to have a Spec). The movement in question in fact takes place for a reason related to the head-complement merger, because something goes wrong with that merger. In other words, we appear to have here target-driven movement that is dissociated from a direct Spec requirement. However, it turns out that even this approach is actually in a rather direct conflict with the spirit of MAR. What MAR actually disprefers is a merger of two phrases. Consider now the relevant structure with respect to T. At the relevant point of the derivation, T already has a complement, which means that we have a phrase. Similarly to Chomsky's (2000, 2001) target-driven, I-need-a-Spec approach to movement in general, what we then have here in Chomsky's (2015) target-requirement approach to the traditional EPP, where T does not explicitly require a Spec, is a phrase which at this point of the derivation directly requires another merger—in other words, we have a direct requirement for a phrase-phrase merger.

The upshot of the above discussion is that target-driven approaches to movement generally rely on requirements that are in a direct conflict with MAR. This is not the case with moving-element driven approaches (or approaches that do not require a formal reason for movement). There, there is either no conflict, or only an indirect conflict, hence these approaches conform better with the spirit of MAR. The traditional EPP requirement to have SpecIP is in most direct conflict with MAR. In fact, we have seen above a rather serious empirical problem with the traditional EPP, a context where SpecIP is quite clearly not there, which we have suggested in fact arises due to MAR-related reasons. This is not to say that EPP effects do not exist at all—the point

of the above discussion is that an approach that deduces EPP effects in a way that avoids a direct conflict with MAR would be preferable both conceptually (because of MAR) and empirically (to give us a shot at capturing the *who left* effect, i.e. the lack of SpecIP in such constructions).[48]

Before closing this section, recall that, as briefly noted in section 3, languages with obligatory wh-fronting, i.e. languages that must move at least one wh-phrase, differ regarding the exact landing site of wh-fronting (see e.g. Aldridge 2010, 2018, Bošković 2002a, Horvath 1995, Roberts 2004 (within split CP), Tuller 1992, Yuan 2017). It seems that this variation in the landing site of wh-fronting is easier to accommodate in an approach where the driving force of wh-fronting is in the wh-phrases themselves, than in an approach where the driving force is in the target head. In fact, there are languages that appear to simply require all wh-phrases to move where they quite clearly do not all move to the same position even within a single language (see Bošković 2002a and references therein), which seems to indicate that they are uninterpretable in situ (and crucially not simply interpretable in a single unique position), which fits better with moving-element rather than target-driven systems (see also Watanabe 2002 for a case where the uK of the moving wh-phrase (which drives wh-movement in the moving-element-driven system), is morphologically realized, with the loss of this morphological realization leading to the loss of wh-fronting).

At any rate, in addition to having consequences for the broader issues regarding the driving force of movement (and EPP effects), the discussion section 5 has unified the *who left* effect with other phenomena and mechanisms that were previously unified from the MAR perspective.

6. Conclusion

6.1. MAR and its consequences

Kayne's (1994) seminal work has established the importance of asymmetric relations in the domain of word order. This paper has expanded the domain of asymmetricity with a number of phenomena that are independent of word order, making a case for a Maximize Asymmetric Relations preference (MAR) as a general property of the language faculty by showing that a number of phenomena, which are independent of word order, can be brought together under this perspective (and thus unified with Kayne's original word-order related concern, i.e. the LCA). These include the diachronic loss of specifiers, the avoidance of specifiers in language acquisition, the Phase Impenetrability Condition, the rarity of multiple Spec constructions (cf. e.g. the rarity of the multiple wh-fronting strategy), the no-Spec-without-complement aspect of Bare Phrase Structure (which, under asymetricity, is unified with the notion of the head of a phrase), and the *who left* effect (where movement to SpecCP cannot proceed via SpecIP). What is behind all this is the Maximize Asymmetric Relations (MAR) preference, one consequence of which is that it favors complements over specifiers, since specifiers enter the structure through a symmetric phrase-phrase merger, which is in conflict with MAR. This aspect of MAR provides a unified perspective on superficially very different mechanisms. Thus, the diachronic loss of specifiers, noted by Dadan (2019), the LCA, and the Phase Impenetrability Condition are all different strategies for dealing with a symmetric merger situation: with the first one, one of the relevant elements is lost, with the second one, movement of one of the elements is forced, and the third one changes the status of one of the relevant elements by making part of the structure inaccessible. They thus all resolve symmetric merger situations. The PIC can in fact be looked at as a derivational (hence synchronic) manifestation of the diachronic (and acquisitional) loss of specifiers: since the PIC changes the status of a specifier derivationally, both the diachronic loss of phrasal movement and the PIC

[48] Needless to say, a number of issues were left open above; the goal of the discussion was not to provide a comprehensive account of traditional EPP effects (or comprehensively compare existing accounts) but simply to note a consequence of MAR in this respect and point out some of the ingredients that the eventual account should have (there is really no existing account that captures everything related to EPP effects).

involve a loss of Specs.[49] Furthemore, we have seen that bringing the PIC into the diachronic loss of specifiers makes a prediction that non-phasal specifiers would be more likely to get lost than phasal specifiers.

MAR also has a number of theoretical consequences, in that it favors certain mechanisms and theoretical concepts over their alternatives. Thus, MAR has relevance for the more general issue of whether movement is target- or moving-element driven. MAR favors the latter approaches (or approaches where movement is not formally driven) over the former approaches, which are generally based on requirements that are in a direct conflict with MAR (this in fact holds for the traditional EPP requirement). FL apparently really does not like Specs. Given this, it would be strange to have a pervasive requirement (in fact all over the place) to take a Spec, as in Chomsky's (2000, 2001) target-driven system.

MAR also has consequences for structure building. The notion of the head of a phrase (or any merger situation), which is inherently asymmetric, rather directly reflects the spirit of MAR. MAR also favors complements over specifiers because, in contrast to the latter, the former enter the structure through an asymmetric (head-phrase) merger. As a result, MAR favors BPS structure building over GB structure building (for the latter reason) as well as over the labeling framework (for the former reason). MAR also favors asymmetric approaches to Case-licensing (e.g. Case-valuation over Case-checking). Regarding phases, it favors one particular approach to the Phase Impenetrability Condition as well as approaches to phases that maximize phasal points.

6.2. Broader issues

In this section, I will examine the relevance of the above discussion for more general issues within a broader formalism vs functionalism setting.

As noted above, the reason for the existence of specifiers is the same as the reason for the existence of movement (if it does not involve adjunction movement always creates a specifier)—it concerns functional, or more broadly, non-syntactic considerations—to be able to express notions that go beyond the basic argument structure involving a simple predicate with a single internal argument as well as more complex semantic notions involving issues like scopal ambiguities, pragmatic notions like topic/focus interpretation, specificity… (see also fn 16). In other words, they are needed to express various semantic and pragmatic relations. A question that arises now is to which extent have these notions led to the development of formal requirements which then drive movement? In other words, do the functional reasons in question directly motivate movement, with movement being directly interpretation driven, or are there formal requirements that serve as intermediaries, leaving syntax itself free of semantic considerations of the kind discussed above, with essentially one-way traffic in this respect? To make the question more concrete: when α moves to e.g. SpecTopP and receives topic interpretation, does it move there in order to receive such interpretation or there is a formal reason behind this movement, with α moving to satisfy this formal reason, as a result of which it ends up being interpreted as a topic.[50]

The question may seem innocent in this particular case (there does not seem to be much of a difference between the two positions), but this isn't always the case. Consider e.g. the different behavior of Bulgarian and Japanese in multiple questions, where in Bulgarian all wh-phrases move to the interrogative SpecCP while in Japanese they all remain in situ. There are various ways of implementing this difference. However, without appealing to formal properties that would cause the difference, i.e. on the direct syntax-interpretation mapping approach, we would be led to the conclusion that Bulgarian and Japanese wh-phrases are subject to different interpretation, which

[49] In this respect, note that successive-cyclic movement in a sense also involves a derivational loss of a Spec, due to its moving away aspect.

[50] There is a cartology vs mapping debate regarding discourse notions like topic and focus within formalist approaches (see e.g. Rizzi 1997 and Neeleman and van de Koot 2008; see also Lacerda in prep for a comparison). I have stated the question under consideration in terms of the former for expository reasons.

means they differ semantically (the gist here would be that Bulgarian wh-phrases must function as operators binding a variable; this would not be the case in Japanese, where wh-phrases are interpretable in situ[51]). This also means that on the direct syntax-interpretation mapping approach, Bulgarian and Japanese questions cannot have the same LF.[52]

The two cases just noted can thus be handled through either directly interpretation-driven movement[53] or formal requirements as intermediaries (which would be an indirect syntax-semantics mapping approach), though with non-trivial differences in the semantics of questions/wh-phrases. There is, however, a case where the two can be teased apart, where something formal must be involved: successive-cyclic movement via phasal edges.

Consider again the step of movement to the embedded clause SpecCP in Bulgarian (40).[54]

(40) Kakvo$_i$ misliš [$_{CP}$ t'$_i$ če [$_{IP}$ prodava t$_i$]]]?
what thinks$_{2sg}$ that sells$_{3sg}$
'What do you think he sells?'

The wh-phrase in (40) is not interpreted in the embedded clause SpecCP (*that* is a declarative, not an interrogative complementizer; the embedded clause being a declarative the wh-phrase is really not interpretable in the embedded clause SpecCP). This step of movement then has to be taking place for formal reasons. Phases/PIC provide the formal reason in question: as discussed above, α cannot move out of a phase (and CP is a phase) unless it moves through the phasal edge. In fact, under the target-driven approach (see section 5.2), this movement is completely disconnected from interpretation: as noted above, in Chomsky's (2000) target-driven movement approach, *that* is optionally given the property I-need-a-Spec to drive movement to the Spec of *that* (with the further proviso that *that* can be given the I-need-a-Spec property only when this is needed to make successive-cyclic movement possible, an instance of look ahead). On the other hand, under Bošković's (2007) moving-element driven approach, the wh-phrase has a formal feature that makes it uninterpretable in any position other than interrogative SpecCP. This feature forces movement to the edge of the embedded clause, a phase, so that the wh-phrase escapes being sent to spell-out, which would freeze it in place. Under this approach, movement to the embedded clause SpecCP is more directly related to the interpretation than under the target-driven approach. The declarative complementizer *that*, which drives the movement in the target-driven approach, really has nothing to do with the interpretation of wh-phrases/questions; the wh-phrase, on the other hand, obviously does. True, under both approaches a formal property is involved, but on the movement-driven approach this formal property is a rather direct reflex of interpretation, while on

[51] There are various ways of implementing this, e.g. unselective binding or choice functions (Pesetsky 1987, Nishigauchi 1990, Reinhart 1998, Hagstrom 1998, Cable 2010, see also Shimoyama 2006). English would have to be different from both Bulgarian and Japanese, where a wh-phrase could be interpreted in situ as long as one wh-phrase is interpreted through an operator-variable relation. The notion of absorption in the sense of Higginbotham and May (1981) could be relevant here (see, however, Pesetsky 2000 for a pronunciation-of-a-lower-copy approach where there is no syntactic or semantic difference between Bulgarian and English, the only difference residing in PF).

[52] Positing LF wh-movement in wh-in-situ languages like Japanese (see e.g. Huang 1982), so that Japanese and Bulgarian still end up with the same LF, would not work—we would still need a formal difference between Japanese and Bulgarian that would be responsible for the overt/covert movement difference, which would go against the spirit of the direct syntax-semantics mapping approach.

[53] QR might be the best candidate for such movement, though there is some controversy regarding the existence of QR (see e.g. Hornstein 1995) and there are also attempts to implement it as formally-driven movement (i.e. with a formal requirement as an intermediary, see Beghelli 1995).

[54] I am illustrating the point here with Bulgarian rather than English due to the complication noted in fn 51 (the point can actually also be made with English, though in slightly more roundabout way).

the target-driven approach it really has nothing to do with interpretation (again, the complementizer *that* has nothing to do with question interpretation).

Pursuing one of the tenets of minimalism that as much as possible, and as directly as possible, should follow from bare output conditions, which means the nature of language, formal reasons that have nothing to do with the nature of language should be minimized. This again favors the moving-element-driven approach. It is the very nature of successive-cyclic movement that it lands in a position where the interpretation does not take place, and the head of the phrase where the movement lands has nothing to do with the relevant interpretation. Even in the case of the final landing site, it is a bit of a stretch to tie the reason to the head of the relevant phrase, not the moving element itself—with successive-cyclic movement it is not merely a stretch, it is simply not possible. It seems that the moving element itself must be involved, in a way that what motivates movement is that the moving element itself is not interpretable where it is, hence it undergoes movement.[55] This is in fact the gist of Bošković's moving-element driven approach. Note also that what I have referred to above as the minimalist tenet, tie as much as possible, and as directly as possible, to the very nature of language, is essentially the Occam's razor strategy—simplify to what is unavoidable. So, Occam's razor, which is to say the nature of language, the bare output conditions (see Chomsky 1995 on this notion), favor the moving-element-driven approach. EPP, which drives movement in the target-driven approach, is really a conceptual abomination from this point of view, especially when it comes to successive-cyclic movement, where it is unconnectable, directly or indirectly, to anything having to do with interpretation—those intermediate heads simply have nothing to do with it.

In this respect, it is worth noting that Chomsky (1995) argues for elimination of Agr (i.e. agreement phrases) on the grounds that this is a strictly formal element which does not contribute to the interpretation. The same reasoning should lead to the elimination of the EPP—it is in fact rather difficult to see how one can argue against AgrPs on these grounds while still accepting the concept of an EPP property on e.g. complementizer *that/če* in (39)-(40) (which is as purely formal as Agr phrases).

The above discussion may make both formalists and functionalists feel uncomfortable, the former because of the emphasis on what is essentially a maximize-functional-considerations-strategy (given that the nature of language certainly includes the function of language—as noted above, the strategy in this context really follows from Occam's razor),[56] and the functionalists because of the rather formal/technical nature of the preceding discussion. This isn't unintentional. Neither formalism nor functionalism are something to shy away from. In fact, the discussion here

[55] This is also the general idea behind quantifier raising, which Bošković (2011b) actually treats as a case of moving-element-driven movement (see fn 45).

[56] Chomsky's (1995) stand toward Agr can be interpreted as a minimize-purely-formal-considerations strategy. There is a point to be made here though. By maximizing functional considerations I do not mean incorporating everything, including the general worldview, of functional linguistics. Functionalists rely on much poorer syntax than generativists. The reason for this is methodological, which has to do with their main goal: for them pragmatics and semantics are more basic than syntax; they look at syntax simply as a tool for expressing pragmatic functions and semantic roles—as a result, they generally do not consider syntactic relations and dependencies that go beyond the tool role of syntax (they also generally do not investigate what is not possible, which is the central interest of generativists, since their goal is to capture what is possible, i.e. to express the needed pragmatic and semantic notions). A generativist pursues a different methodology here, which is again connected to their primary goal that gives primacy to syntax: they are interested in examining the full complexity of syntactic relations and discovering syntactic principles that determine well-formed as well as ill-formed sentences—pragmatics and semantics then take the former, assigning them interpretation and pragmatic use. By maximizing functional considerations I do not mean adopting the pragmatics-is-primary/syntax-is-just-a-tool attitude; in fact, nothing in the standard generativist attitude in this respect needs to change.

can be considered a plea for a new formal functionalism, which incorporates insights of both approaches.

Consider in this respect even MAR itself. The question here is whether MAR is simply an irreducible formal property of FL, or whether it follows from something else, i.e. whether it can be traced back to FL-external factors, related to broader cognitive mechanisms. To me the former seems unlikely—in fact, the preference nature of MAR itself may suggest that if, as is often assumed, tendencies and preferences, as opposed to absolute properties, are more likely to have functional/broader cognitive than formal/FL-specific explanations. (There is also a personal methodological aversion to principles—many linguists seem to be delighted when they discover what they consider new principles—I detest such moments; principles are things that we do not understand—reaching a principle stage should be detested on the methodological grounds of the nature of scientific inquiry, which is striving to understand). Moreover, Kayne (2010) notes that both parsing and language production are asymmetric, in that they show a beginning vs end asymmetry. If there is indeed a connection here, we would be dealing with a broader property of language which goes beyond linguistic competence, i.e. knowledge of language, that the generativists are concerned with—the property would extend to performance, i.e. usage of that knowledge. What is also relevant here is that Minimalism explores the possibility that UG, a genetic endowment that helps children acquire language, is an optimal way of satisfying requirements imposed on FL by the external systems that FL interfaces with. From this perspective, asymmetricity can be encoded in UG even if the ultimate source of asymmetricity is the external systems FL interfaces with—asymmetricity would be imposed on UG as the optimal way of satisfying external system requirements. At any rate, if it turns out that we are indeed dealing here with a broader property, this paper can be taken as an illustration of how looking seriously into formal properties of language can help us elucidate those broader mechanisms, a kind of investigation I have in mind with the term formal functionalism.

In fact, to be a minimalist also means to be a functionalist. Not just in theory—no sane minimalist would claim that there is no aspect of language that can be explained by looking at the function of language—but also a practicing functionalist, in the sense that the minimalist should constantly bear in mind potential functional considerations, in fact as the source of what we refer to as FL external properties. There shouldn't be anything controversial in saying that as much as possible should follow from the nature of language, which inevitably brings in functional considerations. But they may end up being incorporated in a rather abstract, rigorous, and formal way, once the formal properties of the computational system of the language which functionalists themselves generally shy away from are brought into the picture.

In fact, the divide between the traditional generativist and the traditional functionalist/ typological camp seems to be bigger in the slogans that are used to characterize the respective camps than in actual practice.[57] For the sake of illustration, one of the tenets of the latter camp, as perceived by the former camp (it was in fact repeated a number of times at the Heidelberg workshop by Martin Haspelmath), is that every language should be described in its own right, which is often perceived by generativists as a there-is-no-universal-grammar attitude.[58] I am not

[57] The term traditional functionalist/typological camp is used to reflect the fact that traditional typology, about to be discussed, is generally associated with functional approaches.

[58] The position goes back to structural linguistics works like Boas (1911), for discussion and references see Haspelmath (2010). The word tenet may be too strong since the view in question is not held by all non-generative typologists (see again the references in Haspelmath 2010), which I am putting aside here. The perception among the generativists that this position is a reflection of a no-UG attitude is actually misguided since typological works investigate less abstract/more easily observable phenomena that traditional descriptive linguistics is concerned with, which Chomskian generativists generally do not consider to be the main source of universals of UG (those concern more abstract properties like phrase structure/structural

sure the practitioners of the latter camp truly believe the slogan in question. After all, the typologists from that camp have made incredibly important contributions to the field in terms of Greenberg-style crosslinguistic generalizations, which the generativists are also increasingly relying on. Reaching such generalizations would not have been possible if they really believed the slogan in question —you really cannot be a typologist if you truly believe it.[59] Haspelmath (2010) actually tries to show that typological work can be done while fully adhering to it regarding the domain of grammatical categories. However, he also gives universal definitions of the relevant categories that are applicable to all languages (and which make typological work possible). He actually claims that they are artificial linguistic constructs and not real (i.e. not part of the grammar of individual languages), but there is really no reason why they should not be considered real. He makes a distinction between a language particular descriptive category (referred to as X below), and its crosslinguistically applicable comparative concept (referred to as Y), which is used in typological generalizations. But there is no real issue here—it may be that in some language X is just like Y, while in another language, where X is not exactly like Y, we simply have a more complex situation where X=Y+Z; there is then Y in that language as well. To illustrate, consider the generalization in (41a) and the definition of the relevant element from Haspelmath (2010:671) in (41b), which is an example of Y.

(41) a. GENERALIZATION: In all languages, markers of future tense are less bound than markers of present tense or past tense, or equally bound, but never more so. (Ultan 1987:91)

b. DEFINITION: A future tense is a grammatical marker associated with the verb that has future time reference as one prominent meaning.

Haspelmath notes crosslinguistic differences regarding future tense, e.g. in Spanish it is also used to express probability (but not habituality), while in Lezgian it is also used to express habituality (but not probability). This indicates that future tenses are not synonymous crosslinguistically, which then requires a distinction between a language particular descriptive category (X from above) and its crosslinguistically applicable comparative concept (Y from above). But notice that what we have here is the more complex X=Y+Z situation, where Y is still always present. Haspelmath gives a similar definition of question words, and notes that in many languages question words have additional usages, e.g. as indefinite pronouns or as relative pronouns, with languages differing in this respect. This is also an instantiation of the X=Y+Z situation. As another illustration, consider the generalization in (42a) and the definition in (42b) (Haspelmath 2010:673).

---

dependencies, islandhood and locality relations…, which traditional typologists generally do not explore (see e.g. Baker 2015)). To illustrate, (i) gives one such generalization from Bošković (2012) (based on earlier work considering only some Slavic languages by Uriagereka 1988 and Corver 1992). (ia) is the relevant statement from Bošković (2012), while (ib) restates it in a more run-of-the-mill implicational universal way. ((ia) is more intuitive since the presence of definite articles (more precisely, the DP projection, which is present only in languages with definite articles) is what blocks (ii), though there are other conditions on the possibility of (ii), like agreement—see Bošković 2012. Note also that when testing (i), tests need to be done to ensure that (ib) in the relevant language does not involve a base-generated topic (something like "as for expensive (things), John likes expensive cars") or NP ellipsis (something like 'as for expensive ~~cars~~, John likes expensive cars", where 'likes' or 'John' would likely be focalized)).

(i) a. Left-branch extraction of adjectives (and adjectival-like elements), as in (ii), may be allowed only in languages without definite articles.

b. If a language allows left-branch extraction of adjectives (and adjectival-like elements), as in (ii), then the language does not have definite articles.

(ii) *Expensive$_i$, John likes [t_i cars]

[59] Most American structuralists, on the other hand, did believe it, as a result of which they generally did not engage in typological work (see e.g. Greenberg 1974, Haspelmath 2010).

(42) a. GENERALIZATION: In all languages with an ergative case, it has at least some overt allomorphs (Dixon 1979)

b. DEFINITION: An ergative case is a morphological marker that has among its functions the coding of the agent of typical transitive clauses, when this is coded differently from the single argument of intransitive clauses.

Haspelmath observes that ergative case from (42b) is not the same in all languages, since in many languages it has additional functions, like instrumental, locative, possessive, general oblique. This is again the X=Y+Z situation. Haspelmath considers what I am referring to as Y as concepts created by linguists for the purpose of formulating crosslinguistic generalizations. True, they are applicable to all languages but they are supposed to be artificial and not real (i.e. not psychologically real and not part of particular language systems). But there is no real reason why those Ys could not be (psychologically) real, and in fact part of UG (I am putting aside the question of what the primitives of UG in the relevant domain are, which is not relevant to the more general point made here). In some languages those abstract categories would map straightforwardly to surface categories, and in some languages that would not be the case: in the latter case we would have the X=Y+Z relationship, with Y still applicable to all languages. This kind of relationship would become more obvious if we accept that the grammar of each language that is studied in its own terms is, as Baker (2015:936) puts it, "abstract to some non-trivial degree" (which is what generativists generally accept, while functionalists/traditional typologists generally don't)—this would result in more Y=X situations and would make the Y-X relationship more transparent). In this respect, it should be noted that generative typology, an emerging and increasingly influential strand of work within the formalist camp that attempts to reach Greenberg-style typological generalizations and then provide generative explanations for them, also assumes what I have referred to as X and Y above (so there is really no disagreement here), the only difference is that Y, which Haspelmath considers a linguist's construct, is treated as real and in fact part of UG.[60]

At any rate, we can define the relevant concepts differently for each language, or in a way that makes them universally applicable (which would be more abstract but abstract need not to mean not real and artificial), a necessary prerequisite for typological work. Occam's razor, as a general scientific principle, is again relevant here. As Haspelmath (2010) notes, a number of non-generative typologists explicitly refute the view that 'every language must be described in its own terms' (e.g. Dahl 1985, Bybee and Dahl 1989, Lehmann 1989). Thus, Lehmann (1989:142) says: "Describe your language in such a way that the maxim of your description could serve, at the same time, as the principle of general comparative grammar—and thus, the maxim of description of any other language." This will result in fewer mechanisms, categories, concepts…, which simply reflects Occam's razor as a general scientific principle (in addition to being a prerequisite for typological work). Haspelmath's position is essentially a result of accepting a certain level of abstractness in doing typological work but not in doing analyses of individual languages, which basically leads to separating the two into different fields ("the analysis of particular languages and the comparison of languages are thus independent of each other as theoretical enterprises", Haspelmath 2010:682). Allowing the same level of abstractness for both, which would also be in the spirit of Occam's razor, would, however, dissolve this distinction.

There is, however, an aspect of the describe-every-language-in-its-own-right slogan which is generally not explicitly acknowledged and which should not be dismissed by the generativists

[60]Thus, the point made with (41)-(42) can also be illustrated with Bošković's (2012) generalizations regarding definite articles, one of which is given in (i) in fn 58, and Bošković's definition of definite article (see Bošković 2016c), which superficially shows similar variation across languages as future tense and ergative case.

as a denial of UG but taken as a methodological warning to be careful before jumping to the conclusion that something is in UG—it's a warning not to follow up a detailed investigation of a single language with the proclamation that it is all UG. This tendency is to still there to some extent among the generativists—I don't have in mind here abstract properties like e.g. c-command and domination where a single language can be used as an illustration for ease of exposition—I am not aware of any language where c-command/domination (which basically means structure) do not hold, but investigations of minute details of structure of a single language which is followed up by a claim that all languages are like that without checking other languages (or simply by forcing other languages into the mold set up by that detailed investigation of a single language without looking at, or ignoring, what does not fit).[61] From this perspective, I have to admit that I understand and am sympathetic to the functionalist/typology mantra look-at-languages-in-their-own-right; it should not be dismissed off hand, and we should not jump to the conclusion that everything we see in one particular language is universal grammar at work.[62]

It should, however, be noted in this connection that there is nothing wrong with Chomsky's position that one can learn a great deal about UG by studying a single language—this is true of what would be considered principles of UG (see e.g. Epstein 1999, where one language is used as an illustration to examine c-command); examining parametric variation does require going beyond a single language (though focusing on one language may run the risk of improperly treating a parametric point as a principle of UG). At any rate, for additional discussion to the effect that differences between the traditional generativist and the traditional functionalist/typological camp are bigger in the slogans associated with these approaches than in actual practice, the reader is referred to Bošković (in prep). One point made there concerns Greenberg-style generalizations, which at the right level of abstractness that also dissolves exceptions to them do reflect UG at work; from this perspective the practitioners of the traditional functionalist/typological camp have contributed a great deal to our understanding of UG (although they may deny it for reasons discussed in Bošković in prep).

References

Adams, Marianne. 1987. From Old French to the theory of pro-drop. *Natural Language and Linguistic Theory* 5: 1-32.

Agbayani, Brian. 2000. Wh-subjects in English and the vacuous movement hypothesis. *Linguistic Inquiry* 31:703-713.

Aldridge, Edith. 2009. Short *wh*-movement in Old Japanese. *Japanese/Korean Linguistics* 17:549-563.

Aldridge, Edith. 2010. Clause internal *wh*-movement in Archaic Chinese. *Journal of East Asian Linguistics* 19:1-36.

Aldridge, Edith. 2011. Neg-to-Q: Historical development of one clause-final particle in Chinese. *The Linguistic Review* 28.4: 411-447.

Aldridge, Edith. 2018. C-T inheritance and the left periphery in Old Japanese. *Glossa* 3: 1-22.

[61] While there has been a strong emphasis on investigations of understudied languages in the generativist camp in recent years, Germanic and Romance still hold a privileged place when it comes to claims about UG, with understudied languages often being used to confirm those claims (which means they are often used as mold fitters).

[62]As discussed in Bošković (in prep), the look-at-languages-in-their-own-right mantra also seems to be a reaction to what functionalists/traditional typologists seem to take to be an assumption held by everyone in the generative camp, namely, that there is a universal sentence structure holding for all languages (the assumption being that if we were to look at languages in their own right it would become clear that the position is untenable). There is, however, a misconception here regarding the relevant state of affairs within the generative camp. While it is true that this claim is often made by the generativists, it is often given up in actual practice (see Bošković in prep).

Alexiadou, Artemis. 2002. Word order patterns in Greek nominals: Aspects of diachronic change. *ZAS Papers in Linguistics* 27, 91-107.
Alexiadou, Artemis and Elena Anagnostopoulou. 1998. Parameterizing AGR: Word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16: 491-539.
Axel-Tober, Katrin. 2017. The development of the declarative complementizers in German. *Language* 93: e39-e65.
Baker, Mark C. 2015. Formal generative typology. In *The Oxford handbook of linguistic analysis 2nd edition*, ed. by Bernd Heine and Heiko Narrog, 925-951. Oxford: Oxford University Press.
Baltin, Mark. 2007. Deletion versus pro-forms: A false dichotomy? Ms., New York University.
Barbosa, Pilar. 1995. Null subjects. Doctoral dissertation, MIT, Cambridge, MA.
Beghelli, Filippo. 1995. The phrase structure of quantifier scope. Doctoral dissertation, University of California, Los Angeles.
Biberauer, Theresa and Ian Roberts. 2006. Loss of residual "head final" orders and remnant fronting in Late Middle English: Causes and consequences. In *Comparative studies in Germanic Syntax: From Afrikaans to Zürich German*, ed. by Jutta M. Hartmann and László Molnárfi, 263-297. Amsterdam: John Benjamins.
Boas, Franz. 1911. Introduction. Handbook of American Indian languages, vol. 1. (Bureau of American Ethnology, bulletin 40.) Washington, DC: Government Printing Office.
Bobaljik, Jonathan. 1995. Morphosyntax: The syntax of verbal inflection. Doctoral dissertation, MIT, Cambridge, MA.
Boeckx, Cedric, and Kleanthes K. Grohmann. 2007. Putting phases in perspective. *Syntax* 10: 204-222.
Borer, Hagit. 1986. I-Subjects. *Linguistic Inquiry* 17: 375–416.
Bošković, Željko. 1997. *The syntax of nonfinite complementation: An economy approach*. Cambridge, MA: MIT Press.
Bošković, Željko. 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena.* Amsterdam: Elsevier Science.
Bošković, Željko. 2002a. On multiple wh-fronting. *Linguistic Inquiry* 33: 351-383.
Bošković, Željko. 2002b. A-movement and the EPP. *Syntax* 5: 167-218.
Bošković, Željko. 2006. Case checking vs Case assignment and the Case of adverbial NPs. *Linguistic Inquiry* 37, 522-533.
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 the operator freezing effect. *Natural Language and Linguistic Theory* 26, 249-287.
Bošković, Željko. 2008b. 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 Hegedűs, and Henk van Riemsdijk, 195–233. Amsterdam: Elsevier.
Bošković, Željko. 2011a. On Unvalued Uninterpretable Features. In *Proceedings of 39th Conference of the North-Eastern Linguistic Society (NELS 39)*, ed. by Suzi Lima, Kevin Mullin, and Brian Smith. 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 Günther Grewendorf and Thomas Ede Zimmermann, 179-242. Berlin: Mouton de Gruyter.
Bošković, Željko. 2013. Traces do not head islands: What can PF deletion rescue? In *Deep insights, broad perspectives: Essays in honor of Mamoru Saito*, ed. by Yoichi Miyamoto,

Daiko Takahashi, Hideki Maki, Masao Ochi, Koji Sugisaki, and Asako Uchibori, 56−93. Tokyo: Kaitakusha.
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. What is sent to spell-out is phases, not phasal complements. *Linguistica* 56: 25–56.
Bošković, Željko. 2016c. On second position clitics crosslinguistically. In *Formal Studies in Slovenian Syntax. In honor of Janez Orešnik,* ed. by Franc Lanko Marušič and Rok Žaucer, 23-44. Amsterdam: John Benjamins.
Bošković, Željko. in press. 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. Berlin: Mouton De Gruyter.
Bošković, Željko. in prep. Formalism and not vs Functionalism. Ms., University of Connecticut.
Bošković, Željko, and Daiko Takahashi. 1998. Scrambling and last resort. *Linguistic Inquiry* 29: 347-366.
Bresnan, Joan. 1971. Contraction and the transformational cycle. Ms., MIT.
Brown, H. Paul, Brian D. Joseph, and Rex E. Wallace. 2009. Questions and answers. In *New perspectives on historical Latin syntax volume 1: Syntax of the sentence*, ed. by Philip Baldi and Pierluigi Cuzzolin, 489-530. Berlin: Mouton de Gruyter.
Bybee, Joan L. and Östen Dahl. 1989. The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13:51-103.
Cable, Seth. 2010. *The grammar of Q: Q-particles, wh-movement and pied-piping*. Oxford: Oxford University Press.
Carstens, Vicki, Norbert Hornstein, and Daniel Seely. 2016. *The Linguistic Review* 33, 67-86.
Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20,* ed. by Kenneth Hale and Samuel J. Keyser, 1-52. Cambridge, MA: MIT Press.
Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: 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, MA: MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.
Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. by Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 133–166. Cambridge, MA: MIT Press.
Chomsky, Noam. 2013. Problems of projection. *Lingua* 130: 33–49.
Chomsky, Noam. 2015. Problems of projection: Extensions. In *Structures, strategies and beyond: Studies in honour of Adriana Belletti*, ed. by Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, 3–16. Amsterdam: John Benjamins.
Corver, Norbert. 1992. On deriving left branch extraction asymmetries: a case study in parametric syntax. In *Proceedings of 22nd Conference of the North-Eastern Linguistic Society (NELS 22)*, ed. by Kimberly Broderick, 67-84. Amherst: GLSA, University of Massachusetts.
Coveney, Aidan. 1995. The use of the QU- final interrogative structure in spoken French. *Journal*

of French Language Studies 5: 143-171.
Dadan, Marcin. 2019. Head labeling preference and language change. Doctoral dissertation, University of Connecticut, Storrs.
Dadan, Marcin. in press. Case sharing: Evidence from Slavic and beyond. In *Formal Approaches to Slavic Linguistics 26,* ed. by Tania Ionin and Jonathan MacDonald. Ann Arbor, MI: Michigan Slavic Publications.
Dahl, Östen. 1985. *Tense and aspect systems*. Oxford: Blackwell.
Danckaert, Lieven. 2012. *Latin embedded clauses: The left periphery.*Amsterdam: John Benjamins
Dékány, Éva. 2018. Approaches to head movement: A critical assessment. *Glossa* 3(1): 65. 1–43.
den Dikken, Marcel. 2007. Phase extension: Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33(1). 1−41.
den Dikken, Marcel and James Griffiths. 2018. Ellipsis and spec-head agreement. Ms., Hungarian Academy of the Sciences and Lieden.
Diesing, Molly. 1996. Semantic variables and object shift. In *Studies in comparative Germanic syntax.* ed. by Höskuldur Thráinsson, Samuel D. Epstein, and Steve Peter, 66-84. Dordrecht: Kluwer.
Di Sciullo, Anna Maria. 2011. A biolinguistic approach to variation. In *The biolinguistic entreprise: New perspectives on the evolution and nature of the human language faculty*, ed. by Anna Maria Di Sciullo and Cedric Boeckx, 305-328. Oxford: Oxford University Press.
Di Sciullo, Anna Maria. 2015. On the domain specificity of the human language faculty and the effects of principles of computational efficiency: contrasting language and mathematics. *Revista Linguística* 11(1), 57-70.
Di Sciullo, Anna Maria. 2017. Asymmetry and the Language Faculty. *Revista Linguistica. Gramatica Generativa: celebrando os 60 anos de Syntactic Structures (1957-2017),* 13: 88-107.
Di Sciullo, Anna Maria, Marco Nicolis, and Stanca Somesfalean. 2020. Comitative P. In *Variation in P: Comparative approaches to adpositional phrases*, ed. by Jacopo Garzonio and Silvia Ross. Oxford: Oxford University Press.
Duguine, Maia and Aritz Irurtzun. 2014. From obligatory WH-movement to optional WH-in-situ in Labourdin Basque. *Language* 90: 1-30.
Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68: 81-138.
Erlewine, Michael Yoshitaka, 2016. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language and Linguistic Theory* 34: 429–479.
Epstein, Samuel David. 1999. Un-principled syntax: The derivation of syntactic relations. In *Working Minimalism*, ed. by Samuel Epstein and Norbert Hornstein, 317-345. Cambridge, MA: MIT Press.
Epstein, Samuel D. 2007. On i(internalist)-functional explanation in Minimalism. *Linguistic Analysis* 33:20–53.
Epstein, Samuel D. 2009. The unification of theta relations: How TRANSFER renders SpecvP a theta marked complement. Ms., University of Michigan.
Epstein, Samuel D. and T. Daniel Seely. 2002. Rule Applications as Cycles in a Level-Free Syntax. In *Derivation and Explanation in the Minimalist Program,* ed. by Samuel D. Epstein and T. Daniel Seely, 65-89. Oxford: Blackwell.
Epstein, Samuel D. and T. Daniel Seely. 2006. *Derivations in Minimalism: Exploring the elimination of A-chains and the EPP*. Cambridge: Cambridge University Press.
Figueiredo Silva, Maria Cristina and Elaine Grolla. 2016. Some syntactic and pragmatic aspects of WH-in-situ in Brazilian Portuguese. In *The morphosyntax of Portuguese and Spanish in Latin America*, ed. by Mary A. Kato and Francisco Ordonez, 259-285. Oxford: Oxford University Press.
Fortson IV, Benjamin W. 2004, *Indo-European Language and Culture: An Introduction.*

Blackwell Publishing.
Gallego, Ángel and Juan Uriagereka. 2007. Conditions on sub-extraction. In *Coreference, modality, and focus*, ed. by Luis Eguren and Olga Fernández-Soriano, 45−70. Amsterdam: John Benjamins.
van Gelderen, Elly. 2004. *Grammaticalization as Economy.* Amsterdam: John Benjamins.
van Gelderen, Elly. 2009. Renewal in the left periphery: Economy and the complementizer layer. *Transactions of the Philological Society* 107.2 (2009): 131-195.
van Gelderen, Elly. 2011. *The Linguistic Cycle.* Oxford: OUP.
van Gelderen, Elly. 2015. Problems of Projection: The role of language change in labeling paradoxes. *Studia Linguistica*: 1-15.
Ginzburg, Jonathan and Ivan A. Sag. 2000. *Interrogative Investigations*. Stanford: CSLI Publications.
Goto, Nobu. 2013. Labeling and scrambling in Japanese. *Tohoku: Essays and Studies in English Language and Literature* 46. 39–73.
Greenberg, Joseph H. 1974. *Language typology: A historical and analytic overview*. The Hague: Mouton.
Gribanova, Vera. 2017. Head movement and ellipsis in the expression of Russian polarity focus. *Natural Language and Linguistic Theory* 35(4):1079–1121.
Haddican, Bill. 2007. The structural deficiency of verbal pro-forms. *Linguistic Inquiry* 38:539–547.
Haegeman, Liliane. 2000. Negative preposing, negative inversion, and the split CP. In *Negation and polarity: syntactic and semantic perspectives*, ed. by Laurence Horn and Yasuhiko Kato, 21-61. Oxford: Oxford University Press.
Hagstrom, Paul. 1998. *Decomposing questions*. Doctoral dissertation. MIT, Cambridge, MA.
Hale, Mark. 1987. *Studies in the Comparative Syntax of the Oldest Indo-Iranian Languages.* Doctoral dissertation, Harvard University.
Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from Building 20: Essays in linguistics in honor of Sylvian Bromberger*, ed. by Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge, MA: MIT Press.
Harris, Alice C., and Lyle Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.
Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3): 663-687.
Heim, Irene, and Angelika Kratzer. 1998. *Semantics in Generative Grammar.* MA: Blackwell.
Higginbotham, James and Robert May. 1981. Questions, quantifiers and crossing. *The Linguistic Review* 1: 41-79.
Hornstein Norbert. 1995. *Logical Form: From GB to Minimalism*. Oxford: Blackwell.
Horvath, Julia. 1995. Structural Focus, structural Case, and the notion of feature-assignment. In *Discourse configurational languages*, ed. by Katalin É Kiss, 28-64. Oxford: Oxford University Press.
Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT, Cambridge, MA.
Huddleston, Rodney and Pullum, Geoffrey K. 2002. *The Cambridge Grammar of the English Language*, Cambridge: Cambridge University Press.
Kaisse, Ellen M. 1983. The syntax of auxiliary reduction in English. *Language* 59: 93-122.
Kang, Jungmin. 2014. On the Absence of TP and its Consequences: Evidence from Korean. Doctoral dissertation, University of Connecticut, Storrs.
Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
Kayne, Richard. 2010. Why are there no directionality parameters? In *Proceedings of the 28th West Coast Conference on Formal Linguistics*, ed. by Mary Byram Washburn, Katherine

McKinney-Bock, Erika Varis, Ann Sawyer, and Barbara Tomaszewicz, 1-23. Somerville: Cascadilla Press.
Kiparsky, Paul. 1996. The Shift to Head-Initial VP in Germanic. In *Studies in comparative Germanic syntax*, ed. by H. Thráinsson, S. Epstein, and S. Peter, 140-179. Dordrecht: Kluwer.
Koizumi, Masatoshi. 1994. Layered specifiers. In *Proceedings of the North East Linguistic Society 24*, ed. by Mercè Gonzàlez, 255-269. GLSA, University of Massachusetts, Amherst.
Kuroda, Shige-Yuki. 2007. On the syntax of Old Japanese. In *Current issues in the history and structure of Japanese*, ed. by Bjarke Frellesvig, J.C. Smith, and Masayoshi Shibatani, 263-318. Tokyo: Kurosio.
Lacerda, Renato. in preparation. Middle-field syntax and information structure in Brazilian Portuguese. Doctoral dissertation, University of Connecticut.
Lambova, Mariana. 2004. On triggers of movement and effects at the interfaces. In *Triggers*, ed. by Anne Breibarth and Henk C. van Reimsdijk, 231–258. Berlin: de Gruyter.
Lasnik, Howard. 1995a. Case and expletives revisited: On Greed and other human failings. *Linguistic Inquiry* 26:615-633.
Lasnik, Howard. 1995b. Verbal morphology: *Syntactic Structures* meets the minimalist program. In *Evolution and Revolution in Linguistic Theory: Essays in Honor of Carlos Otero*, ed. by Héctor Campos & Paula Kempchinsky, 251-275. Georgetown: Georgetown University Press.
Ledgeway, Adam. 2012. *From Latin to Romance. Morphosyntactic Typology and Change.* New York: Oxford University Press.
Lehmann, Christian. 1989. Language description and general comparative grammar. In *Reference grammars and modern linguistic theory*, ed. by Gottfried Graustein and Gerhard Leitner, 133-62. Tübingen: Niemeyer.
Lightfoot, David, 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.
Lobeck, Anne. 1990. Functional Heads as Proper Governors. In *Proceedings of the North East Linguistic Society* 20, 348-362.
Lopes Rossi, Maria Aparecida. 1996. *As Orações Interrogativas-q no Português de Brasil: Um Estudo Diacrônico*. Doctoral dissertation, Universidade Estadual de Campinos.
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): 13. 1-12.
Meyer, Roland. 2017. The C System and Complement Clauses in the History of Slavic Languages. *Language* 93(2): e97-e113.
Mizuguchi, Manabu. 2014. Phases, labeling, and wh-movement of the subject. Presented at the 32nd Conference of the English Linguistic Society of Japan, Gakushuin University, Tokyo.
Moro, Andrea. 1997. *The raising of predicates. Predicative noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press.
Moro, Andrea. 2000. *Dynamic Antisymmetry*. Cambridge, MA: MIT Press.
Müller, Gereon. 2010. On deriving CED effects from the PIC. *Linguistic Inquiry* 41: 35–82.
Narita, Hiroki. 2011. Phasing in Full Interpretation. Doctoral dissertation, Harvard University.
Narita, Hiroki. 2012. Phase cycles in service of projection-free syntax. In *Phases: Developing the framework*, ed. by Ángel J. Gallego, 125–172. Berlin: Mouton de Gruyter.
Neeleman, Ad, and Hans van de Koot. 2008. Dutch scrambling and the nature of discourse templates. *The Journal of Comparative Germanic Linguistics* 11: 137–189.
Nishigauchi, Taisuke. 1990. *Quantification in the theory of grammar*. Dordrecht: Kluwer.
Nunes, Jairo. 2014. Adjunct control and edge features. In *Minimalism and beyond: Radicalizing the interfaces*, ed. by Peter Kosta, Steven L. Franks, Teodora Radeva-Bork and Lilia Schürcks, 79-108. Amsterdam: John Benjamins.

Nunes, Jairo. 2019. Edge features and phase head allomorphy. Paper presented at GLOW 42, the University of Oslo, May 7-11, 2019.
Ogawa, Kunihiko. 1976. Where diachronic and synchronic rules meet: A Case study from Japanese interrogatives and Kakari-Musubi. *Papers in Japanese Linguistics* 5, 193-242.
Oiry, Magda. 2011. A case of true optionality: Wh-in-situ patterns like Long Movement in French. *Linguistic Analysis* 37: 115-142.
Ott, Dennis. 2012. *Local Instability: Split Topicalization and Quantifier Float in German*. Berlin: De Gruyter.
Pesetsky, David. 1987. *Wh*-in situ: Movement and unselective binding. In *The representation of (in)definiteness*, ed. by Eric J. Reuland and Alice G. B. ter Meulen, 98-129. Cambridge, MA: MIT Press.
Pesetsky, David. 2000. *Phrasal movement and its kin*. Cambridge, MA: MIT Press.
Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 355-426. Cambridge, MA: MIT Press.
Reglero, Lara. 2007. Wh-in-situ interrogatives in Spanish. *Probus* 19: 267-297.
Reglero, Lara and Emma Ticio. 2013. A unified analysis of wh-in-situ in Spanish. *The Linguistic Review* 30: 1-47.
Reinhart, Tanya. 1998. *Wh*-in-situ in the framework of the minimalist program. *Natural Language Semantics* 6:29-56.
Richards, Norvin. 2001. *Movement in language: Interactions and architectures*. Oxford: Oxford University Press.
Richards, Norvin. 2010. *Uttering trees*. Cambridge, Mass.: MIT Press.
Richards, Norvin. 2016. *Contiguity theory*. 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: Kluwer.
Rizzi, Luigi. 2001. On the position "int(errogative)" in the left periphery of the clause. In *Current studies in Italian syntax: Essays offered to Lorenzo Renzi*, ed. by Guglielmo Cinque and Giampaolo Salvi, 267–296. Amsterdam: Elsevier.
Rizzi, Luigi. 2006. On the form of chains: Criterial positions and ECP effects. In *Wh-movement: Moving on*, ed. by Lisa Lai-Shen Cheng and Norbert Corver, 97-133. Cambridge: MIT Press.
Roberts, Ian. 1991. Excorporation and minimality. *Linguistic Inquiry* 22(1). 209–218.
Roberts, Ian. 1993. *Verbs and Diachronic Syntax: A Comparative History of English and French*. Dordrecht: Kluwer.
Roberts, Ian. 1997. Directionality and word order change in the history of English. In *Parameters of Morphosyntactic Change*, ed. by A. van Kamenade and N. Vincent, 397-426. Cambridge: Cambridge University Press.
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.
Roberts, Ian. 2007. *Diachronic Syntax*. New York: Oxford University Press.
Roberts, Ian and Anders Holmberg 2010: Introduction: Parameters in minimalist theory. In *Parametric variation: Null subjects in minimalist theory*, ed. by Theresa Biberauer, Anders Holmberg, Ian Roberts, & Michelle Sheehan, 1-57. Cambridge: Cambridge University Press.
Roberts, Ian and Anna Roussou. 2003. *Syntactic Change: A Minimalist Approach to Grammaticalization.* Cambridge University Press.
Rudin, Catherine. 1988. On multiple questions and multiple wh-fronting. *Natural Language and Linguistic Theory* 6: 445-501.
Saito, Mamoru. 1989. Scrambling as semantically vacuous A'-movement. In *Alternative conceptions of phrase structure*, ed. by Mark R. Baltin and Anthony S. Kroch, 182-200. Chicago: University of Chicago Press.

Saito, Mamoru. 1992. Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1:69-118.
Saito, Mamoru. 2016. (A) case for labeling: Labeling in languages without phi-feature agreement. In *The Linguistic Review: Special Issue on Labels* 33:129-175.
Saito, Mamoru, and Keiko Murasugi. 1990. N'-deletion in Japanese: A Preliminary Study. *Japanese/Korean Linguistics* 1, 258-301.
Selkirk, Elisabeth. 1972. The phrase phonology of English and French. Doctoral dissertation, MIT.
Şener, Serkan. 2010. (Non-)peripheral matters in Turkish syntax. Doctoral dissertation, University of Connecticut, Storrs.
Shimoyama, Junko. Indeterminate phrase quantification in Japanese. *Natural Language Semantics* 14:139-173.
Spevak, Olga. 2010. *Constituent Order in Classical Latin Prose.* Amsterdan: Benjamins.
Stepanov, Arthur. 2001. Cyclic domains in syntactic theory. Doctoral dissertation, University of Connecticut, Storrs.
Stjepanović, Sandra. 2011. Differential object marking in Serbo-Croatian: Evidence from left branch extraction in negative concord constructions. In *Proceedings of Formal Approaches to Slavic Linguistics 19*, ed. by John Bailyn, Ewan Dunbar, Yakov Kronrod, and Chris LaTerza, 99-115. Ann Arbor, MI: Michigan Slavic Publications.
Takahashi, Daiko. 1994. Minimality of movement. Doctoral dissertation, University of Connecticut, Storrs.
Takita, Kensuke, Nobu Goto, and Yoshiyuki Shibata. 2016. Labeling through Spell-Out. *The Linguistic Review* 33, 177-198.
Tuller, Laurice. 1992. The syntax of postverbal focus constructions in Chadic. *Natural Language and Linguistic Theory* 10, 303-334.
Uriagereka, Juan. 1988. On government. Doctoral dissertation, University of Connecticut, Storrs.
Uriagereka, Juan. 1999. Multiple Spell-Out. In *Working minimalism*, ed. by Sam Epstein and Norbert Hornstein, 251-282. Cambridge, MA: MIT Press.
Uriagereka, Juan. 2012. *Spell-out and the Minimalist Program*. Oxford: Oxford University Press.
Villa-García, Julio. 2015. *The Syntax of Multiple*-que *Sentences in Spanish: Along the Left Periphery*. Amsterdam: John Benjamins.
Walkden, George. 2014. *Syntactic reconstruction and Proto-Germanic*. Oxford University Press.
Wang, Shuyan. 2019. Reconsideration of yi 'one' and classifiers in Mandarin Chinese. Ms., University of Connecticut, Storrs.
Watanabe, Akira. 2002. Loss of overt wh-movement in Old Japanese. In *Syntactic Effects of Morphological Change*, ed. by D. Lightfoot. Oxford.
Whitman, John. 2001. Kayne 1994: p. 143, fn. 3. *The Minimalist Parameter: Selected Papers from the Open Linguistics Forum, Ottawa, 12-23 March 1997*, ed. by G. Alexandrova & O. Arnaudova, 77-100. Amsterdam: John Benjamins.
Willis, David. 2007. Specifier-to-Head reanalyses in the complementizer domain: Evidence from Welsh. *Transactions of the Philological Society* 105 (3):432–480.
Wolfe, Sam. 2018. *Verb Second in Medieval Romance*. Oxford: Oxford University Press.
Wurmbrand, Susi. 2013. QR and selection: Covert evidence for phasehood. In *NELS 42*, 277–290.
Yuan, Michelle. 2017. Movement to complement in Kikuyu and the syntax of focus association. Ms., MIT, Cambridge, MA.
Zuckerman, Shalom. 2001. *The acquisition of optional movement*. Groningen Dissertations in Linguistics 34.
Zyman, Erik. 2018. On the Driving Force for Syntactic Movement. Doctoral dissertation, University of California, Santa Cruz.