

Considerations pertaining to the nature of logodiversity, or How to construct a parametric space without parameters*

Cedric Boeckx^{1,2}

¹Catalan Institute for Advanced Studies and Research (ICREA)

&

²Universitat de Barcelona

`cedric.boeckx@ub.edu`

February 2012—*First draft*

1 Lessons from comparative linguistics

A central explanatory goal of the branch of linguistics that I will be concerned with here (*cf.* Chomsky (1965)) and that I will refer to as *biolinguistics* in what follows

*The present material incorporates ideas presented at the *Workshop on Formal Grammar and Syntactic Variation* held at the Universidad Autonoma de Madrid in October 2011, as well as in lectures at the Universities of Oslo and Porto (Fall 2011). It builds on claims and extends proposals made in Boeckx (2011a, 2010). As always, I have benefited greatly from the comments I received from people in the audience whenever I touched on these ideas. I also want to note that, although I take issue with many of their recent statements on parameters, I have learned a tremendous amount from the writings of R. Kayne, M. Baker, L. Rizzi, and, very especially, I. Roberts concerning linguistic variation, its scope and its limits. I also want to thank Martin Nowak for giving me the opportunity some years back to join the Program for Evolutionary Dynamics at Harvard University, and present an early version of this paper back in 2004. The present work is supported by a Marie Curie International Reintegration Grant from the European Union (PIRG-GA-2009-256413), research funds from the Fundació Bosch i Gimpera, and a grant from the Spanish Ministry of Economy and Competitiveness (FFI-2010-20634; PI: Boeckx).

is to account for the remarkable language acquisition ability that any human child displays (barring severe pathology or highly unusual environmental circumstances). Such an ability has now been extremely well documented. Its absence in otherwise closely related and highly intelligent species or machines strongly suggests the existence of a biological basis for it in humans. This much is common background assumption among many linguists and cognitive scientists, over a broad range of theoretical persuasions. Controversies arise (in a healthy fashion) concerning the specific content of this biological basis. Over much of its history, research in generative grammar, whose research agenda provided, and continues to provide much of the basis for biolinguistic inquiry, has taken this biological basis to be highly specific to the domain of language, and has often referred to it in genetic terms. It is only recently that this stance has been softened somewhat, with generative grammarians recognizing (though, it seems to me, not yet fully embracing—see below) the importance of epigenetic and more generic factors (*cf.* Chomsky (2005)), thereby agreeing in part with claims found in other linguistic traditions.

It is in this context that I would like to reflect on a central empirical fact that any adequate account of the human language acquisition capacity must come to grips with: the fact that the linguistic system that any human child can acquire is not limited to one. I am not here referring to the existence of bi-, or multi-lingualism, but rather to the even more obvious fact that a large variety of languages are, have been, and will be, acquired by children of our species.

Understanding the differences, and the similarities, among the languages acquirable by humans is of paramount importance, and is the focus of inquiry of what is traditionally known as comparative linguistics.¹ Progress in comparative linguistics will enable us to grasp the space of possibilities through which the child will have to navigate successfully in order to acquire a specific language. Once defined, this space of possibilities—call it the grammatical morphospace²—can then become

¹Comparative linguistics in this sense is quite distinct from comparative cognitive psychology, which focuses on homologs and analogs across species. A field that one could call *comparative biolinguistics*, which would focus on the research program advocated by Hauser, Chomsky, and Fitch (2002), in particular their “Faculty of Language in the Broad Sense”, would be much more in line with work in comparative psychology.

²As I mentioned in Boeckx (2006, chap.4), there is an obvious parallelism between the study of the range of possible, acquirable human languages studied by the linguists, and the morphospaces studied by the branch of biology known as Theoretical Morphology, whose essence is well captured in the following quote from McGhee (1999, 2):

The goal [of Theoretical Morphology] is to explore the possible range of morphologic variability that nature could produce by constructing n -dimensional geometric hyper-spaces (termed “theoretical morphospaces”), which can be produced by systematically

the focus of intensive biolinguistic inquiry seeking to disentangle the contribution of the various factors (environmental, genetic, and developmental) that determine the shape of this morphospace, bearing in mind, of course, the fact—well assimilated by biologists (see Lewontin (2000), Pigliucci and Müller (2010)), but, it seems, not so well (yet) by many linguists, judging from passages like Gallego (2011b, 538f.)—that all these factors mutually define one another and are therefore not completely separable. The present work is an attempt in this direction.

It is unquestionable that significant progress has been made in the domain of comparative linguistics over the past sixty years (the period of time I will use here as reference frame, as it corresponds to the modern period of biolinguistic investigation, beginning with Chomsky (1959) and the early work of Eric Lenneberg, culminating in Lenneberg (1967)). Kayne (2011, 1) writes of the “qualitative improvement” of our “knowledge and understanding” of human language (specifically, “of human language syntax”) made possible by “advances” in comparative linguistics (specifically, “in the subfield of syntax called comparative syntax”). Few, I think, would disagree with Kayne. I certainly agree, and like him, I think that “comparative syntax in all its range provides us with a window on the language faculty that is just beginning to bear fruit”. But, while agreeing with him concerning the significant progress made, I disagree with Kayne on what we have come to learn, for reasons that I will detail in this paper (see also Boeckx (2010, 2011a)).

Like Kayne and many others, I take the development of the “Principles-and-Parameters” proposal (Chomsky (1981) and much subsequent work) to have played a major role in shedding light on the nature of logodiversity, and thereby on the general character of the language faculty.³ But Kayne, along with other major pro-

varying the parameter values of a geometric model of form. [...] Once constructed, the range of existent variability in form may be examined in this hypothetical morphospace, both to quantify the range of existent form and to reveal nonexistent organic form. That is, to reveal morphologies that theoretically could exist (...) but that never have been produced in the process of organic evolution on the planet Earth. The ultimate goal of this area of research is to understand why existent form actually exists and why nonexistent form does not.

³Kayne (2011) is certainly right in stressing that

If anything, the primary importance of comparative syntax lies in the fact that it provides us with new kinds of evidence bearing on questions concerning the general character of the language faculty. Figuring out what cross-linguistic generalizations hold and why exactly they hold will invariably help us to narrow down the set of hypotheses and principles that we entertain about the language faculty; it will always help us, whether our work is thought of as minimalist or not, to counteract the disadvantage that we have of being so smart that we can all too readily invent new

ponents and architects of the “Principles-and-Parameters” model, takes the results achieved to point to the fundamental correctness of the model, whereas I take the thirty years of work within the the “Principles-and-Parameters” model to point in the opposite direction. It seems to me that what we have learned from the patterns of variation uncovered over the years is that the classical⁴ notion of parameter put forth in Chomsky (1981) does not hold.

When I voice such an opinion in talks, I consistently get two very defensive responses from advocates of the parametric model (researchers in other traditions need no convincing on my part, and in fact often point out that they have said it all along, although my arguments are quite distinct from their, even if they come to a similar conclusion). The first defensive response is that if we indeed say that the parametric model failed, we are essentially giving up on the explanatory benefits of the model, well discussed by Chomsky and others, the very properties of the model that led many (not without reason) to characterize the Principles-and-Parameters approach as a major breakthrough in the history of linguistics, as it provided us, arguably for the first time, with a way to make linguistic variation finite, and also gave us some hope of “solving” what Chomsky (1986) call “Plato’s problem” (the logical problem of language acquisition).

The second defensive response I consistently get is that of course, the classical notion of parameter that I have in mind (the one in Chomsky (1981)) no longer holds, but every proponent of the parametric approach ceased to adopt this notion of parameter long ago, replacing it with a restrictive notion, often attributed to Borer (1984), which takes all variation to boil down to lexical parameters. This change, I’m told, marks an improvement of the parametric model, not, as I would have it, its demise.

Both responses are interesting, but ultimately fail to reach their targets. First, the explanatory benefits of the original parametric model are not in fact exploited

mechanisms to accommodate new data or new problems.

⁴As I observed in Boeckx (2010), I am aware of the existence of a much watered-down notion of parameter (which I dubbed ‘lower-case *p*’) in the literature, which I think is currently used massively, though not always explicitly, to provide artificial life support to the Principles-and-Parameters model. This notion of parameter is not what I am focusing on here, as it is clearly devoid of any theoretical teeth, unlike the notion of parameter put forth in Chomsky (1981). Of course, if one wishes to use the term ‘parameter’ as a synonym for ‘difference’, then I also agree that there are “parameters” between John’s English and Koji’s Japanese. But used in this way, ‘parameter’ is “nothing but jargon for language-particular rule” (Newmeyer (2005, 53) . I take it that advocates of Principles-and-Parameters model are trying (or at any rate, should try) to advocate something stronger (and more interesting), though. Otherwise, they are the ones who are throwing the baby with the bathwater.

by current proposals. As I will show below, such proposals are not really concerned with Plato’s problem, but rather seek to capture, in a more formal fashion, results from the typological literature, much of which is dedicated to the issue of ‘probable’ (as opposed to ‘possible’) language, and therefore at variance with the focus of the Principles-and-Parameters approach, as Newmeyer (2005) correctly points out. Second, the restrictive character of lexical parameter is a myth, until we have a restrictive theory of the lexicon and of what counts as a possible lexical entry. Attempts in this direction, which I will review below, quickly reveal serious problems for lexical parameters. Moreover, a close reading of Borer’s 1984 proposal concerning the benefits of confining variation to the lexicon reveals a certain degree of skepticism regarding parametric models, contrary to what is often claimed in the literature (see next section). Finally, it is not true that the lexical parametrization hypothesis is universally accepted. Prominent proposals that go well beyond lexical parametrization exist (they tend to revolve around the notion of “macroparameter”, and will be discussed below), and they provide a perfect illustration of the more substantive notion of parameter I have in mind. And because they can be shown to be so problematic, they constitute some of the best arguments against parametric models.

Before engaging with the literature, and back up the statements just made, let me point out that contrary to what is often claimed by proponents of the Principles-and-Parameters approach, the case against parameters of which this paper is an example is not solely, nor even primarily motivated by ‘minimalist’ considerations.⁵ Ultimately, my argument against parameters, much like Newmeyer’s or others’, is downright empirical.⁶ Conceptually, it would be hard to argue against the classic notion of parameter. It’s precisely because it does not seem to be what we find, given what we have learned at the empirical level, that an alternative to solving Plato’s problem must be found.

2 Parameters: Lexical or syntactic?

Despite the ease and frequency with which it is used in the literature, as Koster (2010) correctly points out, “the notion “parameter” has always remained underde-

⁵For example, Baker (2011) writes that “the minimalist ethos puts pressure against [parameters]” and that “We should aim not only for beautiful theories, but for beautiful analyses of languages.” In a similar vein, Kayne (2011) says “the rhetoric of the field sometimes puts disproportionate emphasis on the explanatory frontiers at the expense of the observational/descriptive”; comments that suggest that conceptual arguments prevail over empirical ones in this domain.

⁶I should note, though, that the empirical domain that I concentrate on is the language faculty, which is quite different from ‘languages’, as Chomsky has not tired of pointing out over 60 years.

veloped from a theoretical point of view.” That there are patterns of variation across languages is, I think, beyond doubt: variation is not infinite, it is discrete, but quite dense (in the sense that the number of points of variation is very large indeed).

A few linguists continue to insist on the existence of what Snyder (2011) calls “constructive parameters”. Such parameters, within which, according to Snyder, syntactic parameters fall, have “the effect of adding new ‘building blocks’ or new structure-building operations”.

For example, Snyder puts forth his well-known Compounding Parameter as an example of such constructive parameter. Snyder’s Compounding Parameter rests on the observation that languages differ sharply in whether they allow endocentric, bare-root compounding as a fully productive process. Thus, whereas English allows for things like *university lab space committee decision*, Catalan does not. Snyder noted that English-type languages display certain constructions that are systematically absent from Catalan-type languages, such as Verb-NP-particle constructions, adjectival resultative constructions, and so on. On the basis of this, Snyder put forth the idea that English-type languages allow for a “General Modification” rule⁷ that is lacking in Catalan-type languages. It is in this sense that the Compounding Parameter qualifies as a “constructive parameter”, since “its positive setting provides a semantic composition rule for syntactic combinations that would otherwise be uninterpretable” (Snyder (2011)).

There are, however, several problems with Snyder’s proposal, all of which are characteristic of the difficulties suffered by parameters (in the classical sense) over the years. First, the full range of grammatical consequences predicted by the parameter on the basis of data from English does not always hold across a wider range of languages, as can be gathered from Son (2006). This typically leads to a weakening of the parameter (from biconditional, ‘Property X iff Property Y’ statements to weaker, unidirectional, ‘Property X if Property Y’ claims), an important point I return to below. Second, a proposal like Snyder faces difficulties in explaining the existence of compounds in Catalan-type languages. To be sure, these are not as productive as in English, but their existence begs the question of how they were generated in the first place if the rule behind them—by hypothesis, the General Modification rule

⁷I do not think that the specific details of Snyder’s formulation of the rule matter for the point that I am making in the text, but for the sake of concreteness, here is Snyder’s latest formulation of his General Modification rule (from Snyder (2011)):

If α and β are syntactic sisters under γ , where α is the head of γ and denotes a kind, then interpret γ semantically as a subtype of the kind α , and as standing in a pragmatically suitable relation to the denotation of β .

put forth by Snyder—is set to ‘off’ in the language. Lack of productivity is a matter of language use. Third, Snyder’s proposal reveals the importance of properly formulated principles needed to embed parameters in. In a restrictive semantic framework such as Pietroski (2005), the limited repertoire of interpretive operations makes it impossible for languages to ‘deactivate’ some of them. Indeed, it is difficult to see how a language would be able to completely do away without as general a rule as General Modification. Fourth, there exist much more ‘surfacy’ explanations for the highly limited availability of compounds of the English-type in Catalan-type languages, such as Tokizaki (2010, 2011), who shows how the cross-linguistic variation that Snyder is concerned with can be predicted on the basis of the canonical word-stress location in languages (specifically, the difference between right-oriented stress languages and right-edge stress languages).

As Sugisaki (2011) points out, Tokizaki’s analysis suggests that the core difference between English and Catalan is not due to a constructive parameter, but merely a result of “externalization”—a possibility that is gaining support in the literature concerned with cross-linguistic variation (witness Mathieu (2011), Richards (2010b)), and to which I return below. Similar considerations hold for the much-discussed difference between verb-framed and satellite-framed languages, first discussed by Talmy and often related to Snyder’s Compounding parameter, in light of proposals like Real Puigdollers (2011), who claims that the variation at issue depends on the morphological properties of the functional prepositional domain and not on the availability of a syntactic process of manner incorporation (see also Acedo-Matellán (2010) and Fasanella-Seligrat and Fortuny (2011) for proposals along similar lines). That the morpho-phonological properties of prepositions would be relevant in the context of Snyder’s Compounding Parameter is clear from the fact that Catalan-type languages insert prepositional material to support compounds.

In a vein very similar to Snyder’s assertion, Baker (2008a) contends that there are large-scale patterns in crosslinguistic variation that require “parameters within the statements of the general principles that shape natural language syntax.” Baker (2011) cites the following proposals, mostly taken from his own work, as illustrations of such syntactic parameters:

- The symmetrical object parameter (Baker, 1988)
Languages can have {1, more} “objects” (=structural Acc case)
- The Polysynthesis Parameter (Baker, 1996)
Languages {must, need not} express all theta-roles as morphemes on the verb
- The case dependence of agreement parameter (Baker, 2008b)
Functional heads {must, need not} assign case to an NP they agree with.

- The direction of agreement parameter (Baker, 2008b)
The goal of agreement {must, need not} c-command the agreeing head.
- Parameterization in the minimal link condition (Baker and Collins, 2006)
{The closest, any} NP can move into the Specifier of a functional head.
- Parameterization in the Case filter, whether NPs have case (Diercks, 2011)
NPs {are, are not} generated with an unvalued case feature.
- Parameterization in how case is assigned (Baker and Vinokurova, 2010)
Object case and agreement {are, are not} two sides of the same coin

I do not have the space here to go into the details of each of these proposals, and will limit myself to a few general remarks. The first one is that it is quite striking that many of the parameters listed above pertain to the morphological realization of case and agreement markers, whose syntactic status have been questioned (see, among others, Bobaljik (2008)), and can relatively straightforwardly be reformulated as PF/realizational parameters (i.e., spell-out rules). Second, some of the parameters proposed by Baker clearly exploit a lack of (one hopes, temporary) restrictiveness in syntactic theory. For example, Baker exploits several of the theories of case assignment of the literature, and claims that these various theories correspond to different parametric settings. But the intention of the proponents of such theories are clearly more universalist than this. Most of them would assume that in the fullness of time only one option for case assignment/agreement will be valid. The relevant syntactic parameters are therefore expected to go away. It is worth noting in this context that Baker is not alone in exploiting an unresolved theoretical ambiguity and turn it into a parameter. Lasnik (2000) did the same in trying to capture the well-known difference in verb placement differentiating English from French. As I wrote in Boeckx (In press), “[i]t is indeed quite remarkable to see that all too often it is only lack of understanding that leads one to claim that a certain property attributed to the language faculty is taken to be parametrizable. It is as if variation were the default.” But the recent history of syntactic theory leads us to expect that the better we motivate our syntactic constructs, the more they will turn out to be invariant—which is one of the main reasons for me to claim that narrow syntax is completely immune to variation, and that parameters of the sort Baker has formulated don’t hold.

The third remark I would like to make concerning the type of parameters Baker posits is that the one that is arguably the best worked out of all of them, his Polysynthesis Parameter (Baker, 1996) has been subject to strong criticism, and has been shown to be clearly inadequate. For instance, several authors have pointed out that languages traditionally characterized as polysynthetic do not display the full range

of properties predicted by Baker’s parameters. In part due to this fact, but also due to the fact that some of the main signatures of polysynthesis such as Noun Incorporation, vary so much across polysynthetic languages, some have argued in favor of a reformulation of the Polysynthesis parameter in terms of microparameters (see, e.g., Legate (2002), Adger et al. (2009)), which are much more easily recast as realizational rules (see, e.g., Mathieu and Barrie (2011)).

Last, but not least, it seems clear to me that the theoretical syntax community has (if only tacitly) decided that the reasonable success of accounting for principles like Relativized Minimality in terms of deeper design properties in recent years renders parametrizable versions of such principles unacceptable, even if such versions offer adequate descriptions of the data (though, even here, alternatives readily exist; *cf.* Jeong (2007) for a reanalysis of the facts that led Baker and Collins (2006) to propose their specific parameterization in the minimal link condition). This opinion, which I attribute to the theoretical syntax community, is well captured in the following passage (from van Riemsdijk (2008, 243f.)):

One of the main problems that we now face is the question of how the actual repercussions of such highly general principles of physical/biological organization in the grammar of specific languages can be insightfully represented. ...It would be absurd to propose that the constraint[s] [them]sel[ves] [are] parametrized.

Baker (2011) writes that “collecting a fuller range of good cases [of parameters such as those listed above] should help with the high-level theoretical work of discerning what can be a parameter and why.” But it seems to me that we already have enough of an understanding of what syntactic principles are to be confident that parametrization of the minimal link condition can be excluded. So, I strongly disagree with Baker when he says that “we should remain open to the possibility of deeper/more extreme parameterization, at least until we know more about crosslinguistic syntactic variation at the highest level.” There is, of course, always more to be learned, but I find it remarkable that Baker, who is on the one hand so confident about the robustness of syntactic principles as to be able to write (in support of macroparametric analysis, as opposed to microparametric analysis) that “it is already feasible to compare unrelated languages in an interesting way. This is possible because of the universal principles, which constrain crosslinguistic variation so that the dangers of incomparability and undiscernable interfering variation are not so dire”, can be so *unconfident* about their robustness as to entertain parametrized versions of them.

In sum, I think it is fair to conclude that the position advocated by Snyder and

Baker, a position very close to the original idea of parameter in generative grammar, is now a minority view, for essentially two reasons, one conceptual (or theoretical), and the other, empirical. At the conceptual, or theoretical, level, we have come to understand that (to borrow an observation of Kayne’s, formulated at a meeting in Barcelona in January 2010), “some properties of the language faculty are too deeply built in to be possible loci of variation.”⁸

It is occasionally suggested that the interest away from deep, large scale, higher-order parameters⁹ is due to “the fascination of microcomparative work (and single-language studies)”, which “can distract attention from [macroparameters] (because that can be really interesting too)” (Baker, 2011) (see also Baker (2008a) for reflections along similar lines). But this is not true. I think that the move away from deep parameters is due to the recognition that syntax is invariant.¹⁰ In this respect, I disagree with the statement in Baker (2011) according to which principles of syntax, unlike parameters, “have been vindicated, and are alive and well (details always evolving).” I think that our understanding of principles in syntax has radically changed over the past twenty years. From principles that were easily parametrizable, we have come to entertain ‘principles’ that are far more generic, and (I would contend) (therefore) unparametrizable.

At the empirical level, the position advocated by Snyder and Baker has lost traction mainly because the vast amount of work produced within the Principles-and-Parameters framework (certainly, its greatest virtue) quickly revealed that, to adapt a phrase of Sapir’s, all (macro-)parameters leak. The beauty of early parameters was that the number of grammatical consequences attached to a single on/off switch—what is known as ‘parametric clustering’—made it seem plausible that ten, twenty, or thirty switches at most would suffice to capture the grammars children

⁸*cf.* “Limiting syntactic parameters to features of functional heads is also intended to exclude the possibility that there could be a syntactic parameter that is a feature of no element of the lexicon at all, e.g. there could presumably not be a parameter of the sort ‘language L has or does not have bottom-to-top derivations’.” (Kayne, 2005)

⁹I prefer these adjectives to the notion of ‘macroparameter’, for I agree with Kayne (2005) that the micro/macro-parameter distinction is largely arbitrary. In Kayne’s terms,

It has occasionally been thought that the term parameter itself should only be used when there is such a notable or dramatic range of effects. I will not, however, pursue that way of thinking here. In part that is because what seems dramatic depends on expectations that may themselves be somewhat arbitrary.

¹⁰And perhaps also due to the recognition that large scale effects could be captured by the ‘alignment’ or conjunctive effect of microparameters; *cf.* “macroparametric differences might all turn out to dissolve into arrays of microparametric ones (i.e. into differences produced by the additive effects of some number of microparameters.” (Kayne, 2005)

acquire. As we know now, this estimate was now off by several degrees of magnitude. The need for a larger number of parameters is due in part to the fact that the strong and immediate implications that were part of early parameters were shown not to automatically follow, requiring as they seem to do, the existence of independent, more local parameters. The failure of clustering has been extensively discussed in the literature (see especially Newmeyer (2004, 2005)), so it is not necessary for me to illustrate it here.¹¹ I will come back, though, to one major consequence of the failure of clustering in the next section.

In light of what has been discussed in this section so far, it can be said that the majority view is that all parameters are lexical (*cf.* Kayne (2005, 10, 11), Rizzi (2009)). This view has been dubbed the “Borer-Chomsky conjecture” by Baker (2008a), who characterizes it as follows:

(1) *Borer-Chomsky conjecture*

All parameters of variation are attributable to differences in features of particular items (e.g. the functional heads) in the lexicon.

Although this Conjecture¹² is often touted as progress, I do not think that this is so, at least not until we know what is meant by the term ‘lexicon.’ Indeed it seems to me that the lexical parameter hypothesis was motivated more by empirical reasons (failure of Baker-style syntactic parameters to display all of their consequences) than by explanatory reasons, despite the repeated appeal in the literature to notions like ‘simplicity’, ‘restrictiveness’ and, in fact, ‘learnability considerations’ (“since children have to learn the lexicon anyway, . . .”).¹³ How else to account for the fact that what we wanted to understand (the nature of variation, the character of parameter) has been relegated to the part of the language organ that we understand the least: the

¹¹I cannot resist mentioning the recent and to my mind rather dramatic illustration of the failure of clustering provided by Garbacz (2011), who showed that even among closely related varieties, which are often said to provide an ideal testing ground (“controlled experiment”) for parametric hypotheses, clustering fails. In the case at hand, Garbacz (2011) discusses the parametric difference put forth by Anders Holmberg and Christer Platzack in a number of works, meant to capture salient differences among Insular and Main Scandinavian varieties. Garbacz (2011) shows that the proposed parameter makes the wrong prediction for a number of Swedish dialects. Importantly, as Newmeyer (2004) points out in his critique of parametric hypotheses, the problem noted by Garbacz (2011) for Holmberg and Platzack’s parameter is not one that can easily be fixed. Rather, it appears to be a ‘structural’ problem, inherent to the idea of parameter itself.

¹²I do not like to call the conjecture the “Borer-Chomsky” conjecture because I do not think that Borer actually endorsed it in the (1984) work that is routinely referred to in the literature. For discussion of this point, see Boeckx (2010).

¹³In this respect, see my (2010) discussion of Rizzi (2009, 2010).

lexicon, the part for which we have no theory?¹⁴

As a particular illustration of the point just made, consider the fact that many syntacticians would agree with me that parameters like Ken Hale’s “Configurationality Parameter”—according to which the phrase structure in non-configurational languages is not projected from the lexicon (i.e., Non-configurational languages are not subject to the what was then called the Projection Principle)—no longer fall within the realm of options they are willing to entertain, the sort of lexical parameters that they favor turn out to be far less lexical, and much more syntactic than they appear at first. Thus, many syntacticians in recent years (see, among many others, Giorgi and Pianesi (1997), Pytkänen (2008), Bobaljik and Thráinsson (1998), Fortuny (2008), Sigurdsson (2004), Zanuttini (2010), Savescu Ciucivara and Wood (2010)) have argued for parameters that take the following form:

(2) “*Bundling*” *Parameter*

Given two lexical features f_1 and f_2 , drawn from a universal repertoire (UG), does a given language L project f_1 and f_2 as a bundle or do f_1 and f_2 function as distinct heads in syntax?

Such “bundling” parameters account for a large number of parameters put forth in the literature under the Borer-Chomsky Conjecture.¹⁵ But what is the nature of this

¹⁴It is very curious indeed to see that proponents of the notion of ‘parameter’ have argued against Newmeyer’s (2004, 2005) suggestion to replace parameters by rules by stressing that we have no idea of what counts as a possible rule, when the notion of parameter they propose makes use of a component of the grammar that is equally underdeveloped theoretically speaking.

¹⁵As a matter of fact, the number of bundling parameters proposed in the literature increases exponentially if we take into account those lexical parameters that are in effect implicit bundling parameters. Indeed, many of the lexical parameters that focus on whether a given language L makes use of a given feature f boil down to a bundling issue, for in the vast majority of cases, it is only the presence of *the unvalued version of f* that is being parametrized. But “the unvalued version of f ” is nothing but a feature bundle (*cf.* Boeckx (2010/In progress)): $\{\{ \}, f\}$.

My claim is well illustrated in the following passage from Longobardi (2005b):

Though I agree with Benveniste (1971) that one cannot imagine a natural language where the meaning of person (i.e. the role of individuals talked about with respect to the speech act) is really ineffable, it is the case that some languages have been argued to be deprived of syntactic effects of ϕ -features altogether, including person (e.g. person agreement on predicates or anaphors), a case in point being e.g. Japanese

Likewise, Rizzi’s take on lexical parameters (*cf.* Rizzi (2009))—“A parameter is an instruction for a certain syntactic action expressed as a feature on a lexical item and made operative when the lexical item enters syntax as a head”—also reduces to a bundling parameter, given that instructions for syntactic actions, such as EPP features, Agree features, even Merge features, are nothing but features of features (i.e., feature bundles) (see, in addition to Boeckx (2010/In progress), Adger

lexical bundling operation? As far as I can see (it’s hard to say for sure because the nature of bundling is rarely made explicit in the literature), bundling is nothing but Merge operating ‘in the (pre-syntactic) lexicon’ (on this point, see Boeckx (2010/In progress)).¹⁶ Like Merge, “Bundle” operates on lexical features, it groups them into a syntactically combinable (mergeable) set; in fact, “Bundle” builds syntactic structures (feature trees; cf. Harley and Ritter (2002a,b)). So why do we take bundling parameters to be lexical when they pertain to the availability of particular *syntactic* constructs in a way that is no less radical than some of Baker’s “parameters within the statements of the general principles that shape natural language syntax”?¹⁷ In fact, bundling parameters are no less radical than Hale’s Configurationality parameters or other recent proposals (e.g., Huang’s claim that “highly analytic” languages like Modern Chinese lack l(exical)-syntax, or Reinhart and Siloni’s (2005) “Lexicon-Syntax parameter” according to which UG allows thematic arity operations to apply in the lexicon or in the syntax). *All* of these proposals presuppose an active lexicon, which is more than a mere list of (vocabulary) items; one that allows the application of derivational operations. In fact, bundling parameters turn the lexicon into a syntactic component, and by doing so, it effectively turns all lexical parameters into syntactic parameters, thereby nullifying the alleged explanatory superiority of

(2010), Adger and Svenonius (2011)).

In a similar vein, Kayne’s view on parameters articulated in Kayne (2005) boils down to “bundling”: Kayne writes (p. 15) that “UG imposes a maximum of one interpretable syntactic feature per lexical or functional element”, and since elsewhere (p. 11), he states that “every functional element made available by UG is associated with some syntactic parameter”, such a parameter must pertain to the ‘uninterpretable feature(s)’ (whose cardinality, interestingly, Kayne does not address) on the relevant lexical/functional item, in effect, it pertains to a feature of a feature and the way it bundles with the interpretable feature.

¹⁶As far as I can see, the ambition to reduce “Bundle” to Merge is also a central aspect of the “nano-syntax” approach sketched in Starke (2010), where heads entering the syntax are said to consist of single features. However, in terms of execution, I don’t think that nano-syntax achieves its goal, as it must ascribe more properties (i.e., features) to syntactic heads than just their interpretable features (e.g., selectional restrictions to capture the ordering of Merge, the “functional sequence”, which crucially nano-syntacticians take to be syntactic; as well as ‘probe’ features to capture the size of moving elements). The model I put forth in Boeckx (2010/In progress) avoids these consequences by taking lexical items to consist of a single property (the “edge feature”), with selectional and size restrictions being relegated to post-syntactic filters.

¹⁷If “Bundle” is not Merge, then, we have a division of labor between the syntax and the lexicon as generative engines along the lines of Reinhart and Siloni (2005). If “Bundle” is Merge, Bundling parameters are syntactic parameters. Either way, although they may appear innocuous, bundling parameters fall within the characterization of Snyder’s “constructive parameter” mentioned above, since they either add new ‘building blocks’ (bundles) or new structure-building operations (“Bundle”, or, as Gallego (2011a) suggests, an extra Spell-Out operation).

Borer’s insight of locating all parameters in the lexicon, and exposing such parameters to the same critique of syntactic parameters touched on above.

Given the range of considerations expressed in this paper so far, it seems to me that theoretically most desirable hypothesis boils down to the ‘Strong Uniformity Thesis’ (SUT) I formulated in Boeckx (2011a):

(3) *Strong Uniformity Thesis*

Principles of narrow syntax are not subject to parametrization; nor are they affected by lexical parameters

SUT is clearly in line with Chomsky’s (2001, 2) Uniformity Hypothesis (“In the absence of compelling evidence to the contrary, assume languages to be uniform”), but as we will see shortly, it intends to go one step further by closing the door to the possibility of an indirect parametrization of syntax through the elimination of pre-syntactic ‘lexical’ parameters such as “bundling” parameters. SUT also agrees with Berwick and Chomsky (2011),¹⁸ who state that “parametrization and diversity, then, would be mostly — possibly entirely — restricted to externalization.” (SUT in fact strengthens Chomsky and Berwick’s statement by eliminating the “mostly” and “possibly” of their formulation.

As I observed in Boeckx (2011a), under SUT, points of variation (‘parameters’, if one wishes to call them that, an issue I return to momentarily) would be confined to the margins of narrow syntax, especially the morpho-phonological component (PF). Let me make this clear in the following statement:

(4) *Locus of variation*

All ‘parameters’ reduce to realizational options
(i.e., PF decisions rendered necessary by the need to externalize structures constructed by an underspecified syntactic component¹⁹)

Let me make a few remarks concerning this strong statement. First, although it is certainly motivated by the desire to keep the syntactic component universal and invariant, a desire grounded in the strongest minimalist thesis of Chomsky (2004), it is not just a statement in service of “minimalism”. It is also grounded in empirical results, such as the gradual disintegration of (syntactic) macroparameters into (lexical) microparameters, and, especially, the recognition that the alleged syntactic

¹⁸See also Kandybowicz (2009) for reflections that point in the same direction.

¹⁹We do not yet have a complete typology of such rules, but works focusing on the articulation of the post-syntactic morphological component of the sort found in Distributed Morphology provides a good idea of what such realizational options amount to.

effects of these parameters can be (indeed, in many cases, have already been) recast in morpho-phonological terms: instead of saying that a given language moves certain elements (say, *wh*-words) overtly, we can equivalently say that the language PF-realizes the copy of those elements (say, in SpecCP); instead of talking about pre-syntactic lexical bundles, we can just as easily talk about post-syntactic morpho-phonological bundles, and so on and so forth.

Such a reformulation of parametric options has clearly been influenced by the development of ‘late’ (‘post-syntactic’) morphological models of the distributed (Halle and Marantz (1993)) and exoskeletal (Borer (2005)) kinds (to which we can add the more recent nanosyntax model; Starke (2010)).²⁰ All these models (especially when they are ‘radicalized’ along the lines I suggested in Boeckx (2010/In progress)) have the effect of freeing the syntactic component from following lexical instructions and from being subject to the potential (indeed unavoidable) parametrizability of these instructions.

The reformulation of parameters in terms of realizational properties also receives support from results in the syntactic literature that point to the irrelevance of featural specifications for syntax. The clearest illustration of this that I know of comes from the work of Martina Wiltschko (see Wiltschko (2009), who shows that although languages vary in the substantive content associated with functional categories (e.g., the functional category INFL, whose content, according to Wiltschko, can be ‘tense’, ‘location’ or ‘participant’), the syntax of these functional categories is completely invariant.²¹ This *empirical* result led Wiltschko to formulate her “Parametric Substantiation Hypothesis” according to which “Universal Grammar provides only abstract functional categories (without substantive content)” That is, at the initial state, “UG categories are associated with an abstract feature” (which Wiltschko takes to require valuation) and only later (in the course of acquisition) do these categories acquire their (language-)specific content. Wiltschko’s work points to the complete separation of formal and substantive aspects of lexical items, with formal (syntactic) aspects invariant, and substantive aspects parametrized, along the lines predicted by the idea that all parameters reduce to realizational options (especially if the substance of categories has no semantic effect, a possibility advocated by Harbour (2009), who

²⁰The copy theory of movement has also played an important role, as it allowed the exploration of the idea that a significant portion of logodiversity amounts to options of deletion or doubling of duplicates formed by movement/copying; see Barbiers (2009) for a systematic exploitation of this technical possibility.

²¹Some of the proposals in Miyagawa (2010) can be construed along similar lines, especially his suggestion that languages that are traditionally regarded as lacking agreement in effect substitute ϕ -feature agreement with more discourse-oriented feature—crucially retaining the formal agreement operation, only varying its featural content.

shows that what is required semantically is far more generic than the proliferation of functional flavors in the literature would lead us to expect.²²).

Finally, the realizational view on parameters also feeds on the growing evidence that much parametric lexical variation ‘merely’²³ “involve[s] the non-pronunciation of a lexical or functional item” (Kayne, 2005, 15-16).²⁴

The idea that logodiversity is confined to externalization strategies is, I think, very close in spirit to the language-particular rules of Newmeyer (2005), given that parameters in the classical sense are meant to be principles (parametrized principles) and a compelling case has been made for the absence of ‘principles’ in PF (Bromberger and Halle (1989)). Rules are often criticized for lacking a “format” (as if the format of parameters were clear), but I find this criticism rather weak: realizational rules will be whatever fills the space between the structures made available by the syntax and the content of the morphological component. ‘Caught’ in this way between syntax and phonology, realizational strategies will be akin to Kanizsa triangles. Rules are also often criticized because they would seem to lack the capacity to yield the cascade effects that early parameters were meant to have. But as we saw above, this may be a point in favor of rules, given that these cascade effects don’t seem to hold in a way predicted by the classical parametric model. Moreover, rules may still have a big impact on the overall profile of the language. As Baker (2008a) observes, “it is perfectly possible that a lexical parameter [equivalent to a local rule—CB] could have a substantial impact on the language generated, particularly if it is concerned with some very prominent item (such as the finite Tense node).”

One may, of course, continue to refer to the externalization strategies envisaged here as ‘parameters’ (it is, at bottom, nothing more but a terminological issue), but if one does so, one ought to make clear that there is a discontinuity between these

²²On distinct morphological categorial flavors reducing to the same semantic category, see, in addition to Harbour (2009), Schlenker (2006), Krifka (1992), and Boeckx (2010/In progress).

²³I put ‘merely’ in quotes for null elements are known to impose certain special ‘licensing’ conditions at PF (e.g., affixal properties), which may trigger the application of certain PF readjustment/repair rules that may well affect the overall grammatical, but crucially not syntactic, profile of the language. On a particularly compelling demonstration of cascade effects of ‘early’ PF decisions, see Bošković (2001); see also Acedo-Matellán (2010).

²⁴For Kayne, non-pronunciation still implicates syntax, as it is achieved by moving the relevant element to the edge of the phase (which Kayne says leads to unpronunciation, given that the edge of the phase is not transferred). I resist this view, as I do not think that moving to the edge of a phase ensures unpronunciation, for even if it’s true that the edge of the phase is not transferred to the external system along with the complement domain of the phase in Chomsky (2000), it is part of the spell-out domain in the subsequent cycle/phase, except perhaps in the very topmost, ‘root’ phase edge, although there exist arguments in favor of transferring the entire phase, edge included, in this case (see Boeckx (2010/In progress), building on Ott (2011); see also Obata (2010).

‘parameters’ and the parameters of old (contra Roberts and Holmberg (2009, 55)). One may even continue to call such externalization strategies “lexical parameters”, as long as by ‘lexical’ one means ‘post-syntactic’ morpho-phonological component (what Distributed Morphologists call “List B” or “the vocabulary”). Such a lexicon shares many properties of what Emonds (2000) calls the ‘syntacticon’; it would be a repertoire of structures (treelets)—Constructions, in fact,²⁵ whose degree of abstractness may vary, as Jackendoff (2005, 2010) has insightfully noted.²⁶

Given that the notion of the ‘lexicon’ touched on in the preceding paragraph is quite syntactic, one may even call the realizational strategies I have in mind ‘syntactic parameters’ as long as by syntax one means what Hale and Keyser (1993) called “l-syntax” (which they clearly and I think correctly distinguished from ‘s-syntax’, what I have called here ‘syntax’):

We have proposed that argument structure is a syntax, but we have also separated it from s-syntax, ... probably an onerous distinction, perhaps nothing more than a temporary terminological convenience. [BUT—CB] We must nevertheless assume that there is *something* lexical about any verbal/lexical entry. ... What is it that is lexical about the entry corresponding to *shelve*? Clearly, it is a lexical fact that *shelve* exists as a simple transitive verb in English. ... in reality all verbs are to some extent phrasal idioms, that is, syntactic structures that must be learned as the conventional “names” for various dynamic events.

In effect, Hale and Keyser are pointing out that their “l-syntax” is a syntax in the representational sense (a post-syntax, a morphology, in my terminology), whereas “s-syntax” is a syntax in the dynamic, derivational sense (narrow syntax, for me).²⁷ Despite the fact that Hale and Keyser were criticized for introducing what many regarded as an onerous distinction in the grammar, confusing l-syntax and s-syntax would be like confusing genetics and epigenetics. One is static, the other one dynamic. One relies on the other, but they are not to be collapsed.

Before concluding this section, I would like to address an interesting point made by Roberts (2010, 2011). While recognizing the existence of realizational ‘parameters’, Roberts thinks that it would be wrong to limit variation to the PF-component

²⁵Given the syntax-semantics transparency, the syntactic structures at issue would qualify as abstract Constructional idioms.

²⁶The view of the lexicon advocated here corresponds pretty closely to the one adopted in nanosyntax.

²⁷It could be said that while s-syntax is instructional, l-syntax is selectional, so my (l-)syntactic parameters would be selectional parameters, not instructional parameters of the sort advocated by Rizzi (2009).

of the grammar, as he sees no non-stipulative way to exclude syntactic parameters in current minimalist models, hence, he claims, such syntactic options for variation should be exploited as well.²⁸

Roberts is certainly right in claiming that in current minimalist models of syntax there is no “inherent ban on narrow syntactic variation”. As the following quote from Chomsky (2001, 2) makes clear,

Parametric variation is restricted to the lexicon, and insofar as syntactic computation is concerned, to a narrow category of morphological properties, primarily inflectional.

syntactic variation can arise in current minimalist models via the influence of pre-syntactic lexical parameters. However, this is true only because current minimalist models of narrow syntax are what I have elsewhere (Boeckx (2010/In progress)) called “lexiconcentric”: all properties of syntax are supported by a (hyper)active and far from minimal pre-syntactic lexicon. As soon as such a view of the pre-syntactic lexicon is abandoned, the ban on narrow-syntactic variation is not a stipulation, it automatically follows from the architecture of the grammar, and to the extent that realizational strategies can account for the observed variation, such an architecture is empirically supported.

To conclude, the rich body of work in comparative syntax over 30 years reveals quite a different picture from the one anticipated in the introductory chapter of Chomsky (1981). Variation is still parametric, in the sense of being limited (constrained), but such a limitation is not due to the existence of parameters (in the classical sense). Rather, variation is constrained because it lies at the intersection of two systems with strong demands of their own: syntax, whose structures have to be externalized, and morpho-phonology, which provides the toolkit for externalization. The intersection of these two systems yields a pool of variants, from which children learning their native languages select. Variation, in this context, boils down to learning which options of the universal syntax to pronounce (morphologize/lexicalize/idiomatize) and how. Parametric values are indeed constructional idioms (which narrow syntax constructs), and learning a language indeed reduces to constructing one’s own native (grammatical) vocabulary,²⁹ as Borer (1984) had anticipated, when she advocated the benefits of confining variation “to the one part

²⁸Part of Roberts’s argument in favor of the existence of syntactic, non-PF-parameters is that he takes realizational parameters to be “dumb” and to be unable to give rise to parametric hierarchies. I return to this matter in the next section, where I discuss the adequacy of such hierarchies, and shows that Roberts is wrong on this point.

²⁹I use the term ‘grammatical vocabulary’ to allude to the correctness of Jackendoff’s (1997, 2002, 2007, 2010) position that the dividing line between the lexicon and the grammar is artificial. It is

of a language which clearly must be learned anyway: the lexicon”. What scholars quoting Borer’s conjecture often miss is that she clearly saw that her conjecture meant that “we no longer have to assume that the data to which the child is exposed bear directly on universal principles”—in effect, she saw that her position entailed the demise of syntactic variation. What we have learned from comparative syntax is that classical parameters were not supposed to leak (to adapt Sapir’s well-known statement), but since all parameters turned out to leak, and syntactic principles don’t, parameters can’t be syntactic. Put another way, what we have learned from comparative syntax is that the more languages differ, the more the core syntactic component of the language faculty is invariant.

3 Parameters: In service of Plato’s problem or Greenberg’s problem?

Early proposals within the Principles-and-Parameters approach, such as Rizzi’s (1982) approach to null subjects, suggested to many that perhaps for the first time in the history of the language sciences, a solution to “Plato’s problem” (the logical problem of language acquisition) was within reach. This was due primarily to the fact that the setting of a parameter based on salient surface properties of the language to be acquired was to give rise to a host of seemingly unrelated and quite abstract properties of the language in question. Thus, once the child had determined that his language allowed referential null subjects of the standard Italian type, plausibly from frequent and salient data from the linguistic input the child received, Rizzi’s original formulation of the *pro*-drop parameter, insightfully reviewed in Roberts and Holmberg (2009), automatically entailed that no further learning was needed to account for the fact that Italian allowed expletive null subjects, constructions like free inversion (postverbal subjects), and subject *wh*-extraction from finite embedded clauses headed by an overt complementizer (of the sort that give rise to *that*-trace effects in English). This clustering of properties, expected to hold for all parameters, suggested to many that knowledge of the core properties of a particular language could be acquired on the basis of a rather limited number of parameters.

Unfortunately, subsequent empirical work in the domain of null subjects and indeed in all the areas covered by parameters revealed a much ‘messier’ landscape. In particular, they revealed a landscape where the deep ramifications of parameter settings were not always found outside of the language (family) that provided the

the dividing line between syntax and grammar/lexicon (or between s-syntax and l-syntax) that in my view must be not only maintained, but sharpened.

original motivation for the parameters in question. (For example, as Roberts and Holmberg (2009) discuss, free inversion and extraction facts originally entailed by Rizzi’s formulation of the *pro*-drop parameter were found not to always correlate with null subjects.) This naturally led to a weakening of the clustering claims. Instead of formulating such clustering effects as biconditionals (‘Property X iff Property Y’), comparative linguists provided one-way implicational statements (‘If Property X then possibly Property Y’).

As Roberts and Holmberg (2009, 22) observe, one-way implicational statements are easier to support empirically, and, given what one seems to find empirically, are preferable to stronger bidirectional statements to establish “parameter-based typologies”. Typological concerns occupy a growing portion of the literature on parameters (witness Baker (2010), Baker and McCloskey (2007), Kayne (2004, 2005), Cinque (1996, 2005, 2007)), which attempts to recapture some³⁰ of the universals uncovered in the tradition of Greenberg (1963).

While this work is undoubtedly valuable,³¹ my own feeling is that what I have elsewhere (Boeckx (2010)) called “Greenberg’s problem” (borrowing the phrase from Fasanella-Seligat (2009)) has detracted attention from what was, and, I think, ought to remain the focus of parametric proposals: Plato’s problem. Both problems are, of course, intellectually interesting, but as Newmeyer (2005) discussed extensively, the data sets that the two problems seek to cover are not the same, and in many cases, there is a clear tension between the two, as there often is between descriptive adequacy and explanatory adequacy.

For example, advertising the benefits of a given parameter by listing all its one-way implicational effects is frequently done with Greenberg’s problem, not Plato’s, in mind, for the comparative linguists doing this rarely worry about the nature of these implicational effects from the perspective of the child acquiring a language: one-way implications do no good for the child if the actual presence of these implications in the target language could not be established on the basis of the primary

³⁰As Kayne (2011) observes, the main concern is with “the [type of universals] that [Greenberg] put forth as being exceptionless”, not “the cross-linguistic generalization that he put forth as ‘(overwhelming or strong) tendencies’.” Kayne goes on to write: “As Hawkins (1983) in effect noted, the correct way to interpret these tendencies is to take them to be examples of possible cross-linguistic generalizations that have, however, sharp counterexamples. As in Hawkins work, one can try to reformulate one or another of these tendencies in such a way that the counterexamples disappear. Alternatively, the tendency in question was, in one or another case, simply a mistaken proposal.”

³¹I should note that I view the main value of this work to lie in the establishment of universal constraints (such and such correlation is never found in any language) more than in providing evidence for the one-way implicational statements that the literature on parameters focuses on.

linguistic data the child receives (not to be confused with the data set available by the linguist). (Recall that being one-way implications, their presence in the target language cannot be automatically determined by the setting of the parameter entailing these implications). As an example of this state of affairs, consider many of the one-way implications provided by the NP *vs.* DP parameter put forth by Bošković (2008). According to Bošković, languages that lack the D-layer in nominals may, unlike those that do make use of D, allow for certain sub-extraction patterns (absence of left branch condition effects, possibility of adjunct subextraction from nominals, etc.). But being one-way implications, the child must still rely on the primary linguistic data to determine if his target language indeed allows such extraction patterns. Given the likely rarity of the relevant examples in the linguistic input he receives, figuring this out is much harder than determining if the language makes use of the D-layer. Accordingly, from a learnability point of view (as opposed to a typological perspective), the one-way implications defended by Bošković do not bear on the adequacy of his parameter.³²

Likewise, one-way implications provide little evidence in favor of parameters if the grammatical consequences of the parameter could equally well be acquired independently of the parameter, as they may if they refer to very salient properties found in the primary linguistic data, as is the case for the consequences of the parameter discussed by Holmberg (2010)³³.

Again, only a typological concern provides the motivation for using these one-way implications as evidence for parameters. Clustering effects are only good (for Plato’s problem) insofar as the properties listed in these clusters are abstract (not immediately derivable from the primary linguistic data) and automatic consequences of more surfacy parameters, as was the case for the absence of *that*-trace effects in Italian-type null subject languages discussed by Rizzi.

Fasanella-Seligrat and Fortuny (2011) are correct in stressing that a good parametric model ought to “provide the elements that guide the learner in the process of interpreting the data it receives” to arrive at the knowledge state we call an I-language, and in pointing out that most extant proposals within the Principles-and-Parameters approach (especially, those appealing to macroparameters) fall far short of this goal. Fasanella-Seligrat and Fortuny (2011) list what they call the “locality problem”, given below, as a key problem for parametric proposals.

³²For additional arguments against Bošković (2008), and the severe problems the proposal faces, see Jeong (2011).

³³Lohndal and Uriagereka (2010) make a similar point in their commentary on Holmberg (2010).

(5) *The locality problem*

In order to fix the value of a macroparameter the learner should analyze the data it receives in a global and transverse way, since macroparameters are defined on highly general properties spread across the target language

To remedy this problem, Fasanella-Seligrat and Fortuny (2011) put forth two conditions (in the spirit of the “bootstrapping strategies” reviewed in Gervain and Mehler (2010)): the “atomicity condition” and the “accessibility condition”, which I reproduce here:

(6) *Atomicity condition*

Parameters must be atomic, they cannot be clusters of properties

(7) *Accessibility condition*

Parameters must be set by directly inspecting phonological and morphological properties of utterances

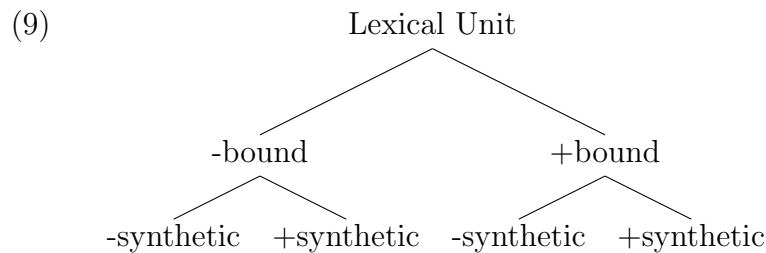
I think that Fasanella-Seligrat and Fortuny (2011) are right to urge parametric proposals to (re-)focus on Plato’s problem, although I don’t think that the locality problem *per se* is the issue. After all, many studies have called for the need to supplement (global and macro-)parameters with (local or micro-)cues to precisely get around the learnability problem identified by Fasanella-Seligrat and Fortuny (2011) (see Dresher (1999), Lightfoot (1999), Westergaard (2009); see also Yang (2002) on the ‘signatures’ of parameters, Clark and Roberts (1993) on ‘parameter expressions’, and Fodor (1998, 2001) on ‘treelets’ as ‘unambiguous triggers’). Nonetheless, insisting on conditions like the “accessibility condition” is certainly a step in the right direction, as it led Fasanella-Seligrat and Fortuny (2011) to propose two basic dimensions of variation which I would like to build on here. One of these dimensions is the bound/unbound distinction: a lexical unit³⁴ is bound if it is phonologically dependent on another lexical unit and unbound otherwise. The other dimension is the analytic/synthetic distinction, according to which a lexical unit is synthetic if it conveys more than one morpheme and non-synthetic (analytic) if it conveys only one morpheme.

³⁴Fasanella-Seligrat and Fortuny (2011) speak of “heads” as the minimal morphological categories, but it is not at all clear how the notion ‘head’ is not “abstract” enough to meet their “accessibility condition”. I will therefore use the more neutral term ‘lexical unit’ to refer to the objects that narrow syntax draws from the ‘narrow (pre-syntactic) lexicon. Lexical units in this sense could well refer to (i.e. have the size of) the conceptual addresses of Boeckx (2010/In progress) or the (nano-sized) features of Starke (2010), and indeed perhaps the “heads” of Fasanella-Seligrat and Fortuny (2011). This is an important empirical task that I have to leave for future work.

Fasanella-Seligrat and Fortuny (2011) argue that the two dimensions of variation just given yield a typology like the following:

- (8) Given a lexical unit U,
- a. Is it +bound or -bound?
 - (i) If +bound, is it +synthetic or -synthetic?
 - (ii) If -bound, is it +synthetic or -synthetic?

As is obvious, this typology has a certain symmetry to it, as is evident in the following representation:

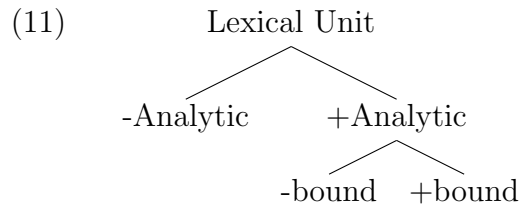


As such it conforms to Roberts’s (2011) characterization of PF-parameters as symmetric, which Roberts contrasts with syntactic parameters, which he takes to be asymmetric and therefore responsible for the typological gaps that the parameter-based studies in formal typology focuses on (Greenberg’s “exceptionless” statements, to use Kayne’s (2011) phrase).

However, upon closer examination, the typology that results from combining the two dimensions given by Fasanella-Seligrat and Fortuny (2011) yields an asymmetric pattern, for one cannot speak of a lexical unit as being expressed synthetically (+synthetic). This property can only be true of two lexical units. Accordingly, if the parametric dimensions are formulated for a single lexical unit, the typology should be rephrased as follows:

- (10) Given a lexical unit LU,
- a. Is it analytically (-synthetically) expressed?
 - (i) If yes, is it +bound or -bound

The following representation highlights the asymmetric character of the resulting typology:



The asymmetrical character of the representation appears to be better in line with the facts; e.g., regarding the morphological expression of [Path] units and the typology of verb-framed *vs.* satellite-framed languages in Acedo-Matellán (2010) or the morphological expression of ergative predicates in Hale and Keyser (1993) and Real Puigdollers (2011), where three language types are found, as opposed to four, which a symmetric typology would expect. (11) also shows that Roberts (2011) is wrong in taking “PF”/“purely realizational” parameters to be “dumber”[sic] than syntactic parameter and unable to form asymmetric hierarchies. Put differently, the existence of typological gaps are not to be taken as arguments in favor of syntactic parameters.

What is true is that (11) predicts that the hierarchical effects of realizational parameters will be limited, given the limited depth of (11). In other words, (11) is quite unlike the parameter hierarchy in Baker (2001) or the many parametric networks explored in Roberts (2011). I think that this is a positive result, for a variety of reasons. First, many of the ‘deep’ parameter hierarchies proposed in the literature are only apparently deep. As an illustration of this (one among many I could have chosen), consider the hierarchy proposed in Biberauer (2011):

- (12) Is F a formal feature in L?
- a. If the answer is no, it is only a semantic feature, but if yes,
 - (i) Is F unvalued?
 - (i) If yes, does F trigger movement of the Goal (= element valuing F)
 - (i) If yes, does the Goal alone move or is there pied-piping (movement of a unit containing the Goal)?

As interesting as Biberauer’s proposal is, it predicts many more ‘options’ than are logically available. For example, the choice of whether F is a formal feature in L does not in fact give rise to the subquestion “Is F unvalued?”, for if F is a formal feature, it must necessarily exist in an unvalued guise (as argued in Chomsky (1995, chap. 4), the valued/unvalued distinction is the distinguishing factor between a semantic feature and a syntactic feature). Likewise, the subquestion of what moves if movement takes place does not arise if pied-piping (*qua* movement of a unit larger

than the actual goal) does not exist, as several authors have argued, see Cable (2007), Narita (2010), Boeckx (2009))³⁵. If these questions don't arise, the hierarchy put forth by Biberauer reduces to the much shallower:³⁶

- (13) Is F a formal feature in L?
 a. does F trigger movement of the Goal (= element valuing F)

A second argument in favor of shallower parametric hierarchies comes from the line of work initiated by Longobardi (2004) (see also Longobardi (2005a), Gianollo et al. (In press), Longobardi and Guardiano (2011)). Longobardi and his colleagues show that (what they call) “syntactic” parameters (which I suspect reduce to realizational properties, although they are not formulated as such by Longobardi, but see Sheehan and Hinzen (2011) for a realizational restatement of Longobardi (1994)) can be used to infer phylogenetic—or, better said (Fitch (2008), glossogenetic—relations among languages. One of the lessons coming out of the work is that “syntax” changes considerably more slowly than the vocabulary. As Longobardi concedes (personal communication), this strongly suggests that cascade effects resulting from resetting of a (‘syntactic’) parameter are quite limited. This would be unexpected if parameter hierarchies were as deep as Baker or Roberts posit, but it would follow quite straightforwardly from the shallower, micro-parametric view advocated here.

Needless to say, the range of consequences of a given parameter value may be quite large, but this need not entail that the structure of the parameter is as rich as Baker or Roberts would have it. As Baker (2008a) himself has noted, “it is perfectly possible that a lexical parameter [equivalent to a local rule—CB] could have a substantial impact on the language generated, particularly if it is concerned with some very prominent item (such as the finite Tense node)” (witness, e.g., the range of consequences of the absence of overt definite determiners discussed in Bošković (2008)), but the limited amount of deductive structure of realizational parameters may in fact have more than an empirical benefit (descriptive adequacy), it may in fact prevent the child from deviating too much from the target language in case of a mis-set parameter value (explanatory adequacy), and may in fact be a generic, “third factor”-type design property of well-behaved networks, more generally (beyond explanatory adequacy)—a point I come back to below in the context of Kauffman’s

³⁵These authors argue that nothing bigger than Goal ever moves.

³⁶The two parameters are clearly realizational, as the decision of whether F is formal or not depends on visible properties of utterances such as agreement [morphological duplication] or movement [pronunciation of higher copy].

The revised hierarchy in the text is very reminiscent of the parameter schema I put forth in Boeckx (2010): (i) Does F have an unvalued variant (*uF*)?; and (ii) which phase head H bears *uF*?

(1993) Boolean networks and his notion of “order for free”.

Before doing this, though, I would like to examine more closely some additional properties of the parameter hierarchies (networks) put forth by Roberts (2011). Pursuing an idea very similar to the one independently put forth in Boeckx (2011a), Roberts has explored the possibility that certain macro-patterns of variation, those that microparameters alone seem unable to capture, such as the alignment of most heads to yield a head-final language like Japanese, may be the result of the aggregation of microparameters “acting in concert for markedness reasons” (Roberts and Holmberg, 2009, 41). The basic idea behind this statement is the language faculty favors the alignment of parametric values for related parameters (e.g., cross-categorical alignment). This is captured by the “Superset” (learning) bias in Boeckx (2011a, 217) and by the “Generalization from the input” in Roberts and Holmberg (2009, 41):³⁷

- (14) *Superset bias*
Strive for parametric-value consistency among similar parameters
- (15) *Generalization from the input*
If acquirers assign a marked value to H, they will assign the same value to all comparable heads

There are good reasons to explore the possibility of such learning tendencies. Empirically, many so-called macroparameters clearly amount to the alignment of parametric values across heads. Thus, Branigan (2011) defends the existence of macroparameters on the basis of a “Head-attraction parameter” “*for all functional heads*” for Algonquian, requiring cross-categorical multiple attraction. Whitman (2008) also points to the tendency to generalize parameter settings across heads. Likewise, many of Baker’s macroparameters listed in section 2 amount to cross-categorical generalizations (*cf.* his “Case dependence of agreement parameter” (Baker, 2008b), which says that *all* functional heads {must, need not} assign case to an NP they agree with, or his “direction of agreement parameter” (Baker, 2008b), which says that for *all* agreeing heads, the goal of agreement {must, need not} c-command the agreeing head.) That such a tendency is formulated as such (as a tendency or bias) leaves the door

³⁷The similarity between the two conditions is even more obvious given the formulation of the Input Generalization in Roberts (2011):

For a given set of features F and a given set of functional heads H , given a trigger for feature $f \in F$ of a functional head $h \in H$, the learning device generalizes f to the maximum quantity of other functional heads $h_i, \dots h_n \in H$.

open for leakages (as I wrote above, all macroparameters leak, for every value in language L , there is always a residue of the opposite value in L ; see, e.g., the “synthetic residue” in “(highly) analytic” languages discussed in Cyrino and Reintges (2011)). A bias, unlike a parameter, allows for exceptions (Boeckx (2011a) claims that this tendency to generalize across heads is abandoned only if the child faces a significant amount of positive evidence going in the opposite direction).

The Superset bias or the Generalization from the Input also have the advantage of predicting more macro-patterns of variation than a purely microparametric approach would. As Roberts (2011) points out in the context of Kayne’s (2005) remark that a big parameter space (as big as 2^{100} , or even 2^{400} ; cf. Cinque and Rizzi (2010)) “is no problem (except, perhaps, for those who think that linguists must study every possible language)”, the microparametric space is perhaps too big, for if parameters are indeed as independent as a strictly microparametric approach would have it, languages “should appear to vary unpredictably and without assignable limits.” Roberts is right in noting that this is not what we seem to find.³⁸ Although we have learned that there are many more subtle ways in which languages can vary than we may have thought at the beginning of the Principles-and-Parameters era, it is still the case that we find recurrent patterns of variation, though not as rigid as a purely macroparametric approaches predict.³⁹ The learning biases listed above intend to capture precisely this.

Although the Superset bias and the Generalization from the Input appear to make similar predictions, my vision departs from Roberts’s in a number of respects. Both agree that macroparameters amount to aggregates of microparameters, but while I take the aggregation of microparameters to result in the emergence of parametric hierarchies (and learning paths) in a bottom up fashion, Roberts does not seem to, as he writes (Roberts and Holmberg (2009)):

Acquirers, and therefore languages, favor a high position in the network. For general reasons of computational conservativity, the acquirer is always trying to stop, as it were [i.e, stay at the top of the hierarchy]. Acquirers only move down the tree if forced to by the P[rimary]L[inguistic]D[ata].

³⁸Roberts (2001, 90) calculates that under certain reasonable assumptions concerning the number of speakers and human generations, and given a grammatical space of 2^{30} , it would take 18,000 centuries for each language type to be realized once! (bear in mind that *Homo sapiens* as only been around for 2000 centuries ...).

³⁹Baker (2008a) addresses this issue and claims that pure types would emerge from the data if it weren’t for the fact that microparameters obscure them. I’m not sure I follow Baker’s reasoning, as macroparametric proposals predict that once parameters values have been set, there is no room left for microparameters to, say, exempt certain heads from the value of the macroparameter.

I find this statement puzzling for it seems to go back to a macroparametric view by reifying the hierarchies outside of the learner, as it were. If I understand the idea behind the Generalization from the input, which I take to be the same as the one behind the Superset bias, it cannot be the case that acquirers ever “move down the tree [parameter hierarchy]”, for it is acquirers that build the hierarchy bottom up (i.e., going up!) by generalizing across heads. Talking about going down the hierarchy only makes sense in a typological context (Greenberg’s problem), not in developmental context (Plato’s problem).⁴⁰

Another difference between my construal of the Superset bias and Roberts’s Generalization from the input—to be fair, this is only a potential difference for the formulation of the hierarchies is still work in progress—is that Roberts gives the impression (apparent in Holmberg and Roberts (2010)) that he thinks that the hierarchies will combine in a way that will approximate Baker’s (Baker (2001)) single-rooted hierarchy, whereas I think that the hierarchies modeled along the lines of (11) will form modules in a network that will resemble a subway map, with hubs or hotspots as basins of attraction.

In the remainder of this section I’d like to make a few more remarks concerning the picture that seems to be emerging from what I have said so far regarding parameters and the Superset bias.

The first remark concerns macro-parametric phenomena. The working of the Superset bias envisaged here is strongly reminiscent of Yang’s (2005) on the acquisition and productivity of morphological rules.⁴¹ According to Yang, “[w]hen learning starts, there are no [specific] rules. That is, the learner constructs rules on an item-by-item basis” (i.e, morphological rules are inductively determined). Equipped with the toolkit of Universal Grammar, the child “seeks to extract generalizations [from] the learning data” *bottom-up*; attempting to broaden the scope of the rule as much as possible—i.e., making the rule more abstract.⁴² As for the productivity of the rules so induced, it boils down to “the cost of exceptions (complexity)”: assuming that the name of the game is to “minimize the structural description of the data”, the learner will “tolerate exceptions to a rule, until the size of the exception list outweighs the benefits of the rule”, in which case, the learner will extract a new rule covering what was until then ‘exceptional’ data (see also Feldman et al. (2000) for the development

⁴⁰Additional problems for Roberts’s hierarchies as learning paths and/or typologies, see Fasanella-Seligrat and Fortuny (2011) and Walkden (2011)

⁴¹If I am correct in the establishment of this parallelism, this could be another point in favor of Newmeyer’s (2004, 2005) argument in favor of rules.

⁴²What will emerge is a continuum of rules with rules progressively containing more variables, along the lines of Jackendoff’s (2005) continuum from rigid idioms to the \bar{X} -schema.

of a similar idea).

Yang’s model gives rise to a system of rules (a grammar) that is in some sense an optimal compromise between two competing (cognitive) forces: drive to generalize over items (akin to the Superset Bias) and the drive to be conservative and avoid overgeneralizations that may be hard to retract from (akin to the Subset principle of Berwick (1985), to which I will come back). Each rule will be a local Goldilocks solution, a middle ground (minimax)⁴³ solution, halfway between a prototype and an exemplar.

Such a situation is very well-known in cognitive science, especially given the renewed interest in “Bayesian learning” approaches, where the discussion often revolves around what Briscoe and Feldman (2006) has called the “bias/variance tradeoff”. As a matter of fact, the picture emerging from my appeal to the Superset bias, and the invariant, impoverished (underspecified) priors of narrow syntax fits rather well with the framework of Kemp et al. (2007), where a (hierarchical) Bayesian learning model is proposed for what they call “overhypotheses” (which correspond pretty closely to macroparameters in the sense used here). Tenenbaum and colleagues show how “knowledge can be simultaneously acquired at multiple level of abstraction, and may help to reconcile two competing approaches to cognitive development: the bottom-up approach [Superset bias] [and] the top-down approach [*cf.* Roberts’s remarks about going down the parametric hierarchy during learning].” The hierarchical Bayesian approach provides a unifying framework that accommodates both top-down and bottom-up learning by starting with very impoverished priors (“so minimal as to be uncontroversial”) and letting both the very local (micro) and the generalized (macro) version of the rules compete, with the winning version of the rule emerging from the data. In a certain sense, this situation brings us back to some version of the evaluation metric of Chomsky (1965), Chomsky and Halle (1968).

The growing literature on Bayesian learning certainly shows how impoverished priors (what Culbertson (2010)⁴⁴ calls “substantive biases”) and even weak biases (of the “regularization kind”, *cf.* Culbertson (2010)) can be powerful enough to reproduce data profiles that previously would have seemed impossible to acquire without richly specified innate knowledge.⁴⁵

⁴³Both Chomsky and Piattelli-Palmarini speculated in Piattelli-Palmarini et al. (2009) that parameters may be ‘minimax’ solutions to the acquisition problem, but I think they had something else in mind from what I am discussing in the text, as both seem to think of a tradeoff between what is learned and what is innate, whereas I am talking about an optimal format for a learned rule.

⁴⁴Thanks to Aritz Irurtzun for drawing my attention to this work.

⁴⁵A very clear example of the strength of this approach comes from Baronchelli et al. (2010, 2011), where weak biases can be shown to counteract the ‘infinite diversification’ tendencies of cultural

The second remark I’d like to make in the context of the Superset bias is that it indeed appears to conflict with the Subset principle that is frequently cited in the generative literature on language acquisition.⁴⁶ According to the Subset principle proposed by Berwick (1985), the learner must guess the smallest possible language compatible with the input at each stage of the learning procedure to avoid any backtracking.

Recently, Snyder (2007, In press) has brought the subset principle back to the fore by showing that in the context of spontaneous speech (i.e., outside of an elicited production experimental setting), children rarely (if ever) make errors of “commission” (the vast majority of errors are errors of omission). That is, children produce utterances that everyone would regard as part of the grammar of their target languages. As Snyder notes, this appears to conflict with claims by various researchers on language acquisition (S. Crain, C. Yang, etc.) according to which children during the course of language acquisition entertain parameter values that are not those of their target languages. If they indeed did so, Snyder points out, how come they make so few commission errors?

Snyder’s answer is what he calls the Grammatical Conservatism thesis:

(16) *Grammatical Conservatism thesis*

Children do not make productive, spontaneous use of a new syntactic structure until they have both determined that the structure is permitted in the adult language, and identified the adults grammatical basis for it.

As Snyder observes, Grammatical Conservatism points to the sort of deterministic learning that the Subset principle was intended to capture (no backtracking in the course of acquisition).⁴⁷ It should be obvious that this view conflicts with the Superset bias, but not, I think, in a way that is irreconcilable. After all, Grammatical conservatism is a thesis about the child’s production, about which the Superset bias is silent. As far as I can tell, nothing in the data that Snyder uses to motivate Grammatical Conservatism indicates that the child does not entertain target-inconsistent options in his head. For all we know, Grammatical Conservatism could be a prag-

norms to yield universals in the domain of color naming. It seems to me that a similar approach could recapture Baker’s (2008a) “bimodal distribution” of languages, with languages tending to cluster around one type or another, with a certain amount of noise and a few outliers from either one of the principal patterns.

⁴⁶Thanks to Aritz Irurtzun for making me think about this issue.

⁴⁷Snyder also points out that Grammatical Conservatism suggests that there are no substantive defaults, in the sense of an unmarked option that could be incorrect in some languages (*cf.* Sugisaki (2007)). Hence it goes against the view sometimes voiced by M. Baker or I. Roberts that macroparametric settings are defaults.

matic principle that the child adopts. Moreover, the scope of Snyder’s Grammatical Conservatism thesis may turn out to be quite limited indeed if Wexler (1998, 25) is correct in claiming that “Basic parameters are set correctly at the earliest observable stages, that is, at least from the time that the child enters the two-word stage, around 18 months of age.” If that is so, the child’s production would take place after many parameters have been set (correctly), that is, after the action of the Superset bias.

If the Superset bias is on the right track, it certainly qualifies as a “third factor” principle in the sense of Chomsky (2005).⁴⁸ Chomsky (2005, 6) defines third factor principles as “fall[ing] into several subtypes”: “(a) principles of data analysis that might be used in language acquisition and other domains; (b) principles of structural architecture and developmental constraints that enter into canalization, organic form, and action over a wide range, including principles of efficient computation, which would be expected to be of particular significance for computational systems such as language.” The Superset bias is a regularization bias in the sense of Culbertson (2010). It is the sort of “force” whose action is responsible for what is known in the Complex Systems literature as “collective behavior” (or “swarm intelligence”).⁴⁹ It is a principle of (self-)organization that does not require pre-specification to give rise to macroscopic patterns. In the words of Stuart Kauffman, it gives rise to “order for free”.

In this context, I find it intriguing that the parameter schema of limited depth in (11), as opposed to other schemas with further levels of embedding found in the literature (see, e.g., Longobardi (2005a) or the parameter hierarchies proposed by Baker (2001) and Roberts (2011)), is—if we take the general conclusions of Kauffman (1993) seriously—of exactly the right (‘optimal’) size to form a network of connections that gives rise to a system that is neither static (immune to change) nor chaotic. It has the sort of Goldilocks size that, according to Kauffman, many biological systems exhibit, and that maintains it at the edge of chaos and gives it exquisite flexibility.

⁴⁸See Mobbs (2008) for a similar view on the “Generalization from the input” principle.

⁴⁹Here is how the Wikipedia entry on “swarm intelligence” defines the latter:

S[warm] I[n]telligence systems are typically made up of a population of simple agents or boids interacting locally with one another and with their environment. The inspiration often comes from nature, especially biological systems. The agents follow very simple rules, and although there is no centralized control structure dictating how individual agents should behave, local, and to a certain degree random, interactions between such agents lead to the emergence of “intelligent” global behavior, unknown to the individual agents. Natural examples of SI include ant colonies, bird flocking, animal herding, bacterial growth, and fish schooling.

Kauffman (1993) has studied the properties of Boolean networks extensively and has reached several conclusions that are of great interest to us in the context of logodiversity and that I would like to discuss briefly here.⁵⁰

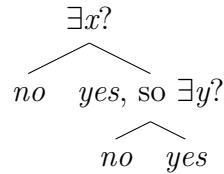
Boolean networks are basically networks of switches which can occupy a certain number of possible states corresponding to the states of each switch. (As Roberts (2001) observes, in the context of language, each possible state would correspond to a grammatical system.) Kauffman has studied systems where the switches are binary and are connected to other switches in Boolean relations (hence the name ‘Boolean network’ that he has given to them). Kauffman has shown that it is possible to restrict what he calls the “state cycle”⁵¹ of such networks by limiting the number of Boolean connections. To do this, Kauffman relies on a certain quantity that he calls K , which characterizes the number of inputs determining the value of each variable in the system (i.e., the degree of the connections in a Boolean network). According to Kauffman, if $K = 1$, that is, if each variable has just a single input, the system is static or frozen. If K is superior to 2, the system is chaotic and in Kauffman’s terms, ‘fluctuates’ wildly. But if K is exactly 2, the system is neither frozen nor chaotic. Kauffman also shows that if K is 2, the size of the state cycle is roughly the square root of n , the number of binary variables. (So, in a 100-parameter system, with $K = 2$, there would be only 10 possible states (language types).)

The parameter schema in (11), repeated here in a schematic form for convenience, is of exactly the right size to avoid freezing or chaos, for each of the tree values (terminal nodes) is a Boolean function of the other two (any system that chooses one value cannot choose either of the other two values:

⁵⁰I am not the first to note the relevance of Kauffman’s work in the context of language variation. Roberts (2001) explicitly discusses Kauffman’s work, and in fact reaches conclusions very similar to the ones I reach in the text, although he appears to have set such conclusions aside in subsequent work.

⁵¹The notion of state cycle in Kauffman’s work arises from the following reflection (here I am following Roberts (2001), who is in turn paraphrasing Kauffman): the system characterized by Boolean connections can be in a finite number of states. If started in one state, the system will over time flow through some sequence of states, a sequence which Kauffman calls a “trajectory”. Since there is only a finite number of states (the system is closed), the system will eventually hit a state that it has been in before. Then the trajectory will repeat, and since the system is deterministic, it will cycle forever around a recurrent loop of states that Kauffman calls the “state cycle”.

(17) *Right “molecule” of variation*



Crucially, if the deductive structure of the parameter schema were any deeper, the system would be completely chaotic, and variation would appear to be more random than it is. We can therefore conclude that the system that emerges is well-behaved, or well designed, even if it lacks the rich specifications of typical Principles-and-Parameters proposals.

4 Conclusion: A qualitative improvement

Kayne (2011) is surely correct in saying that our understand of linguistic variation and of the human language faculty have qualitatively improved over the past 30 years, in no small part due to efforts in the area of comparative syntax. In fact, I think that our understanding has gone further than he seems to think. We have come a long way from the first rigorous attempts to identify the limits of variation, from the concept of universal principles and finite, discrete parameters, “strategically-placed” and “small in number”, as Baker (2011) characterizes them. Such classical Parameters (to be kept separate from *parameters*) and their potential benefits for language acquisition now look to many to have been a mirage—a useful one, to be sure, but a mirage all the same.

Although this conclusion is still resisted in some comparative linguistics circles, certain signs don’t lie. Much like the phlogiston in chemistry, the notion of parameter in modern theoretical linguistics is collapsing under the weight of empirical evidence and conceptual arguments. But many saw this news coming. Indeed, I find it interesting that some ten years ago a major advocate of parameters like Ian Roberts could write the following:

...work by Clark (1994, 1996) has shown that plausible assumptions about the learner impose heavy restrictions on the format for parameters, and in fact are compatible with the abandonment of this notion in favor of indiction over tree fragments [now more often referred to as ‘treelets’—CB]. We can thus think that the existence of parameters themselves is attributable to properties of the learning algorithm interacting with UG; parametric variation might then be an emergent property of

this interaction. UG itself would then contain no “statement” of parameters at all. I will not explore this view in detail here, but it is compatible with everything that I have to say; it is also in the spirit of a minimalist linguistic theory (Roberts, 2001, 84)

Likewise, in the earliest stages of the minimalist program, Uriagereka (1994)⁵² wrote that “[t]he notion ‘parameter’ may be a useful “notational device” but we should not “confus[e] the notion with an ontological claim” (p.247). Instead, “parameters may in principle be states of the language faculty, not primitive notions” (p.249) Much like Roberts a few years later, Uriagereka wrote that “UG contains no internal parameters.” (p.252)

Aware of the difficulties faced by parametric theories in the context of Plato’s problem (or its cousin, “Darwin’s problem”, i.e. the logical problem of language evolution; *cf.* Boeckx (2011b)), proponents of new parameters have sought to motivate them by appealing to “Greenberg’s problem”, arguing that parameters offer us a better, more formal way of doing typology. The problem here is that the typological notions that parameters are intended to capture (“analyticity”, “polysynthesis”, and the like) have long been regarded as inadequate by typologists themselves (see my comments in Boeckx (2010) on this issues, with relevant quotes from M. Haspelmath and others; see also Moreno Cabrera (2000)). In the context of the study of the language faculty, I think we would do well to follow Ernst Mayr’s advice and recognize that biology succeeded once typological thinking was gotten rid of.

Biology offers us other lessons that bear on the notion of ‘parameter’. Much like genocentrism has been criticized for proposing a single-level theory where explanations are likely to require the conspiracy of many factors (see, in this respect, especially Fodor and Piattelli-Palmarini (2010, Part II)), parametric approaches have suffered from the desire to treat every difference as a parameter (a view attributed to Kayne and Manzini, among others, in Smith and Law (2009)). Biology also learned to revive developmental concerns that have been all too absent from studies that in theory at least are about the logical problem of language acquisition (on this point, see Longa and Lorenzo (2008), Lorenzo and Longa (2009)).

To some, it may seem that replacing syntactic parameters with realizational options, as advocated in this paper, may be nothing more than terminology (ditto for Newmeyer’s (2004, 2005) appeal to rules over parameters): it’s six of one and half a dozen of the other. But I would disagree with this view. The system proposed here is still parametric in important respects (variation is finite and limited), but the

⁵²Thanks to Marc Richards for bringing this paper to my attention after I wrote Boeckx (2010, 2011a).

elimination of parameters in syntax makes it not only possible but also plausible to pursue certain questions concerning third factor effects that would be significantly harder to motivate if syntax were parametrized. Inquiry into third factor effects came to prominence in the context of minimalism, but like the minimalist program itself, such concerns extend well beyond narrowly Chomskyan circles and have figured prominently (though perhaps under different names such as ‘biases’) in the non-generative tradition (see, e.g., the work of Simon Kirby and colleagues, beginning with Kirby (1999)). This tradition is bound to be of utmost importance if we are to understand why options left open by an underspecified UG get fixed during the course of language acquisition. As Richards (2010a) observes, this “fixing problem” can no longer be solved by appealing to innate requirements (“parameters must be set”), and is likely to receive a solution in terms of learning and cultural norms.

The unavoidable appeal to learning biases once parameters are removed from UG also has the desirable effect of reviving theoretical concerns for Plato’s problem and the topic of learnability, whose influence had undoubtedly diminished as the parametric model matured.

Needless to say, none of the big questions have yet been solved, but the shift away from classical parameters and the switchboard metaphor and towards a more ‘pointillist painting’ conception of variation, one that takes variation to be patterned, but not pre-patterned, and that hopes for the emergence of macropatterns “for free” is likely to tell us a lot about the character of our language faculty.

True, throughout this paper I have sort of presupposed that we have learned enough about the shape of the linguistic ‘morphospace’ to discard certain analytic options and promote others, and one must always bear in mind that perhaps “collecting a fuller range of good cases [candidate macroparameters]” (Baker, 2011) is still necessary (though, I doubt it⁵³), I find it remarkable that there is a consensus emerging (even within the ‘parametric’ literature) that variation exists (and takes the form that it does) because (in the words of Marc Richards) “UG doesn’t mind”. If UG doesn’t mind (and perhaps doesn’t care) about variation, it strikes me as counterproductive to insist on building in the limits of variation into it by brute force.⁵⁴ If UG doesn’t mind the absence of specification in the domain of variation, perhaps linguists that care about UG shouldn’t mind either.

⁵³And at any rate, aiming for explanatory adequacy (or even aiming to go beyond it, need not wait full satisfaction with the degree of descriptive adequacy attained (*cf.* Chomsky (1965), Kayne (2011)).

⁵⁴This includes appeal to the Baldwin effect of the sort proposed by Smith and Law (2009), which would predict (wrongly) that UG does mind after all, quite apart from the fact that there does not seem to have been enough time to achieve the desired results; on this, see Chater et al. (2009).

References

- Acedo-Matellán, V. 2010. Argument Structure and the Syntax-Morphology Interface. A Case Study in Latin and other Languages. Doctoral Dissertation, Universitat de Barcelona.
- Adger, D. 2010. A minimalist theory of feature structure. In *Features. Perspectives on a key notion in linguistics*, ed. A. Kibort and G. Corbett, 185–218. Oxford: Oxford University Press.
- Adger, D., D. Harbour, and L. J. Watkins. 2009. *Mirrors and macroparameters: Phrase structure beyond Free Word Order*. Cambridge: Cambridge University Press.
- Adger, D., and P. Svenonius. 2011. Features in Minimalist Syntax. In *The Oxford Handbook of Linguistic Minimalism*, ed. C. Boeckx, 27–51. Oxford: Oxford University Press.
- Baker, M. 1988. *Incorporation: a theory of grammatical function changing*. University of Chicago Press.
- Baker, M. 1996. *The Polysynthesis Parameter*. Oxford: Oxford University Press.
- Baker, M. 2001. *The Atoms of Language*. New York: Basic Books.
- Baker, M. 2008a. The macroparameter in a microparametric world. In *The limits of variation*, ed. T. Biberauer, 351–373. Amsterdam: John Benjamins.
- Baker, M. 2008b. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.
- Baker, M. 2010. Formal generative typology. In *Oxford Handbook of Linguistic Analysis*, ed. B. Heine and H. Narrog, 285–310. Oxford: Oxford University Press.
- Baker, M. 2011. Principles and Parameters Set Out from Europe. Presented at 50 Years of Linguistics at MIT.
- Baker, M., and C. Collins. 2006. Linkers and the Internal Structure of vP. *Natural Language & Linguistic Theory* 24:307–354.
- Baker, M., and J. McCloskey. 2007. On the relationship of typology to theoretical syntax. *Linguistic Typology* 11:285–296.
- Baker, M., and N. Vinokurova. 2010. Two modalities of case assignment in sakha. *Natural Language & Linguistic Theory* 28:593–642.
- Barbiers, S. 2009. Locus and limits of syntactic microvariation. *Lingua* 119:1607–1623.
- Baronchelli, A., T. Gong, A. Puglisi, and V. Loreto. 2010. Modeling the emergence of universality in color naming patterns. *Proceedings of the National Academy of Sciences* 107:2403–2407.
- Baronchelli, A., V. Loreto, and A. Puglisi. 2011. Cognitive biases and language

- universals. Ms.
- Berwick, R.C. 1985. *The acquisition of syntactic knowledge*. Cambridge, Mass.: MIT Press.
- Berwick, R.C., and N. Chomsky. 2011. The biolinguistic program: The current state of its evolution and development. In *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, ed. A. M. Di Sciullo and C. Boeckx, 19–41. Oxford: Oxford University Press.
- Biberauer, T. 2011. In defense of lexico-centric parametric variation: Two 3rd Factor-constrained case studies. Presented at the workshop on linguistic variation and the minimalist program, Universidad Autonoma de Madrid.
- Bobaljik, J. D, and H. Thráinsson. 1998. Two heads aren’t always better than one. *Syntax* 1:37–71.
- Bobaljik, J.D. 2008. Where’s Phi? Agreement as a post-syntactic operation. In *Phi theory: phi-features across modules and interfaces*, ed. D. Harbour, D. Adger, and S. Béjar, 295–328. Oxford: Oxford University Press.
- Boeckx, C. 2006. *Linguistic Minimalism: Origins, Concepts, Methods, and Aims*. Oxford: Oxford University Press.
- Boeckx, C. 2009. The locus of asymmetry in UG. *Catalan Journal of Linguistics* 8:41–53.
- Boeckx, C. 2010. What Principles & Parameters Got Wrong. Ms., ICREA & Universitat Autònoma de Barcelona. [Written version of Talk given at the Workshop on Variation in the Minimalist Framework, Barcelona, January 2010.].
- Boeckx, C. 2010/In progress. Elementary syntactic structures. Ms., ICREA–UAB. [Part A, “Defeating lexiconcentrism” available at <http://ling.auf.net/lingBuzz/001130>.].
- Boeckx, C. 2011a. Approaching parameters from below. In *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, ed. A.-M. Di Sciullo and C. Boeckx, 205–221. Oxford: Oxford University Press.
- Boeckx, C. 2011b. Some reflections on Darwin’s Problem in the context of Cartesian Biolinguistics. In *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, ed. A.-M. Di Sciullo and C. Boeckx, 42–64. Oxford: Oxford University Press.
- Boeckx, C. In press. Phases beyond explanatory adequacy. In *Phase theory: developing the framework*, ed. Á. Gallego. Berlin: Mouton/de Gruyter.
- Borer, H. 1984. *Parametric syntax*. Dordrecht: Foris.
- Borer, H. 2005. *Structuring Sense (2 vols.)*. Oxford: Oxford University Press.
- Bošković, Ž. 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena*. Amsterdam: Elsevier.

- Bošković, Ž. 2008. What will you have, DP or NP? In *Proceedings of NELS 37*. GLSA.
- Branigan, P. 2011. In defense of macroparameters: Algonquian evidence. Presented at the workshop on linguistic variation and the minimalist program, Universidad Autonoma de Madrid.
- Briscoe, E., and J. Feldman. 2006. Conceptual complexity and the bias-variance tradeoff. Ms., Rutgers University.
- Bromberger, S., and M. Halle. 1989. Why Phonology is Different. *Linguistic Inquiry* 20:51–70.
- Cable, S. 2007. The grammar of Q. Doctoral Dissertation, MIT.
- Chater, N., F. Real, and M.H. Christiansen. 2009. Restrictions on biological adaptation in language evolution. *Proceedings of the National Academy of Sciences* 106:1015–1020.
- Chomsky, N. 1959. Verbal behavior. *Language* 35:26–58.
- Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Foris: Dordrecht.
- Chomsky, N. 1986. *Knowledge of language*. New York: Praeger.
- Chomsky, N. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.
- Chomsky, N. 2000. Minimalist inquiries: the framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. R. Martin, D. Michaels, and J. Uriagereka, 89–155. Cambridge, Mass.: MIT Press.
- Chomsky, N. 2001. Derivation by phase. In *Ken Hale: A Life in Language*, ed. M. Kenstowicz, 1–52. Cambridge, Mass.: MIT Press.
- Chomsky, N. 2004. Beyond explanatory adequacy. In *Structures and beyond*, ed. A. Belletti, 104–131. New York: Oxford University Press.
- Chomsky, N. 2005. Three factors in the language design. *Linguistic Inquiry* 36:1–22.
- Chomsky, N., and M. Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row. Reprinted in 1991. Cambridge, Mass.: MIT Press.
- Cinque, G. 1996. The antisymmetric programme: theoretical and typological implications. *Journal of Linguistics* 32:447–464.
- Cinque, G. 2005. Deriving Greenberg’s Universal 20 and its exceptions. *Linguistic Inquiry* 36:315–332.
- Cinque, G. 2007. A note on linguistic theory and typology. *Linguistic Typology* 11:93–106.
- Cinque, G., and L. Rizzi. 2010. The cartography of syntactic structures. In *Oxford Handbook of Linguistic Analysis*, ed. B. Heine and H. Narrog. Oxford: Oxford University Press.
- Clark, R., and I.G. Roberts. 1993. A computational model of language learnability

- and language change. *Linguistic Inquiry* 24:299–345.
- Culbertson, J. 2010. Learning biases, regularization, and the emergence of typological universals in syntax. Doctoral Dissertation, Johns Hopkins University.
- Cyrino, S., and C. H. Reintges. 2011. Analyticization and the syntax of the synthetic residue. Presented at the workshop on linguistic variation and the minimalist program, Universidad Autonoma de Madrid.
- Diercks, M. 2011. Parameterizing Case: Evidence from Bantu. Ms., Pomona College.
- Dresher, B.E. 1999. Charting the learning path: Cues to parameter setting. *Linguistic Inquiry* 30:27–67.
- Emonds, J.E. 2000. *Lexicon and grammar: The english syntacticon*. Berlin: Mouton de Gruyter.
- Fasanella-Seligrat, A. 2009. Los parámetros en la teoría sintáctica: historia y revisión crítica. Màster interuniversari de ciència cognitiva i llenguatge, Universitat de Barcelona.
- Fasanella-Seligrat, A., and J. Fortuny. 2011. Deriving linguistic variation from learnability conditions in a parametric approach to UG. Presented at the workshop on linguistic variation and the minimalist program, Universidad Autonoma de Madrid.
- Feldman, J., et al. 2000. Minimization of boolean complexity in human concept learning. *Nature* 407:630–632.
- Fitch, W.T. 2008. Glossogeny and phylogeny: cultural evolution meets genetic evolution. *Trends in Genetics* 24:373–374.
- Fodor, J.A., and M. Piattelli-Palmarini. 2010. *What Darwin Got Wrong*. New York: Farrar, Straus and Giroux.
- Fodor, J.D. 1998. Unambiguous triggers. *Linguistic Inquiry* 29:1–36.
- Fodor, J.D. 2001. Setting syntactic parameters. In *The handbook of contemporary syntactic theory*, ed. M. Baltin and C. Collins, 730–767. Malden: Wiley–Blackwell.
- Fortuny, J. 2008. *The emergence of order in syntax*. Amsterdam: John Benjamins.
- Gallego, Á. 2011a. Lexical items and feature bundling. Presented at the workshop on linguistic variation and the minimalist program, Universidad Autonoma de Madrid.
- Gallego, Á. 2011b. Parameters. In *Oxford Handbook of Linguistic Minimalism*, ed. C. Boeckx, 523–550. Oxford: Oxford University Press.
- Garbacz, P. 2011. Morphology and syntax in the Scandinavian vernaculars of Ovan-siljan. Ms., University of Oslo.
- Gervain, J., and J. Mehler. 2010. Speech perception and language acquisition in the first year of life. *Annual review of psychology* 61:191–218.
- Gianollo, C., C. Guardiano, and G. Longobardi. In press. Historial implications of a

- formal theory of syntactic variation. In *Proceedings of DIGS VIII*, ed. S. Anderson and D. Jonas. Oxford: Oxford University Press.
- Giorgi, A., and F. Pianesi. 1997. *Tense and aspect: From semantics to morphosyntax*. Oxford: Oxford University Press.
- Greenberg, J.H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of language* 2:58–90.
- Hale, K., and S. J. Keyser. 1993. On argument structure and the lexical expression of grammatical relations. In *The view from Building 20: Essays in Linguistics in honor of Sylvain Bromberger*, ed. K. Hale and S. J. Keyser, 53–110. Cambridge, Mass.: MIT Press.
- Halle, M., and A. Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from building 20*, ed. K. Hale and S. J. Keyser, 111–176. Cambridge, Mass.: MIT Press.
- Harbour, D. 2009. The semantics, and generality, of features: or, how not to construct a theory of cognitive evolution. Ms, Queen Mary University of London.
- Harley, H., and E. Ritter. 2002a. Person and number in pronouns: A feature-geometric analysis. *Language* 482–526.
- Harley, H., and E. Ritter. 2002b. Structuring the bundle: A universal morphosyntactic feature geometry. In *Pronouns—Grammar and Representation*, ed. H. Weise and H. Simon, 23–39. Amsterdam: John Benjamins.
- Hauser, M. D., N. Chomsky, and W. T. Fitch. 2002. The Faculty of Language: What is it, who has it, and how did it evolve? *Science* 298:1569–1579.
- Holmberg, A. 2010. Parameters in minimalist theory: The case of Scandinavian. *Theoretical Linguistics* 36:1–48.
- Holmberg, A., and I.G. Roberts. 2010. Macroparameters and the minimalist program. Presented at the workshop on linguistic variation and the minimalist program, Barcelona, January 2010.
- Jackendoff, R. 1997. *The architecture of the language faculty*. Cambridge, Mass.: MIT Press.
- Jackendoff, R. 2002. *Foundations of language*. Oxford: Oxford University Press New York.
- Jackendoff, R. 2005. Alternative minimalist visions of language. In *Proceedings from the 41st Annual Meeting of the Chicago Linguistic Society*, 2, 189–226.
- Jackendoff, R. 2007. *Language, Consciousness, Culture: Essays on Mental Structure*. Cambridge, Mass.: MIT Press.
- Jackendoff, R. 2010. *Meaning and the Lexicon: The Parallel Architecture 1975–2010*. Oxford: Oxford University Press.
- Jeong, Y. 2007. *Applicatives: Structure and interpretation from a minimalist per-*

- spective. Amsterdam: John Benjamins.
- Jeong, Y. 2011. Re-examining the “NP/DP” parameter in light of the diversity of east-asian classifiers. In *Proceedings of the 12th Annual Tokyo Conference on Psycholinguistics*, ed. Y. Otsu, 113–131. Tokyo.
- Kandybowicz, J. 2009. Externalization and emergence: On the status of parameters in the minimalist program. *Biolinguistics* 3:93–98.
- Kauffman, S.A. 1993. *The origins of order: Self-organization and selection in evolution*. Oxford: Oxford University Press.
- Kayne, R. S. 2004. Antisymmetry and japanese. In *Variation and universals in biolinguistics*, ed. L. Jenkins, 3–35. Elsevier B. V.
- Kayne, R.S. 2005. *Movement and silence*. Oxford: Oxford University Press.
- Kayne, R.S. 2011. Comparative syntax. Presented at 50 Years of Linguistics at MIT.
- Kemp, C., A. Perfors, and J.B. Tenenbaum. 2007. Learning overhypotheses with hierarchical Bayesian models. *Developmental Science* 10:307–321.
- Kirby, S. 1999. *Function, selection, and innateness: The emergence of language universals*. Oxford: Oxford University Press.
- Koster, J. 2010. Language and tools. Ms., Universiteit Groningen.
- Krifka, M. 1992. Thematic relations as links between nominal reference and temporal constitution. In *Lexical matters*, ed. I. Sag and A. Szabolcsi, 29–53. Stanford, Calif.: CSLI Publications.
- Lasnik, H. 2000. *Syntactic Structures revisited*. Cambridge, Mass.: MIT Press.
- Legate, J.A. 2002. Walpiri: Theoretical implications. Doctoral Dissertation, Massachusetts Institute of Technology.
- Lenneberg, E.H. 1967. *Biological foundations of language*. New York: Wiley.
- Lewontin, R.C. 2000. *The Triple Helix: Gene, Organism, and Environment*. Cambridge, Mass.: Harvard University Press.
- Lightfoot, D. 1999. *The development of language: Acquisition, change, and evolution*. Malden: Wiley-Blackwell.
- Lohndal, T., and J. Uriagereka. 2010. The logic of parametric theory. *Theoretical Linguistics* 36:69–76.
- Longa, V.M., and G. Lorenzo. 2008. What about a (really) minimalist theory of language acquisition? *Linguistics* 46:541–570.
- Longobardi, G. 1994. Reference and proper names: a theory of N-movement in syntax and logical form. *Linguistic inquiry* 25:609–665.
- Longobardi, G. 2004. Methods in parametric linguistics and cognitive history. *Linguistic Variation Yearbook* 3:101–138.
- Longobardi, G. 2005a. A minimalist program for parametric linguistics? In *Organizing grammar*, ed. H. Broekhuis, N. Corver, R. Huybregts, U. Kleinhenz, and

- J. Koster, 407–414. Berlin: Mouton de Gruyter.
- Longobardi, G. 2005b. Toward a Unified Grammar of Reference. *Zeitschrift für Sprachwissenschaft* 24:5–44.
- Longobardi, G., and C. Guardiano. 2011. The biolinguistic program and historical reconstruction. In *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 266–304. Oxford: Oxford University Press.
- Lorenzo, G., and V.M. Longa. 2009. Beyond generative geneticism: Rethinking language acquisition from a developmentalist point of view. *Lingua* 119:1300–1315.
- Mathieu, E. 2011. Wh-in-situ and external parameters. Paper presented at the workshop on Formal Grammar and Syntactic Variation, Universidad Autonoma de Madrid.
- Mathieu, E., and M. Barrie. 2011. Macroparameters don’t exist: the case of polysynthesis and noun incorporation. Paper presented at the workshop on Formal Grammar and Syntactic Variation.
- McGhee, G. 1999. *Theoretical morphology: The concept and its applications*. New York: Columbia University Press.
- Miyagawa, S. 2010. *Why Agree; Why Move*. Cambridge, Mass.: MIT Press.
- Mobbs, I. 2008. ‘Functionalism’, the design of the language faculty, and (disharmonic) typology. MPhil, University of Cambridge.
- Moreno Cabrera, J.C. 2000. *La dignidad e igualdad de las lenguas*. Madrid: Alianza Editorial.
- Narita, H. 2010. Phasing in Full Interpretation. Doctoral Dissertation, Harvard University.
- Newmeyer, F. 2004. Against a parameter-setting approach to language variation. *Linguistic Variation Yearbook* 4:181–234.
- Newmeyer, F.J. 2005. *Possible and probable languages: A generative perspective on linguistic typology*. Oxford: Oxford University Press.
- Obata, M. 2010. Root, Successive-cyclic and Feature-Splitting Internal Merge: Implications for Feature-Inheritance and Transfer. Doctoral Dissertation, University of Michigan.
- Ott, D. 2011. A note on free relative clauses in the theory of phases. *Linguistic Inquiry* 42:183–192.
- Piattelli-Palmarini, M., P. Salaburu, and J. Uriagereka, ed. 2009. *Of minds and language: A Basque encounter with Noam Chomsky*. Oxford: Oxford University Press.
- Pietroski, P. M. 2005. *Events and semantic architecture*. Oxford: Oxford University

- Press.
- Pigliucci, M., and G. Müller, ed. 2010. *Evolution—The Extended Synthesis*. Cambridge, Mass.: MIT Press.
- Pylkkänen, L. 2008. *Introducing arguments*. Cambridge, Mass.: MIT Press.
- Real Puigdollers, C. 2011. Towards a non-syntactic approach of manner incorporation. Paper presented at the Workshop on Verbal Elasticity, Bellaterra, October 2011.
- Reinhart, T., and T. Siloni. 2005. The Lexicon-Syntax Parameter: Reflexivization and Other Arity Operations. *Linguistic Inquiry* 36:389–436.
- Richards, M. 2010a. The fixing problem (A dead-end for the minimalist program?). Presented at the European Science Foundation Exploratory Workshop “Exploring the Roots of Linguistic Diversity — Bilingual Perspectives”, Bellaterra, September 2010.
- Richards, N. 2010b. *Uttering trees*. Cambridge, Mass.: MIT Press.
- van Riemsdijk, H. 2008. Identity Avoidance: OCP-Effects in Swiss Relatives. In *Foundational issues in linguistics*, ed. R. Freidin, C. Otero, and M.-L. Zubizarreta, 227–250. Cambridge, Mass.: MIT Press.
- Rizzi, L. 1982. *Italian Syntax*. Dordrecht. Foris.
- Rizzi, L. 2009. Some elements of syntactic computation. In *Biological Foundations and Origin of Syntax*, ed. D. Bickerton and E. Szathmáry, 63–88. Cambridge, Mass.: MIT Press.
- Rizzi, L. 2010. On the elements of syntactic variation. Presented at the workshop on linguistic variation and the minimalist program, Barcelona, January 2010.
- Roberts, I.G. 2001. Language change and learnability. In *Language acquisition and learnability*, ed. S. Bertolo, 81–125. Cambridge: Cambridge University Press.
- Roberts, I.G. 2010. On the nature of syntactic parameters: A programme for research. Presented at the 2010 Mayfest on “Bridging Typology and Acquisition”.
- Roberts, I.G. 2011. Parametric hierarchies: Some observations. Presented at the workshop on linguistic variation and the minimalist program, Universidad Autónoma de Madrid.
- Roberts, I.G., and A. Holmberg. 2009. Introduction: Parameters in minimalist theory. In *Parametric Variation: Null Subjects in Minimalist Theory*, ed. T. Biberauer, A. Holmberg, I. Roberts, and M. Sheehan, 1–57. Cambridge: Cambridge University Press.
- Savescu Ciucivara, O., and J. Wood. 2010. Re-prefixation and talmy’s parameter. In *Proceedings of NELS 36*, 2–13.
- Schlenker, P. 2006. Ontological symmetry in language: A brief manifesto. *Mind & language* 21:504–539.

- Sheehan, M., and W. Hinzen. 2011. PF-parameters and clausal/nominal denotation. Presented at the workshop on linguistic variation and the minimalist program, Universidad Autonoma de Madrid.
- Sigurdsson, H. Á. 2004. Meaningful silence, meaningless sounds. *Linguistic Variation Yearbook* 4:235–259.
- Smith, N., and A. Law. 2009. On Parametric (and Non-Parametric) Variation. *Biolinguistics* 3:332–343.
- Snyder, W. 2007. *Child Language: The Parametric Approach*. Oxford: Oxford University Press.
- Snyder, W. 2011. On language acquisition and syntactic theory. Presented at 50 Years of Linguistics at MIT.
- Snyder, W. In press. Children’s Grammatical Conservatism: Implications for Syntactic Theory. In *BUCLD 35 Proceedings*. Somerville, Mass.: Cascadilla Press.
- Son, M. 2006. Directed Motion and Non-Predicative Path P. *Nordlyd: Tromsø Working Papers on Language and Linguistics* 176–99.
- Starke, M. 2010. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36.
- Sugisaki, K. 2007. A note on the default values of parameters. *Biolinguistics* 1:114–117.
- Sugisaki, K. 2011. Preposition stranding: Its parametric variation and acquisition. Paper presented at the Workshop on Verbal Elasticity, Bellaterra, October 2011.
- Tokisaki, H. 2010. Recursive compounds and wordstress location. Paper presented at On Linguistic Interfaces II, University of Ulster, Belfast, Ireland.
- Tokisaki, H. 2011. Paper presented at on linguistic interfaces ii, university of ulster, belfast, ireland. stress location and the acquisition of morphosyntactic parameters. WCCFL 28 Online Proceedings. <https://sites.google.com/site/wccfl28pro/tokizaki>.
- Uriagereka, J. 1994. A conjecture on the form of parameters. *UConn Papers in Linguistics* 4:235–259.
- Walkden, G. 2011. The parameter is dead. Long live the parameter? Presented at the Conference on the past and future of Universal Grammar, Durham University, December 2011.
- Westergaard, M.R. 2009. *The acquisition of word order: micro-cues, information structure, and economy*. Amsterdam: John Benjamins.
- Wexler, K. 1998. Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua* 106:23–79.
- Whitman, J. 2008. The classification of constituent order generalizations and diachronic explanation. In *Linguistic universals and language change*, ed. J. Good,

- 233–253. Oxford: Oxford University Press.
- Wiltschko, M. 2009. The composition of INFL. An exploration of tense, tenseless languages and tenseless constructions. Ms., University of British Columbia.
- Yang, C. 2002. *Knowledge and learning in natural language*. Oxford: Oxford University Press.
- Yang, C. 2005. On productivity. *Linguistic Variation Yearbook* 5:265–302.
- Zanuttini, R. 2010. A syntactic analysis of interpretive restrictions on imperative, promissive, and exhortative subjects. Ms., Yale University.