Principles and Parameters/Minimalism

Terje Lohndal & Juan Uriagereka
Norwegian University of Science and Technology & University of Maryland

Draft September 9, 2012

1. Introduction

The Minimalist Program (MP) grew out of the theory of Government and Binding (GB), the first instantiation of a theory of Principles and Parameters (P&P) as developed in Chomsky (1981). So both GB and MP are part of the P&P approach to the study of the Faculty of Language (Freidin and Vergnaud 2001, Hornstein, Nunes and Grohmann 2005, Lasnik and Uriagereka 2005, Boeckx and Uriagereka 2006, Freidin in press, Lasnik and Lohndal in press). Rather than conceptualizing them as different frameworks, we view the MP as an attempt to rationalize the principles of GB.

In what follows we focus on the distinction between minimalism conceived as applying principles of science and as making a substantial claim about the architecture of the Faculty of Language. The chapter organizes this general thesis as follows. In section 2 we provide some background for the MP, describing how it evolved. Section 3 is the substantial part of the chapter, where we discuss the consequences of the different views one may have of the MP. In section 4 we discuss some prospects and challenges for minimalism, relating them to the notion of a program. Section 5 concludes the chapter.

2. Background

The first explicitly minimalist paper, Chomsky (1993), was concerned with unifying a certain set of data and theories that had become prominent in the 1980s. What we may thing of as "early minimalism" took what was known from GB for granted, and attempted to unify/eliminate relevant conditions. A good example of this approach is the suggestion that Case checking happens in functional agreement projections.[1]

Consider the data in (1).

(1) a. John likes **him**.
b. **She** chased the cat.
c. Mary proved [$_{IP}$ **him** wrong].

In (1a) we have an object that gets accusative Case, in (1b) a subject that gets nominative Case, and in (1c) an exceptionally Case-marked subject that gets accusative Case (all relevant items in bold). Within GB, the phrasal configurations for these three Cases are different. Accusative Case is assigned to the complement of the verb, nominative Case to the specifier of TP, and exceptionally Case marked subjects in SpecIP of the embedded sentence somehow get their Case from the main verb.[2] Chomsky (1993) reanalyzes these

[1] The capital 'c' denotes abstract case, following Vergnaud (1977) and Chomsky (1981).

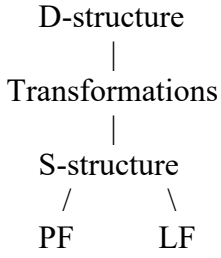
[2] There is a substantial difference between Case *assignment* and Case *checking*, which we set aside here. See Lasnik (2008: 24) for discussion and a historical overview of Case considerations.

cases and argues that all Case checking happens in specifier-head relationships. This is made possible by invoking abstract agreement projections for objects: Chomsky argues that Case checking happens in the agreement projection for objects in both (1a) and (1c). Of course, it is not obvious that **him** in either instance occupies a specifier position in overt syntax. However, the possibility emerges that the checking takes place in the covert component of LF, where the specifier-head relationship instantiate Case checking.[3]

This example illustrates the overall logic. (i) The basic tenets of the GB theory are assumed to be correct: that lexical items get Case, what the Case-bearing heads are, the positions in which lexical items that get Case occupy. (ii) However, the lack of unification in the domains in which lexical items get Case is suspect. Therefore: (iii) The new theory rationalizes what a Case checking domain is, by suggesting the specifier-head relation, independently needed for subject-verb agreement, as the locus for all checking operations – even if abstract moves are thereby invoked (e.g. Case checking happens at LF).[4] In this way the MP attempts to rationalize GB and provide a more principled theory by eliminating overlapping conditions (Freidin and Vergnaud 2001: 642).

Chomsky also argues in the same paper that it is desirable and possible to dispense with two grammar-internal components assumed within GB:

(2)



This is the "T", or "inverted Y", model: The grammar takes lexical items from the lexicon and constructs a D-structure, the locus for theta roles and the Theta Criterion. Transformations map this structure into an S-structure representation (as is patent in the case of a passive or a *wh*-structure). This is the point at which the structure is simultaneously transferred to the sound and the meaning interfaces. Clearly S-structure, and arguably D-structure, are grammar internal representations. Chomsky then explores whether they can be dispensed with, while maintaining empirical coverage:

> Each derivation determines a linguistic expression, an SD, which contains a pair (π, λ) meeting the interface conditions. Ideally, that would be the end of the story: each linguistic expression is an optimal realization of interface conditions expressed in elementary terms (chain link, local X-bar-theoretic relations), a pair (π, λ) satisfying these conditions and generated in the most economical way. Any additional structure or assumptions require empirical justification (Chomsky 1993/1995: 186-187).

---

[3] The other logical possibility, pursued by Koizumi (1995), is that the object in each instance is in fact displaced to a specifier position – but the verb has displaced to an even higher position.

[4] The other presupposition is that conditions obtaining overtly in one language may obtain covertly in others. At the time a variety of studies analyzed object agreement in such languages as Basque, Hindi or Navajo, and the observation was generalized to languages without overt object agreement.

Chomsky notes that the Projection Principle and the Theta Criterion (see Chomsky 1981) argue for D-structure only if their empirical properties must be captured at that level of representation. Work on the configurational properties of theta roles (Hale and Keyser 1993, 2002) and the treatment of theta roles as features (Hornstein 1999) challenges this view (though see Uriagereka 2008). Chomsky (1993) furthermore recalls the problem posed by complex adjectival ("*tough*") constructions, such as (3a). These are assumed to have the S-structure representation in (3b), where *t* represents the trace of an empty operator *Op*.

(3) a. John is easy to please.
b. John is easy [$_{CP}$ Op [$_{IP}$ PRO to please *t*]]
c. it is easy [$_{CP}$ [$_{IP}$ PRO to please John]]

(3c) shows that the matrix subject is not a theta-position; therefore, *John* in (3b) has no role to play in D-structure. Chomsky (1981) liberalizes lexical insertion so as to allow *John* to be inserted in the course of the derivation, thereby being assigned a theta-role at LF. However, as Howard Lasnik observed, this technical solution does not work: a subject of arbitrary complexity may be inserted in place of *John*:

(4) [That Mary is easy to please] is easy to discover.

Literally inserting an entirely formed sentence (which contains the very structure we are analyzing) is of course senseless. Instead, the system is then driven to a version of "generalized transformations" in the sense of Chomsky (1955). That, however, indicates that there is no D-structure level of representation that (4) could start from, given all that we have said so far.[5]

Typical arguments for S-structure involve differences between languages. Some move *wh*-items overtly, others covertly (Huang 1982, Lasnik and Uriagereka 1988). A language may move V to T overtly or covertly (Pollock 1989). The question is whether S-structure is required to account for such differences. Chomsky (1993) argues that it is not, and that languages differ in terms of which features they make use of and whether those features require overt or covert movement for their appropriate checking (see Hornstein, Nunes and Grohmann 2005 for perspective).[6]

The elimination of D- and S-structure paved the way for a perspective on parametric variation that dates back to Borer (1984):

> [P]arametric differences must be reduced to morphological properties if the Minimalist Program is framed in the terms so far assumed. There are strong reasons to suspect that LF conditions are not relevant. We expect languages to be very similar at the LF level, differing only as a reflex of properties detectable at PF; the reasons basically reduce to conditions of learnability. Thus, we expect that

[5] Curiously, it is not clear how to analyze "tough"-movement in minimalist terms, given the strange nature of the null operator. See Hicks (2009) for a review and a recent analysis.

[6] Chomsky (1995b: 197) asserts: "the Minimalist Program permits only one solution to the problem: PF conditions reflecting morphological properties must force V-raising in French but not in English".

> at the LF level there will be no relevant difference between languages with phrases overtly raised or in situ (e.g., *wh*-phrases or verbs). Hence, we are led to seek morphological properties that are reflected at PF (Chomsky 1993/1995: 192).

We will briefly return to features, which essentially trigger and drive derivations, in section 4 (see Adger (2010), Adger and Svenonius (2011), Boeckx (2011)).

An important difference between GB and the MP stems from the focus on *interfaces* and the role they play in a theory. Freidin and Vergnaud (2001: 640-641) see three assumptions that are unique to the MP: (1) the interface levels LF and PF are the only relevant linguistic levels, (2) all conditions are interface conditions, and (3) a linguistic expression is the *optimal* realization of these conditions.[7] Together these assumptions constitute the Strong Minimalist Thesis (Chomsky 2000), which is a conjecture about the organization of the Faculty of Language. The thesis, as such, could of course be right or wrong – but it serves as a working hypothesis to guide theorizing.

That said, in our view there is no paradigm change here: The focus on interfaces is a natural continuation of the approach initiated by Chomsky and Lasnik (1977) with the idea of representational *filters*. Lasnik and Saito's (1984, 1992) very broad *Affect Alpha* transformation takes this idea to its extreme: There are no constraints on whether a category alpha can be affected; only output constraints or filters that apply after the representation has been generated. In our view, directly or indirectly current work tries to rationalize the filters (witness Case conditions) in order to understand why the interfaces are structured the way they are (for some work on this, see Uriagereka 2008, Lohndal 2012, Samuels 2012).

Needless to say, in order to understand the interfaces it is pertinent to seek conditions outside of the internal properties of the linguistic system, which poses a difficult question: How do language-internal properties interact with the rest of the mind? (Boeckx and Uriagereka 2006: 542) take such concerns to 'have become all the more relevant within the MP, especially because this system explores the rational conjecture that some fundamental properties of the language faculty are the way they are precisely *because of* the system's interaction with the rest of the mind'. They mention language acquisition e.g., Crain and Thornton 1998), language change (e.g., Lightfoot 1999) and language use (e.g., Berwick and Weinberg 1984) as examples of cases where linguists have looked at the relevant interactions between the internal system and external properties.

Now, there is a further difference between GB that the MP that runs deeper than the focus on interfaces. We consider this matter in the next section.

---

[7] Optimal realization is another way of saying that the syntax is perfectly organized for the purposes of satisfying the requirements that the interfaces set. Freidin and Vergnaud (2001: 649) provide a useful clarification of the notion of 'perfection':

> […] the notion of 'perfection' often invoked within the MP is *ultimately* a mathematical notion, calling for a higher level of mathematical formalization of syntax [footnote left out]. The Minimalist conjecture that C_{HL} [the computational system of human language] is a 'perfect system' is a tentative claim about the form and the complexity of each computation. The claim is (i) that each computation can be represented as an abstract mathematical structure completely defined by interface (output) conditions and (ii) that this structure is an extremum in some mathematical space.

3. What is Minimalism?

There are various ways to characterize the MP. So far we have discussed how the program largely focuses on *eliminating* theoretical primitives that GB postulated, while hopefully maintaining the same empirical coverage. Following Martin and Uriagereka (2000), we can call this approach *methodological minimalism*:

> What one might call a "weak minimalist thesis" is nothing new. The drive for simple and nonredundant theories of the world (or Occam's razor) is taken for granted in the core sciences. Even within the more specialized science of linguistics, this working methodology has brought undeniable success. From such a perspective, minimalism is just a new way to refer to what many people have been doing for a long time: seeking the best way to theorize about a particular domain of inquiry. We think of this thesis as *methodological minimalism* (Martin and Uriagereka 2000: 1)

We can contrast the methodological view with what the same authors call *ontological minimalism*. This is another name for the Strong Minimalist Thesis, discussed above. Chomsky (2000: 92) states things as follows:

> Suppose that a super-engineer were given design specifications for language: "Here are the conditions that FL must satisfy; your task is to design a device that satisfies these conditions in some optimal manner (the solution might not be unique)." The question is, how close does language come to such optimal design?

From this perspective, the question is not how good our theory of the Faculty of Language is but, rather, how good the Faculty itself is (cf. Chomsky 2000: 141, n. 12). Put differently, we are asking whether the faculty of language is optimally designed. Related to this question is why the computational system of human language is organized the way it is. *Why* do certain principles hold and not others? (Chomsky 2000: 92, Chomsky 2004).

It is important to be aware of the immediate problems related to asking such *why*-questions. Chomsky is rather clear on this, as the following quote illustrates:

> Questions of this kind are not often studied and might not be appropriate at the current level of understanding, which is, after all, still quite thin in a young and rapidly changing approach to the study of a central component of the human brain, perhaps the most complex object in the world, and not well understood beyond its most elementary properties (Chomsky 2000: 93).

The difficulty in providing substantial answers in this domain is not empirical. Rather, it is a question of design: Why is one design preferred to a number of other ones that anyone could easily imagine? Even though such a question is central to minimalism, it is not totally novel (see Freidin and Lasnik 2011 for some roots of minimalism). Berwick and Weinberg (1982: 167) put it like this:

> […] if the "language faculty" is even roughly analogous to other organs of the body (like the heart), then we might reasonably expect, just as in the case of other systems of the boy, that it has been "well designed" according to some as yet undetermined criteria of efficiency. This scenario clearly takes for granted the usual backdrop of natural selection. Since one of the evolutionary "design criteria" could well have been ease of language processing, it is certainly conceivable that efficient parsability has played a role in the shaping of the language faculty (Berwick and Weinberg 1982: 167).

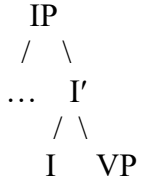
Aside from explicitly raising design questions, this quote foreshadows concerns about the evolution of language. Berwick and Weinberg mention natural selection, though today we know that there are other ways in which good design can arise in nature (cf. Kaufmann (1933) and Gould (2002), among many others).

Because the questions raised by ontological minimalism are extremely challenging, most minimalists have focused on the methodological aspects of minimalism. Though in actual practice it is hard to differentiate between methodological minimalism and ontological minimalism, since the latter by necessity contains elements of the former. Let us briefly consider the development from X-bar theory to Bare Phrase Structure as an example.

3.1. From X-bar theory to Bare Phrase Structure

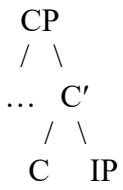
Chomsky (1986a, henceforth *Barriers*) provides a generalization of X-bar structure (Chomsky 1970, Jackendoff 1970), though systematic attempts had already been made in Chomsky (1981), Stowell (1981), den Besten (1983) or Thiersch (1985), to mention some important works.[8] Prior to *Barriers*, the maximal projections were taken to be VP, NP, AP and PP. In addition, there was S (rewritten as NP Infl VP), and S′ (rewritten as Comp S). Comp includes at least C and *wh*-expressions. The immediate problem is that S does not conform to X-bar theory: It is not endocentric since it has no head, which means that there is no projection from a head to a maximal projection. S′ is also not uniformly endocentric since when Comp is filled by complex phrasal material, it cannot the *head* of S′, by definition (a head is not a complex phrase). Because of these problems, Stowell (1981: chapter 6) suggests that the head of S is Infl, as illustrated in (5). This is very similar to the proposal in Williams 1981: 251, which suggests that S is headed by Tense, and also Pesetsky (1982:00), which presents similar considerations:

(5)



[8] We will not be concerned here with the historical development of phrase structure. For a comprehensive review, see Freidin (2007, in press), Lasnik and Lohndal (in press).

Once IP replaces S, a natural step is to reconsider S′. Stowell (1981: chapter 6) proposes that C is the head of S′. The optional specifier then becomes the target of *wh*-movement (see Thiersch 1985 for a detailed elaboration). We then have the structure in (6) (see also Chomsky 1986a).

(6)


With this in place, it is possible to formulate restrictions on movement based on what can appear in a head position and what can appear in a specifier position, cf. Travis (1984) and Rizzi (1990), among many others.

The reanalysis of S and S′ paves the way for a generalization of X-bar theory. Chomsky (1986: 3) proposes that X-bar is a structure as in (7), where X* stands for zero or more occurrences of some maximal projection and X = X^0. [9]

(7) a. $X' = X\ X''^*$
b. $X'' = X''^*\ X'$

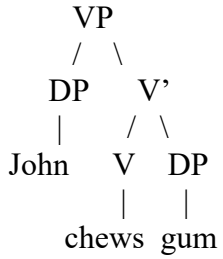
(7) does not force binarity (a node may have more than two daughters). One can either restrict X-bar theory so that it does observe binarity – assuming that it is empirically correct – by "hard-wiring" that condition into the X-bar theory, or else follow proposals like those in Kayne 1984, 1994 to the effect that independent grammatical constraints require all relations of immediate dominance to be binary.

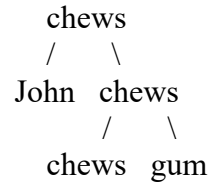
For reasons that we cannot go into now, Kayne's theory forces the elimination of the distinction between X′ and XP. Chomsky (1995a, b) goes further, arguing that X-bar levels should be eliminated altogether. This is an essential feature of the theory of Bare Phrase Structure (BPS). The gist of this theory is summarized as follows: "Minimal and maximal projections must be determined from the structure in which they appear without any specific marking; as proposed by Muysken (1982) they are relational properties of categories, not inherent to them" (Chomsky 1995a: 61).[10] Chomsky (1995b: 242) ties the ban on marking such properties and maximal and minimal projections to the Inclusiveness Condition:[11]

[9] This is what Chomsky said, but it cannot be exactly what he meant. (7a) should read $X' = X\ Y''^*$ because otherwise a verb, for example, could only take a VP complement, and similarly for (7b) and specifiers.
[10] Muysken's (1982) proposal is "that bar level is not a primitive of the grammar at all, rather 'maximal projection' and 'minimal projection' are defined terms, and intermediate projections are simply the elsewhere case" (Muysken 1982). This proposal is closely related to Speas (1990: 35).
[11] This condition could be seen as extension of an idea in Katz and Postal (1964: 44-45), developed in Chomsky (1965: 132), where transformations are taken not to introduce meaning-bearing elements.

(8) *Inclusiveness Condition*
Any structure formed by the computation is constituted of elements already present in the lexical items. No new objects are added in the course of computation apart from rearrangements of lexical properties. (Chomsky 1995b: 228)

Another way to look at BPS is to say that phrase structure consists solely of lexical items. This effectively means that, instead of being represented as in (9), phrases would more accurately be represented as in (10) (setting aside how verbs get their inflection and where the arguments ultimately belong in the structure):

(9)


(10)


These lexical items are accessed at the LF interface, and no units apart from the lexical items can be part of the computation. Thus bar-levels, as such, have no ontological status within BPS. For a critical discussion of BPS see Starke (2004), Jayaseelan (2008), and Lohndal (2012).

The change from X-bar theory to BPS showcases several aspects of minimalism. It shows the methodological part, since BPS makes use of fewer theoretical primitives than X-bar theory. It also shows the ontological part, since BPS makes a different claim about the phrase structure component: There are only lexical items and these, together with constraints on Merge, generate hierarchical phrase structures. We also see how the methodological aspect, by pushing it to some appropriate limit, actually leads to the ontological aspect: (10) involves less information than (9), and the information that it does involve is linguistically more relevant.

Let us now move on to discuss the emphasis on economy conditions that has become a characteristic of minimalism.

3.2. Minimalism and economy

Lappin, Levine and Johnson 1999 claim that economy conditions are new to the MP. However Freidin and Vergnaud (2001) and Freidin and Lasnik (2011) show that this is not the case. Chomsky has been explicit regarding this as well:

> In early work, economy considerations entered as part of the evaluation metric, which, it was assumed, selected a particular instantiation of the permitted format for rule systems, given PLD. […] it seems that economy principles of the kind explored in early work play a significant role in accounting for properties of language. With a proper formulation of such principles, it may be possible to move toward the minimalist design: a theory of language that takes a linguistic expression to be nothing other than a formal object that satisfies the interface conditions in the optimal way. A still further step would be to show that the basic principles of language are formulated in terms of notions drawn from the domain of (virtual) conceptual necessity (Chomsky 1993/1995: 171)

Freidin and Vergnaud (2001: 642) offer the Case Filter as one example, and we will use their discussion as an illustration.

In its original formulation, the Case Filter prohibits a phonetically realized nominal expression that has not been marked for abstract Case in the syntax. The Filter is a stipulation added in order to account for the facts. Within the MP there is no Case Filter as such. Rather a more general principle is assumed: Full Interpretation, originally suggested in Chomsky (1986b). Full Interpretation can be seen as an economy condition: derivations that adhere to it are taken to be better than derivations that do not. The principle requires that all features in a derivation be legible at the relevant interfaces. Thus the principle bans superfluous symbols in general, ruling out vacuous quantification and features that have no values at the interfaces. Case is an instance of the latter, since an uninterpretable Case feature will cause a crash at the LF interface. The underlying assumption is that all features need an interpretable value in order to be legible at the interfaces. While Full Interpretation can be argued to subsume the Case Filter, the converse is not the case.[12]

Freidin and Vergnaud argue that the same logic extends to another example, involving a standard case of an Empty Category Principle (ECP) violation,[13] as in (11), where *t* is the trace of *John* and T is the category Tense.

(11) *John T is believed [that *t* T is happy]

Within GB the ungrammaticality of this example was taken to follow from the fact that the trace *t* was not properly licensed (see fn. 13). Freidin and Vergnaud (2001: 642-642) propose to subsume the ungrammaticality of (11) under Full Interpretation. They assume that the nominative Case feature of *John* is checked in the embedded clause by T_2, as is the nominative Case feature of T_2. When *John* moves to the subject position of the matrix clause, the nominative Case feature of *John* has already been checked. This means that the nominative Case feature on T_1 cannot be checked. An unchecked feature causes the derivation to crash. Thus Full Interpretation accounts for data that previously were

[12] This discussion also presupposes that the Case Filter is an interface condition, cf. Chomsky (1995b: 197), and of course poses the question of what Case features ultimately are or why they are part of the computation (see Pesetsky and Torrego 2001, Uriagereka 2002: chapter 8, and 2009).

[13] The ECP presents conditions under which traces (of movement) are allowed to exist within given configurations, relating to whether such configurations are directly associated to lexical verbs or similar elements, or whether their antecedent is in some definable sense local to the trace.

accounted for using the Case Filter and the ECP. This is a good example of the eliminativist angle of minimalism at work: Two apparently unrelated conditions are reduced to a single one at a different level of abstraction.

Freidin and Vergnaud (2001: 643) welcome the use of Full Interpretation instead of the Case Filter and the ECP:

> [..] this analysis, which explains deviance on the basis of legibility conditions imposed by cognitive systems that interface with C_{HL} [the computational system of human languages], strikes us as a more promising explanatory account than the postulation of various constraints internal to C_{HL} that basically reflect the complexity of the phenomena in an essentially descriptive fashion.

As an interface condition, Full Interpretation demands that all elements introduced in a derivation be legible at the interfaces. This has an interface filtering effect: derivations that do not fulfill this requirement do not converge. Surely Full Interpretation is more fundamental than the more ad-hoc Case Filter or ECP.

The discussion above has provided a concrete example of how minimalism focuses on interface conditions. The current hypothesis is actually that there are multiple points at which the derivation transfers to the interfaces (Uriagereka 1999, building on Bresnan 1972; Chomsky 2000, 2001). These transfer points are usually called 'phases', and there has been much work regarding their structure and nature. See Gallego (2012) for a comprehensive discussion that space considerations prevents us from undertaking.

Another example of economy comes from what Rizzi (1990) calls Relativized Minimality.[14] Rizzi proposes a unified account of the following cases:

(12) a. *[$_{\alpha}$ Fix] John [$_{\beta}$ can] t(α) the car
b. *[$_{\alpha}$ John] seems [$_{\beta}$ it] is certain t(α) to be here.
c. *[$_{\alpha}$ how] did John wonder [$_{CP}$ [$_{\beta}$ what] Mary fixed t(β) t(α)]

In (12a) the Head Movement Constraint is violated, (12b) is an instance of superraising, and in (12c) the *wh*-island constraint is violated. Contrary to the traditional picture we just outlined, Rizzi argues that the same violation obtains in these cases. He proposed Relativized Minimality, given in (13).

(13) *Relativized Minimality*
α cannot cross (= move to a position c-commanding) β if β c-commands α, and β is the same type as α.

The relevant types are characterized as follows (Rizzi 1990).

(14) a. *Head positions*
If α adjoins to a head, β is a head.
b. *A-positions*
If α adjoins to an A-position, β is an A-specifier.

[14] The following summary is based on Kitahara (1997), to which we refer the reader for a more comprehensive discussion.

c. *A-bar positions*
If α adjoins to an A-bar specifier, β is an A-bar specifier.

The account is that in (12a), the head α moves across another head β, in (12b) α crosses the embedded A-specifier β, and in (12c) α crosses the embedded A-bar specifier β.

Rizzi's Relativized Minimality is a representational principle – it applies to the entire representation that has been generated in the transformational component. Chomsky and Lasnik (1993), in contrast, interpret the principle in terms of 'least effort', an economy consideration. We can illustrate by considering the following data.

(15) a. Guess who bought what?
b. *Guess what who bought?

This is a 'superiority effect', which Chomsky and Lasnik analyze in terms of Relativized Minimality. There might seem to exist two derivational options in (12), as one ought to be able to front either *who* or *what*. As the contrast shows, however, that is not the case, and only *who* can be fronted. Studying the relevant paradigms the generalization appears to be that one always has to involve the question word that is 'closest to the position where it ends up moving', as first observed by Chomsky (1973). Another way to put this is that the distance 'traveled' by the moving element has to be minimized. This account can be generalized to the cases in (12) as well. The minimalist approach is to take this as an instance of economy in the derivation and not as a condition on representations.

More generally, conditions arising with respect to movement operations have been captured in terms of Relativized Minimality (see Starke 2001 for comprehensive discussion), when combined with considerations about phases (Uriagereka 1998, Chomsky 2008, Müller 2011). Space considerations prevent us from discussing this here.

For more discussion of economy, see e.g. Collins (1997), Kitahara (1997), Reinhart (2006), and Lasnik and Lohndal (in press).

3.3. The programmatic nature of minimalism

Since Chomsky (1993), Chomsky and other scholars have consistently insisted that minimalism is a program, not a theory. The following quote makes this clear:

> [t]here are minimalist questions, but no minimalist answers, apart from those found in pursuing the program: perhaps that it makes no sense, or that it makes sense but is premature (Chomsky 2000: 92).

The most comprehensive attempt at justifying the programmatic nature of minimalism can be found in Boeckx (2006), which we now rely on.

Programs are by nature abstract. Boeckx uses Hilbert's Program as an illustration. David Hilbert proposed this program in the early 1920s as a call for a formalization of all of mathematics in axiomatic form, together with a proof that this axiomatization of mathematics is consistent (Boeckx 2006: 86). He understood programs as a set of guidelines and proposed boundary conditions that could be used to determine whether the program was successful. Although Kurt Gödel proved that Hilbert's program cannot be

carried out to completion, it has had an important influence not just on mathematics, but also in logic, computer science and beyond.

As Boeckx points out, physicists rarely use programs. However, he quotes Feyerabend to illustrates that the notion can be applied to physics (Boeckx 2006: 86):

> There may not exist a single theory, one 'quantum theory', that is used in the same way by all physicists. The difference between Bohr, Dirac, Feynman and von Neumann suggests that this is more than a distant possibility […] (Feyerabend 1993:191).

Lakatos (1970) narrows programs down to *research* programs, identifying properties that characterize these: programs have a core, are slow to mature and their rigor may not be stellar at first – but they present an openness and flexibility that may lead to theoretical insight. For the MP, the core is arguably the optimal connection with interfaces: All research within the MP is to create a theory of syntax that mediates sound and meaning.[15] Beyond this core, there is a lot of variation, which Boeckx (2006: 94-94) contextualizes:

> As Lakatos stressed, programs are more than a core. They consist of auxiliary hypotheses that may vary from one researcher to the next, and they have a heuristic, a plan for addressing problems... [. . .] Different auxiliary hypotheses, or different arrangement of the elements of a program's core, may lead to radically different questions, empirical problems, solutions, etc.

In the programmatic perspective, the object of study has to be approached with an open mind and different ways of answering a question may need to be entertained – or there will be unanswered questions. Freidin and Vergnaud (2001: 647) provide the following quote from Chomsky (1980: 9-10) regarding the consequences of implementing what is typically called the Galilean style:

> [it entails a] readiness to tolerate unexplained phenomena or even as yet unexplained counterevidence to theoretical constructions that have achieved a certain degree of explanatory depth in some limited domain, much as Galileo did not abandon his enterprise because he was unable to give a coherent explanation for the fact that objects do not fly off the earth's surface.

Evidently, the challenge for a research program is to determine whether it is 'progressive' or 'degenerative', to use Lakatos's terms: Whether it produces results that move the program forward (e.g., unexpected predictions), or assumptions are just made in order to accommodate the facts. Although we have our own opinions, we will not take a position on that matter here with regards to MP, since it is too early to make predictions.

4. Prospects and challenges

[15] Boeckx himself argues in (2006) that the core of minimalism consists of a focus on (i) economy, (ii) virtual conceptual necessity, and (iii) symmetry.

Empirical progress within MP is evident to us in various areas, such as the theory of ellipsis (see e.g. Merchant 2001, Lasnik 2001). The framework is also able to analyze scores of different languages, continuing in the tradition of its predecessor. Most work adopt a rich number of features, since features are seen as the "driving force" of a derivation; see Adger (2010) and Adger and Svenonius (2011) for comprehensive discussion, and Boeckx (2010) for some problems. These features and their flexibility make it generally possible to achieve cross-linguistic coverage. Together with categorical assumptions made in the cartographic approach (Rizzi 1997, Cinque 1999, and many others), features have made it possible to achieve a great deal of descriptive adequacy. However, questions remain about their explanatory adequacy, both in terms of how a child acquires the relevant subsets of projections and features, and more generally in terms of the nature of the features themselves and how they relate to one another.

Some scholars have argued that it is time to produce theories that can be evaluated and falsified (see e.g., Hornstein 2009). Soon after the MP was initiated, Lasnik (1999: 6) stated that 'there is not yet anything close to a Minimalist theory of language'. Differences in the underlying philosophy of science will in part determine whether a scholar regards programs as sufficient or whether theories are required for genuine progress. The basic question here is related to whether an approach has to be falsified. Lakatos (1968, 1970) argues that it should not be easy to refute a program; rather, one sticks to the core assumptions and tries to develop it in various ways.

In many ways this is exactly what has happened within the MP. It has generated new work on the interfaces and their relationship to syntax, a sign that the program is still progressive (see for instance Hauser et al. 2002 and the various reactions it provoked, or the array of authors in Di Sciullo et al. 2009 or Piattelli-Palmarini et al. 2011, most of whom would term their approach as both "minimalist" and "interdisciplinary"). However, that does not mean that practitioners can avoid addressing fundamental questions such as why we the system has the features it does, or why the interfaces are structured the way they are. Issues along these lines, surrounding features and connections between the narrow and broad perspectives of the Faculty of Language (in the sense in Hauser et al. 2002), will likely be at the forefront of much forthcoming research.

5. Conclusions

This chapter has surveyed some core issues within the Principles and Parameters approach to the study of language. We have focused on the transition from GB to the MP, and have tried to explain the rationale behind the change. We have argued that the MP is an attempt to rationalize GB and to move beyond explanatory adequacy. We then discussed the nature of minimalism, and presented two approaches: methodological and ontological minimalism. We also used the change from X-bar theory to Bare Phrase Structure to illustrate these conceptions. Then we discussed why minimalism is a program and what that label entails, focusing on Lakatos's work. Lastly, we discussed a few challenges for minimalism and suggested that whether minimalism should be seen as a program or a theory by and large depends on one's view of philosophy of science.

References

Adger, David. 2010. A minimalist theory of feature structure. In *Features:*

Perspectives on a Key Notion in Linguistics, ed. by Anna Kibort and Greville G. Corbett, 185-218. Oxford: Oxford University Press.
Adger, David and Peter Svenonius. 2011. Features in Minimalist Syntax. In *The Oxford Handbook of Linguistic Minimalism*, ed. by Cedric Boeckx, 27-51. Oxford: Oxford University Press.
Berwick, Robert C. and Amy S. Weinberg. 1982. Parsing Efficiency, Computational Complexity, and the Evaluation of Grammatical Theories. *Linguistic Inquiry* 13: 165-191.
Berwick, Robert C. and Amy S. Weinberg. 1984. *The Grammatical Basis of Linguistic Performance*. Cambridge, MA: The MIT Press.
den Besten, Hans. 1983. On the interaction of root transformations and lexical deletive rules. In *On the Formal Syntax of the Westgermania*, ed. by Werner Abraham , 47-131. Amsterdam: John Benjamins.
Boeckx, Cedric. 2006. *Linguistic Minimalism*. Oxford: Oxford University Press.
Boeckx, Cedric. 2010. A tale of two minimalisms – Reflections on the plausibility of Crash-Proof syntax and its Free Merge alternative. In *Derivational Simplicity in Minimalist Syntax*, ed. by Michael Putnam, 105-124. Amsterdam: John Benjamins.
Boeckx, Cedric and Juan Uriagereka. 2006. Minimalism. In *The Oxford Handbook of Linguistic Minimalism*, ed. by Gillian Ramchand and Charles Reiss, 541-573. Oxford: Oxford University Press.
Borer, Hagit. 1984. *Parametric Syntax*. Dordrecht: Foris.
Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press.
Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in English Transformational Grammar*, ed. by Roderick A. Jacobs and Peter S. Rosenbaum, 184-221. Waltham, Mass.: Ginn.
Chomsky, Noam. 1980. On binding. *Linguistic Inquiry* 11: 1-46.
Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
Chomsky, Noam. 1986a. *Barriers*. Cambridge, MA: The MIT Press.
Chomsky, Noam. 1986b. *Knowledge of Language*. New York: Praeger.
Chomsky, Noam. 1993. A Minimalist Program for Linguistic Theory. In *The View from Building 20*, ed. by Kenneth Hale and Samuel Jay Keyser, 1-52. Cambridge, MA: The MIT Press.
Chomsky, Noam. 1995a. Bare phrase structure. *Evolution and Revolution in Linguistic Theory: A Festschrift in Honor of Carlos Otero*, ed. by Héctor Campos and Paula Kempchinsky, 51-109. Washington D.C.: Georgetown University Press.
Chomsky, Noam. 1995b. *The Minimalist Program*. Cambridge, MA: The MIT Press.
Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. 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: The MIT Press.
Chomsky, Noam. 2001. Derivation by Phase. In *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 1-50. Cambridge, MA: The MIT Press.
Chomsky, Noam. 2004. Beyond Explanatory Adequacy. In *Structure and Beyond: The*

Cartography of Syntactic Structures, ed. by Adriana Belletti, 104-131. Oxford: Oxford University Press.
Chomsky, Noam. 2008. On Phases. Chomsky, Noam. 2008. On Phases. *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, ed. by Carlos Otero, Robert Freidin and María-Luísa Zubizarreta, 133-166. Cambridge, MA: The MIT Press.
Chomsky, Noam and Howard Lasnik. 1993. The theory of principles and parameters. In *Syntax: An International Handbook of Contemporary Research*, ed. by Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld and Theo Venneman, 506-569. New York: Walter de Gruyter.
Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective*. Oxford: Oxford University Press.
Crain, Stephen and Rosalind Thornton. 1998. *Investigations in Universal Grammar*. Cambridge, MA: The MIT Press.
Feyerabend, Paul. 1993. *Against Method*. New York: Verso.
Freidin, Robert. 2007. *Generative Grammar: Theory and its History*. London: Routledge.
Freidin, Robert. In press. A brief history of generative grammar. In *The Routledge Handbook of the Philosophy of Language*, ed. by Delia Fara Graf and Gillian Russell. London: Routledge.
Freidin, Robert and Howard Lasnik. 2011. Some Roots of Minimalism. In *The Oxford Handbook of Linguistic Minimalism*, ed. by Cedric Boeckx, 1-26. Oxford: Oxford University Press.
Freidin, Robert and Jean-Roger Vergnaud. 2001. Exquisite Connections: Some Remarks on the Evolution of Linguistic Theory. *Lingua* 111: 639-666.
Ángel J. Gallego (ed.). 2012. *Phases: Developing the Framework*. Berlin: Mouton de Gruyter.
Gould, Stephen Jay. 2002. *The Structure of Evolutionary Theory*. Cambridge: Harvard University Press.
Hale, Kenneth and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of grammatical relations. In *The View from Building 20*, ed. by Kenneth Hale and Samuel Jay Keyser, 53-110. Cambridge, MA: The MIT Press.
Hale, Kenneth and Samuel Jay Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. Cambridge, MA: The MIT Press.
Hauser, Marc D., Noam Chomsky and W. Tecumseh Fitch. 2002. The Faculty of Language: What Is It, Who Has It, and How Did it Evolve? *Science* 298: 1569-1579.
Hicks, Glyn. 2009. *Tough*-Constructions and Their Derivation. *Linguistic Inquiry* 40: 535-566.
Hornstein, Norbert. 1999. Movement and Control. *Linguistic Inquiry* 30: 69-96.
Hornstein, Norbert. 2009. *A Theory of Syntax*. Cambridge: Cambridge University Press.
Hornstein, Norbert, Jairo Nunes and Kleanthes K. Grohmann. 2005. *Understanding Minimalism*. Cambridge: Cambridge University Press.
Huang, C. T. James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT.

Jayaseelan, K. A. 2008. Bare phrase structure and specifier-less syntax. *Biolinguistics* 2: 87-106.
Jenkins, Lyle. 2000. *Biolinguistics*. Cambridge: Cambridge University Press.
Katz, Jerrold J. and Paul Postal. 1964. *An Integrated Theory of Linguistic Descriptions*. Cambridge, MA: The MIT Press.
Kaufmann, Stuart. 1993. *Origins of Order: Self-organization and Selection in Evolution*. Oxford: Oxford University Press.
Kayne, Richard S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.
Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: The MIT Press.
Kitahara, Hisatsugu. 1997. *Elementary Operations and Optimal Derivations*. Cambridge, MA: The MIT Press.
Lakatos, Imre. 1968. Criticism and the Methodology of Scientific Research Programmes. *Proceedings of the Aristotelian Society* 69: 149-186.
Lakatos, Imre. 1970. Falsifications and the methodology of scientific research programs. In *Criticism and the Growth of Knowledge*, ed. by Imre Lakatos and Alan Musgrave, 91-195. Cambridge: Cambridge University Press.
Lappin, Shalom, Robert Levine and David Johnson. 2000. The structure of unscientific revolutions. *Natural Language and Linguistic Theory* 18: 665-671.
Lasnik, Howard. 1999. *Minimalist Analysis*. Oxford: Blackwell.
Lasnik, Howard. 2001. When Can You Save a Structure by Destroying It? *Proceedings of NELS* 31: 301-320.
Lasnik, Howard. 2008. On the Development of Case Theory: Triumphs and Challenges. In *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, ed. by Carlos Otero, Robert Freidin and María-Luísa Zubizarreta, 17-41. Cambridge, MA: The MIT Press.
Lasnik, Howard and Terje Lohndal. In press. Brief Overview of the History of Generative Syntax. In *The Cambridge Handbook of Generative Syntax*, ed. by Marcel den Dikken. Cambridge: Cambridge University Press.
Lasnik, Howard and Juan Uriagereka. 1988. *A course in GB Syntax*. Cambridge, MA: The MIT Press.
Lightfoot, David W. 1999. *The Development of Language*. Malden: Blackwell.
Lohndal, Terje. 2012. Without Specifiers: Phrase Structure and Events. Doctoral dissertation, University of Maryland.
Lohndal, Terje and Juan Uriagereka. In press. Third factor explanations and Universal Grammar. In *The Oxford Handbook of Universal Grammar*, ed. by Ian Roberts. Oxford: Oxford University Press.
Martin, Roger and Juan Uriagereka. 2000. Introduction: Some Possible Foundations of the Minimalist Program. In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. by Roger Martin, David Michaels and Juan Uriagereka, 1-29. Cambridge, MA: The MIT Press.
Merchant, Jason. 2001. *The Syntax of Silence*. Oxford: Oxford University Press.
Muysken, Pieter. 1982. Parametrizing the notion "Head". *Journal of Linguistic Research* 2: 57-75.
Müller, Gereon. 2011. *Constraints on Displacement. A Phase-Based Approach*. Amsterdam : John Benjamins.
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 : The MIT Press.
Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20: 365-424.
Popper, Karl. 1959. *The Logic of Scientific Discovery*. London : Hutchinson.
Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, MA: The MIT Press.
Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In *Elements of Grammar: A Handbook of Generative Syntax*, ed. by Liliane Haegeman, 281-337. Dordrecht: Kluwer.
Samuels, Bridget. 2012. *Phonological Architecture: A Biolinguistic Perspective*. Oxford: Oxford University Press.
Speas, Margaret J. 1990. *Phrase Structure in Natural Language*. Dordrecht: Kluwer.
Starke, Michal. 2004. On the inexistence of specifiers and the nature of heads. In *Structure and Beyond: The Cartography of Syntactic Structures*, ed. by Adriana Belletti, 252-268. Oxford: Oxford University Press.
Stowell, Tim. 1981. Origins of Phrase Structure. Doctoral dissertation, MIT.
Thiersch, Craig. 1985. VP and Scrambling in the German Mittelfeld. Ms., University of Tilburg.
Travis, Lisa. 1984. Parameters and Effects of Word Order Variation. Doctoral dissertation, MIT.
Uriagereka, Juan. 1998. *Rhyme and Reason*. Cambridge, MA: The MIT Press.
Uriagereka, Juan. 1999. Multiple Spell-Out. *Working Minimalism*, ed. by Samuel David Epstein and Norbert Hornstein, 251-282. Cambridge, MA: The MIT Press.
Uriagereka, Juan. 2002. *Derivations*. London: Routledge.
Uriagereka, Juan. 2008. *Syntactic Anchors*. Cambridge: Cambridge University Press.
Uriagereka, Juan. 2009. Uninterpretable Features in Syntactic Evolution. In *Of Minds and Language: A Dialogue with Noam Chomsky in the Basque Country*, ed. by Massimo Piattelli-Palmarini, Juan Uriagereka and Pello Salaburu, xx-xx. Oxford: Oxford University Press.
Williams, Edwin. 1981. On the notions “lexically related” and “head of a word”. *Linguistic Inquiry* 12: 245-274.