# Chapter 25

# Where, if anywhere, are parameters? A critical historical overview of parametric theory

Frederick J. Newmeyer

University of Washington, University of British Columbia, and Simon Fraser University

Since the late 1970s, crosslinguistic variation has generally been handled by means of UG-specified parameters. On the positive side, thinking of variation in terms of parameterized principles unleashed an unprecedented amount of work in comparative syntax, leading to the discovery of heretofore unknown morphosyntactic phenomena and crosslinguistic generalizations pertaining to them. On the negative side, however, both macroparameters and microparameters have proven themselves to be empirically inadequate and conceptually nonminimalist. Alternatives to parameters are grounded approaches, epigenetic approaches, and reductionist approaches, the last two of which seem both empirically and conceptually quite promising.

## 1 Introduction

The existence of crosslinguistic variation has always been problematic for syntacticians. If there is a universal grammar, one might ask, then why aren't all languages exactly the same? In the earliest work in generative syntax, characterizing the space in which languages could differ, whether at the surface or at a deep level, was not a priority. At the time, surface differences between languages and dialects were generally attributed to language-particular rules or filters.

In the late 1970s, however, a strategy was developed that allowed the simultaneous development of a rich theory of Universal Grammar (UG) along with a detailed account of the limits of crosslinguistic morphosyntactic variation. In this view, syntactic complexity results from the interaction of grammatical subsystems, each characterizable in terms of its own set of general principles. The central goal of syntactic theory now became to identify such systems and to characterize the degree to which they might vary (be "parameterized") from language to language. Chomsky (1995) describes succinctly

how such variation might be accounted for in what, by the early 1980s, was called the "principles-and-parameters" (P&P) approach.

> Within the P&P approach the problems of typology and language variation arise in somewhat different form than before. Language differences and typology should be reducible to choice of values of parameters. A major research problem is to determine just what these options are, and in what components of language they are to be found. (Chomsky 1995: 6)

The first mention of parameters, I believe, was in Chomsky (1976):

> Even if conditions are language- or rule-particular, there are limits to the possible diversity of grammar. Thus, such conditions can be regarded as parameters that have to be fixed (for the language, or for the particular rules, in the worst case), in language learning … It has often been supposed that conditions on applications of rules must be quite general, even universal, to be significant, but that need not be the case if establishing a "parameteric" condition permits us to reduce substantially the class of possible rules. (Chomsky 1976: 315)

An interesting question is why Chomsky at this point would propose parameters, since there is nothing in his 1976 paper that suggests that they need to be incorporated into the theory. A possible answer is that in the same year an MIT dissertation appeared (Kim 1976) that showed that Korean obeys a form of the Tensed-S-Condition, even though Korean does not distinguish formally between tensed and non-tensed clauses. That fact might have planted the seed for the idea of parameterized principles. At around the same time, an "external" inspiration for parameters was provided by the work of Jacques Monod and François Jacob (Monod 1972; "Darwinism reconsidered"). Their idea was that slight differences in timing and arrangement of regulatory mechanisms that activate genes could result in enormous differences. Berwick & Chomsky (2011: 28) has claimed that "Jacob's model in turn provided part of the inspiration for the Principles and Parameters (P&P) approach to language …"

Whatever the direct inspiration for parameterized principles might have been, their adoption triggered an unprecedented explosion of work in comparative syntax. One unquestionably positive consequence of the P&P approach to linguistic theory was to spur investigation of a wide variety of languages, particularly those with structures markedly different from some of the more familiar Western ones. The explanation for this is straightforward. In earlier transformational grammar (oversimplifying somewhat), one worked on the grammar of English, the grammar of Thai, the grammar of Cherokee, and so on, and attempted to extract universal properties of grammars from the principles one found in common among these constructed grammars. But now the essential unity of all grammars, within the limits of parametric variation, was taken as a starting point. One could not even begin to address the grammar of some language without asking the question of how principles of Case, binding, bounding, and so on are parameterized in that language. And that in turn demanded that one have a rough feel for the degree of

parameterization possible for the principle. As Chomsky noted, to delimit the domain of core grammar, we "rely heavily on grammar-internal considerations and comparative evidence, that is, on the possibilities for constructing a reasonable theory of UG and considering its explanatory power in a variety of language types, with an eye open to the eventual possibility of adducing evidence of other kinds" (Chomsky 1981: 9).

The core idea of the P&P approach is that both the principles of UG and the possible parameter settings are part of our genetic endowment:

> [W]hat we "know innately" are the principles of the various subsystems of $S_0$ [= the initial state of the language faculty – FJN] and the manner of their interaction, and the parameters associated with these principles. What we learn are the values of these parameters and the elements of the periphery (along with the lexicon, to which similar considerations apply). The language that we then know is a system of principles with parameters fixed, along with a periphery of marked exceptions. (Chomsky 1986: 150–151)

The original idea was that there are a small number of parameters and small number of settings. This idea allowed two birds to be killed with one stone. Parametric theory could explain the rapidity of acquisition, given the poor input, and explain the crosslinguistic distribution of grammatical elements. As Norbert Hornstein noted:

> The second reason in favor of parameter setting models has been their ability to provide (at least in principle) an answer to Plato's Problem [the fact that we know so much about language based on so little direct evidence – FJN]. The idea is that construing language acquisition as parameter setting eases the problem faced by the child, for setting parameter values is easier than learning the myriad possible rules of one's native language. In other words, the PLD [= Primary Linguistic Data – FJN] can be mined for parameter values more easily than it can be for rules. (Hornstein 2009: 165)

The need to base a theory of parametric variation on the investigation of a wide variety of languages resulted in what Bernard Comrie, always a major critic of the generative approach, referred to approvingly as "one of the most interesting recent developments in linguistic typology … the entry of generative syntax into the field" (Comrie 1988: 458). Comparative studies of the distribution of null-subjects, binding domains, configurationality, and so on became routine by the 1980s and provided a generative interpretation of the kind of crosslinguistic typological studies that were initiated by the work of Joseph Greenberg. In this regard, it is instructive to observe Chomsky's changing rhetorical evaluation of Greenbergian typological work. His first reference to Greenberg was somewhat dismissive, noting that "Insofar as attention is restricted to surface structures, the most that can be expected is the discovery of statistical tendencies, such as those presented by Greenberg (1963)" (Chomsky 1965: 118). In 1981, Chomsky offered what was perhaps his first favorable reference to this line of research:

> Universals of the sort explored by Joseph Greenberg and others have obvious relevance to determining just which properties of the lexicon have to be learned in this

> manner in particular grammars – and to put it in other terms just how much has to be learned as grammar develops in the course of language acquisition. (Chomsky 1981: 95)

By 1982 he was writing that "Greenbergian universals … are ultimately going to be very rich. … They have all the difficulties that people know, they are "surfacy," they are statistical, and so on and so forth, but nevertheless they are very suggestive" (Chomsky 1982: 111). And in 1986, they are "important, … yielding many generalizations that require explanation …" (Chomsky 1986: 21).

In this paper, I do not question the fertility of the research program that was launched by the P&P approach. What I do is to provide a critical review of the various approaches that have been taken to parameters since the late 1970s, discussing their *conceptual* strengths and weaknesses. Given space limitations, my overview will in places be unavoidably somewhat superficial. The paper is organized as follows. Sections 2 through 5 outline various approaches that have been taken with respect to parameters: UG-principle-based, microparametric, macroparametric, and interface-based, respectively. Some of the major conceptual and empirical problems with the classical view of parameters are outlined in §6, and §7 discusses alternatives to the classical approach. §8 is a brief conclusion.

## 2  Parameterized UG principles

All of the subsystems of principles in the Government-Binding Theory were assumed to be parameterized. Consider a few concrete examples:

(1)   Examples of parameterized UG principles:

   a. BINDING (Lasnik 1991). Principle C is parameterized to allow for sentences of the form *John$_i$ thinks that John$_i$ is smart* in languages like Thai and Vietnamese.

   b. GOVERNMENT (Manzini & Wexler 1987). The notion "Governing Category" is defined differently in different languages.

   c. BOUNDING (Rizzi 1982). In English, NP and S are bounding nodes for Subjacency, in Italian NP and S'.

   d. X-BAR ("Origins of phrase structure"). In English, heads precede their complements; in Japanese heads follow their complements.

   e. CASE and THETA-THEORY (Travis 1989). Some languages assign Case and/or Theta-roles to the left, some to the right.

Fewer and fewer parameterized UG principles have been proposed in recent years for the simple reason that there are fewer and fewer widely accepted UG principles. The thrust of the Minimalist Program (MP) has been to reduce the narrow syntactic component and to reinterpret broad universal principles as economy effects of efficient computation. Economy principles are generally assumed not to be parameterized:

> There is simply no way for principles of efficient computation to be parameterized […], it strikes me as implausible to entertain the possibility that a principle like "Shortest Move" could be active in some languages, but not in others. Put differently, […] there can be no parameters within the statements of the general principles that shape natural language syntax. (Boeckx 2011: 210)

On the same page Boeckx proposes the "Strong Uniformity Thesis": Principles of narrow syntax are not subject to parameterization; nor are they affected by lexical parameters.

It should be noted that the very idea of looking for principles of UG has fallen into disrepute in recent years. For example, Chomsky has attributed to them what can only be described as negative qualities:

> [T]ake the LCA (Linear Correspondence Axiom) [Kayne 1994]. If that theory is true, then the phrase structure is just more complicated. Suppose you find out that government is really an operative property. Then the theory is more complicated. If ECP really works, well, too bad; language is more like the spine [i.e., poorly designed – FJN] than like a snowflake [i.e., optimally designed]. (Chomsky 2002: 136)

So if in theory there are very few UG principles and no parameters associated with them, then the question is where and how to capture systematic crosslinguistic variation. Given the organization of grammars in a P&P-type model, the simplest assumption to make is that one group of languages contains a particular feature attraction mechanism that another group lacks, thus allowing the presence or absence of this mechanism to divide languages into typological classes. Some early examples can be illustrated by whether or not a feature setting determines whether V moves to I in a particular language (Emonds 1978; Pollock 1989), whether V moves to C (to derive V2 order) (Besten 1977), and whether N incorporates into V (Baker 1988).

A major debate within parametric theory has centered on the host of the attracting feature. In "microparametric" approaches, the locus of variation lies in individual functional heads. "Macroparametric" approaches are not so restricted. They will be discussed in §3 and §4 respectively.

# 3  The Borer-Chomsky Conjecture and microparametric approaches

Hagit Borer, in *Parametric Syntax* (Borer 1984), made two proposals, which she may or may not have regarded as variants of each other. One is that parameters are restricted to the idiosyncratic properties of lexical items, the other that they are restricted to the inflectional system. Borer wrote:

> …interlanguage variation would be restricted to the idiosyncratic properties of lexical items. These idiosyncrasies, which are clearly learned, would then interact with general principles of UG in a particular way. (Borer 1984: 2–3)

By way of example, she discussed a rule that inserts a preposition in Lebanese Arabic – a rule that does not exist in Hebrew:

(2)   ø ––––––> *la* / [$_{PP}$ ... NP]

Along the same lines, Manzini & Wexler (1987) pointed to language-particular anaphors that have to be associated with parameters: *cakicasin* and *caki* in Korean; *sig* and *hann* in Icelandic.

Now every language has thousands of lexical items, but nobody ever entertained the possibility that every lexical item might be a locus for parametric variation. Borer's proposal that only inflectional elements provide the locus for parametric variation was designed to forestall this possibility. In the same book she wrote:

> It is worth concluding this chapter by reiterating the conceptual advantage that reduced all interlanguage variation to the properties of the inflectional system. The inventory of inflectional rules and of grammatical formatives in any given language is idiosyncratic and learned on the basis of input data. (Borer 1984: 29)

The restriction of parameters to the inflectional system is a somewhat different proposal than their restriction to lexical items. After all, not all lexical items are part of the inflectional system and not all inflections are lexical. However, Borer took "inflectional" in a pretty broad sense, namely, to encompass Case and agreement relations, theta-role assignment, and so on. She recognized an immediate problem here: Inflection-based parameters could not handle some of the best known cases of crosslinguistic variation such as differences in head-order and extraction possibilities.

In any event, the hypothesis that the locus of parametric variation is restricted to exclude major lexical categories came to be known as the "Borer-Chomsky Conjecture".

Borer (1984) appeared before the distinction between lexical and functional categories had been elaborated. Once this distinction had become well accepted, it seemed natural to associate parameters with functional heads, rather than with inflectional items. This idea was first proposed as The Functional Parameterization Hypothesis (FPH) in Fukui (1988). In this view, only functional elements in the lexicon (that is, elements such as Complementizer, Agreement, Tense, etc.) are subject to parametric variation.[1]

It is important to stress that FPH is not a simple extension of the idea that parameters are inflection-located. There have been countless functional categories proposed that have nothing to do with inflection, no matter how broadly this concept is interpreted. So adverbs, topic, focus, and so on are typically thought to be housed in functional categories, even though they are not in many languages "inflectional".

Associating parameters with functional heads has been claimed to have both methodological and theoretical advantages. Methodologically, it allows "experiments" to be constructed comparing two closely-related variants, thereby pinpointing the possible degree of variation. The ideal situation then would be to compare speech varieties that differ from each other only in terms of (most ideally) one or, failing that, only a few variables. Richard Kayne remarks:

---

[1] Fukui himself exempted ordering restrictions from this hypothesis.

> If it were possible to experiment on languages, a syntactician would construct an experiment of the following type: take a language, alter a single one of its observable syntactic properties, examine the result and see what, if any, other property has changed as a consequence. If some property has changed, conclude that it and the property that was altered are linked to one another by some abstract parameter. Although such experiments cannot be performed, I think that by examining pairs (and larger sets) of ever more closely related languages, one can begin to approximate the results of such an experiment. To the extent that one can find languages that are syntactically extremely similar to one another, yet clearly distinguishable and readily examinable, one can hope to reach a point such that the number of observable differences is so small that one can virtually see one property covarying with another. (Kayne 2000: 5–6)

In other words, in Kayne's view, this "microparametric variation" (as he called it) is the best testing ground for the hypothesis that syntactic variation can be reduced to a finite set of parameters.

Along more theoretical lines, it has been claimed that functional-category-situated microparameters impose a strong limit on what can vary, crosslinguistic differences now being reduced to differences in features, thereby restricting learning to the lexicon (Kayne 2000; Roberts 2010; Thornton & Crain 2013).[2] Indeed, Chomsky has often asserted that microparameters are necessary in order to solve Plato's Problem:

> Apart from lexicon, [the set of possible human languages] is a finite set, surprisingly; in fact, a one-membered set if parameters are in fact reducible to lexical properties [associated with functional categories – FJN] … How else could Plato's problem be resolved? (Chomsky 1991: 26)

## 4  Macroparameters

Not all minimalists have embraced the Borer-Chomsky Conjecture and consequent turn to microparameters. Mark Baker, in particular, while not denying that there are micro-level points of variation between languages, has defended what he calls "macroparameters" (Baker 1996), that is, parametric differences that cannot be localized in simple differences in attracting features of individual functional heads. He gives as examples, among others, the Head Directionality Parameter (i.e. VO vs. OV), where functional categories play no obvious role, the Polysynthesis Parameter, which in his account refers to the lexical category "Verb", and an agreement parameter (Baker 2008) distinguishing Niger-Congo languages from Indo-European languages, which, in opposition to a strong interpretation of the Borer-Chomsky Conjecture, applies to the full range of functional categories. Another example of a macroparameter is the compounding parameter of Snyder (2001: 328), which divides languages into those that allow formation of endocentric

---

[2] As an anonymous reviewer points out, this claim is highly dependent on the nature of the features and the role that they play in the system.

compounds during the syntactic derivation and those that do not. The NP/DP macroparameter of Bošković & Gajewski (2011) distinguishes "NP languages", which lack articles, permit left-branch extraction and scrambling, but disallow NEG-raising, from "DP languages", which can have articles, disallow left-branch extraction and scrambling, but allow NEG-raising. And Huang (2007) points to many features that distinguish Chinese-type languages from English-type languages, including a generalized classifier system, no plural morphology, extensive use of light verbs, no agreement, tense, or case morphology, no overt *wh*-movement, and radical pro-drop.

Baker and other advocates of macroparameters share the conviction long held by advocates of holistic typology that languages can be partitioned into macro-scope broad classes, typically (or, at least, ideally) where the setting of one feature entails a cascade of shared typological properties. As Baker puts it, "the macroparametric view is that there are at least a few simple (not composite) parameters that define typologically distinct sorts of languages" (Baker 2008: 355).

# 5 Parameters as being stated at the interfaces

Under the perspective that parameters are stated at the interfaces, lexical items are subject to a process of generalized late insertion of semantic, formal, and morphophonological features after the syntax, which is where all variation would take place. Or, as another possibility, the parametric differences would derive from the way in which such features are interpreted by the interfaces or by processes that manipulate the features on the path from spell out to the interfaces. There has been some debate as to whether there is parametric variation at the Conceptual-Intentional (C-I) interface. Angel Gallego remarks:

> … it would be odd for semantic features to be a source of variation, which leaves us with formal and morphophonological features as more likely suspects. … Considered together, the observations by Chomsky (2001) and Kayne (2005; 2008) appear to place variation in the morphophonological manifestation of closed classes (i.e. functional categories, which contain unvalued features)." (Gallego 2011: 543–544)

However, for Ramchand & Svenonius (2008) the narrow syntax provides a "basic skeleton" to C-I, but languages vary in terms of how much their lexical items explicitly encode about the reference of variables like T, Asp, and D.

# 6 Conceptual and empirical problems with parameters

Before moving on to nonparametric approaches to variation, it would be useful to highlight some of the main problems with the classic view of parameters as being innately-provided grammatical constructs (for an earlier discussion, see Newmeyer 2005).

## 6.1 No macroparameter has come close to working

The promise of parameters in general and macroparameters in particular is that from the interaction of a small number of simply-stated parameters, the vast complexity of human language morphosyntax might be derived. As Martin Haspelmath put it:

> According to the principles and parameters vision, it should be possible at some point to describe the syntax of a language by simply specifying the settings of all syntactic parameters of Universal Grammar. We would no longer have any need for thick books with titles like *The Syntax of Haida* (cf. Enrico 2003's 1300-page work), and instead we would have a simple two-column table with the parameters in the first column and the positive or negative settings in the second column. (Haspelmath 2008: 80)

Needless to say, nothing remotely like that has been achieved. The problem is that "few of the implicational statements at the heart of the traditional Principles-and-Parameters approach have stood the test of time" (Boeckx 2011: 216). The clustering effects are simply not very robust. The two most-studied macroparameters, I believe, are the Null Subject (Pro-drop) and the Subjacency parameters, neither of which is much evoked in recent work. As for the former: "History has not been kind to to the Pro-drop Parameter as originally stated" (Baker 2008: 352). And Luigi Rizzi notes that "In retrospect, [subjacency effects] turned out to be a rather peripheral kind of variation. Judgments are complex, graded, affected by many factors, difficult to compare across languages, and in fact this kind of variation is not easily amenable to the general format of parameters ..." (Rizzi 2014: 16).

## 6.2 "Microparameter" is just another word for "language-particular rule"

Let's say that we observe two Italian dialects, one with a *do*-support-like structure and one without. We could posit a microparametric difference between the dialects, perhaps hypothesizing that one contains an attracting feature that leads to *do*-support and one that does not. But how would such an hypothesis differ in substance from saying that one dialect has a rule of *do*-support that the other one lacks? Indeed, Norbert Hornstein has stressed that "microparameter" is just another words for "rule":[3]

> Last of all, if parameters are stated in the lexicon (the current view), then parametric differences reduce to whether a given language contains a certain lexical item or not. As the lexicon is quite open ended, even concerning functional items as a glance at current cartographic work makes clear, the range of variation between grammars/languages is also open ended. In this regard it is not different from a rule-based approach in that both countenance the possibility that there is no bound on the possible differences between languages. (Hornstein 2009: 165)

---

[3] See Rizzi (2014: 22–27) for a defense of the idea that microparameters are not merely rules under a different name and Boeckx (2014) for a reply to Rizzi.

Michal Starke has made a similar observation:

> Thirty years ago, if some element moved in one language but not in another, a movement rule would be added to one language but not to the other. Today, a feature "I want to move" ("EPP', "strength', etc.) is added to the elements of one language but not of the other. In both cases (and in all attempts between them), variation is expressed by stipulating it. Instead of a theory, we have brute force markers. (Starke 2014: 140)

## 6.3 There would have to be hundreds, if not thousands, of parameters

Tying parameters to functional categories was a strong conjecture in the 1980s, since there were so few generally recognized functional categories at the time. There were so few, in fact, that it was easy to believe that only a small number of parameters would be needed. Pinker (1994: 112), for example, speculated that there are just "a few mental switches'. Lightfoot (1999: 259) suggested that there are about 30 to 40 parameters. For Adger (2003: 16), "There are only a few parameters'. Roberts & Holmberg (2005) increased the presumed total to between 50 and 100. Fodor (2001: 734) was certainly correct when she observed that "it is standardly assumed that there are fewer parameters than there are possible rules in a rule-based framework; otherwise, it would be less obvious that the amount of learning to be done is reduced in a parametric framework'. At this point in time, many hundreds of parameters have been proposed. Gianollo, Guardiano & Longobardi (2008) propose 47 parameters for DP alone on the basis of 24 languages, only five of which are non-Indo-European, and in total representing only 3 families. Longobardi & Guardiano (2011) up the total to 63 binary parameters in DP. As Cedric Boeckx has stressed: "It is not at all clear that the exponential growth of parameters that syntacticians are willing to entertain is so much better a situation for the learner than a model without parameters at all" (Boeckx 2014: 157).

One way to circumvent this problem would be to posit nonparametric differences among languages, thereby maintaining the possibility of a small number of parameters. Let us examine this idea now.

## 6.4 Nonparametric differences among languages undercut the entire parametric program

Are all morphosyntactic differences among languages due to differences in parameter setting? Generally that has been assumed not to be the case. Charles Yang was expressing mainstream opinion when he wrote that "… it seems highly unlikely that all possibilities of language variation are innately specified …' (Yang 2011: 191). From the beginning of the parameteric program it has been assumed that some features are extraparametric. Outside of (parametrically relevant) core grammar are:

> … borrowings, historical residues, inventions, and so on, which we can hardly expect to – and indeed would not want to – incorporate within a principled theory

of UG. ... How do we delimit the domain of core grammar as distinct from marked periphery? ... [We] rely heavily on grammar-internal considerations and comparative evidence, that is, on the possibilities for constructing a reasonable theory of UG and considering its explanatory potential in a variety of language types ... (Chomsky 1981: 8–9)

In other words, some language-particular features are products of extraparametric language-particular rules. Consider, for example, the treatment of Hixkaryana in Baker (2001), based on an earlier proposal in Kayne (1994). This language for the most part manifests OVS word order:

(3)   Hixkaryana (Derbyshire 1985)
      Kanawa yano toto
      canoe   took person
      'The man took the canoe.'

One's first thought might be that what is needed is a parameter allowing for OVS order. But in fact Baker rejects the idea that a special word order parameter is involved here. Rather, he argues that Hixkaryana is (parametrically) SOV and allows the fronting of VP by a movement rule:

(4)   S[OV] → [OV]S

In other words, in this account word order is determined *both* by a parameter and a language-specific rule.

It is quite implausible that every syntactic difference between languages and dialects results from a difference in parameter settings. Consider the fact that there are several dozen systematic morphosyntactic differences between the Norfolk dialect and standard British English (Trudgill 2003), most of which appear to be analytically independent. If each were to be handled by a difference in parameter setting, then, extrapolating to all of the syntactic distinctions in the world's languages, there would have to be thousands – if not millions – of parameters. That is obviously an unacceptable conclusion from an evolutionary standpoint, given that the set of parameters and their possible settings is, by hypothesis, innate. Furthermore, many processes that can hardly be described as "marginal" have been assumed to apply in PF syntax (where the standard view, I believe, is that parameters are not at work), including extraposition and scrambling (Chomsky 1995); object shift (Holmberg 1999; Erteschik-Shir 2005); head movements (Boeckx & Stjepanovic 2001); the movement deriving V2 order (Chomsky 2001); linearization (i.e. VO vs. OV) (Chomsky 1995; Takano 1996; Fukui & Takano 1998; Uriagereka 1999); and even *Wh*-movement (Erteschik-Shir 2005).

I think that it is fair to say that, after 35 years of investigation, nobody has a clear idea about which syntactic differences should be considered parametric and which should not be.[4] But one thing seems clear: If learners need to learn rules anyway, very little is gained by positing parameters.

---

[4] But see Smith & Law (2009) for an interesting discussion of criteria for distinguishing parametric and nonparametric differences.

## 6.5 Parametric theory is arguably inherently unminimalist

There are a number of ways that the assumptions of the Minimalist Program have entailed a rethinking of parameters and the division of labor among the various components for the handling of variation. In one well-known formulation, "FLN [= the faculty of language in the narrow sense – FJN] comprises only the core computational mechanisms of recursion as they appear in narrow syntax and the mapping to the interfaces' (Hauser, Chomsky & Fitch 2002: 1573). Where might parameters fit into such a scenario? In one view, "… if minimalists are right, there cannot be any parameterized principles, and the notion of parametric variation must be rethought.' (Boeckx 2011: 206). That is, given that the main thrust of the minimalist program is the reduction to the greatest extent possible of the elements of UG, there would seem to be no place for innately-specified parameters.

Despite the above, a great deal of work within the general envelope of the MP is still devoted to fleshing out parameters, whether micro or macro. For example, Yang (2011: 202–203) writes that "a finite space of parameters or constraints is still our best bet on the logical problem of language acquisition'. Note also that in many approaches, "the mapping to the interfaces' encompasses a wide variety of operations. To give one example, "UG makes available a set **F** of features (linguistic properties) and operations $C_{HL}$ … that access F to generate expressions" (Chomsky 2000: 100). In addition to features and the relevant operations on them, minimalists have attributed to the narrow syntax principles governing agreement, labelling, transfer, probes, goals, deletion, and economy principles such as Last Resort, Relativized Minimality (or Minimize Chain Links), and Anti-Locality. None of these fall out from recursion *per se*, but rather represent conditions that underlie it or that need to be imposed on it. To that we can add the entire set of mechanisms pertaining to phases, including what nodes count for phasehood and the various conditions that need to be imposed on their functioning, like the Phase Impenetrability Condition. And then there is the categorial inventory (lexical and functional), as well as the formal features they manifest. The question, still unresolved, is whether any of these principles, conditions, and substantive universals could be parameterized, in violation of the Strong Uniformity Thesis, but not of weaker proposals. If so, that would seem to allow for parametric variation to be manifested in the journey towards the interfaces.

# 7 Alternatives to the classic Principles-and-Parameters model

Chomsky (2005) refers to "the three factors in language design', namely, genetic endowment, experience, and principles not specific to the faculty of language. The last-named "third factor explanations', which include principles of data analysis and efficient com-

putation, among other things, provide a potential alternative to the nonminimalist polifieration of parameters and their settings.[5]

The following subsections discuss alternatives to the classic P&P model, all appealing to one degree or another to third factor explanations. They are grounded approaches (§7.1), epigenetic (or emergentist) approaches (§7.2), and reductionist approaches (§7.3).

## 7.1 Grounded approaches

A grounded approach is one in which some principle of UG is grounded in – that is, ultimately derived from – some third factor principle. Along these lines, a long tradition points to a particular constraint, often an island constraint, and posits that it is a grammaticalized processing principle. One of the first publications to argue for grounded constraints was Fodor (1978), where two island constraints are posited, one of which is the Nested Dependency Constraint (NDC):

(5)  The Nested Dependency Constraint: If there are two or more filler-gap dependencies in the same sentence, their scopes may not intersect if either disjoint or nested dependencies are compatible with the well-formedness conditions of the language.

As Fodor noted, the processing-based origins of this constraint seem quite straightforward.

Another example is the Final Over Final Constraint (FOFC), proposed originally in Holmberg (2000):

(6)  Final Over Final Constraint: If α is a head-initial phrase and β is a phrase immediately dominating α, then β must be head-initial. If α is a head-final phrase, and β is a phrase immediately dominating α, then β can be head-initial or head-final.

As one consequence of the FOFC, there are COMP-TP languages that are verb-final, but there are no TP-COMP languages that are verb-initial. Holmberg and his colleagues interpret this constraint as the following UG principle:

(7)  A theoretical reinterpretation of the FOFC: If a phase-head PH has an EPP-feature, then all the heads in its complement domain from which it is nondistinct in categorial features must have an EPP-feature. (Biberauer, Holmberg & Roberts 2008: 13)

Walkden (2009) points out that FOFC effects are accounted for by the processing theory developed in Hawkins (2004) and hence suggests that (7) is a good example of a grounded UG principle.[6]

---

[5] In what follows, I consider classical functional explanations of grammatical structure to be of the third factor type. It is not clear whether Chomsky shares that view.

[6] Mobbs (2014) builds practically all of Hawkins's parsing theory into UG.

Note that neither Fodor nor Walkden have reduced the number of UG constraints; they have merely attributed the origins of these constraints to what in Chomsky's account would be deemeed a third factor. Naturally, the question arises as to whether these principles would need to be parameterized. The answer is "apparently so', since the NDC does not govern Swedish grammar (Engdahl 1985: 75) and the FOFC is not at work in Chinese (Chan 2013). In other words, grounded approaches, whatever intrinsic interest they might possess, do not prima facie reduce the number of UG principles and parameters.

## 7.2 Epigenetic approaches

Let us turn now to "epigenetic' or "emergentist" approaches to variation, where parameters are not provided by an innate UG. Rather, parametric effects arise in the course of the acquisition process through the interaction of certain third factor learning biases and experience. UG creates the space for parametric variation by leaving certain features underspecified. There are several proposals along these lines, among which are Gianollo, Guardiano & Longobardi (2008); Boeckx (2011); and Biberauer et al. (2014) (preceded by many papers by the same four authors). For reasons of space, I focus exclusively on Biberauer et al. (2014). In their way of looking at things, the child is conservative in the complexity of the formal features that it assumes are needed (what they call "feature economy') and liberal in its preference for particular features to extend beyond the input (what they call "input generalization" and which is a form of the superset bias). The idea is that these principles drive acquisition and thus render parameters unnecessary, while deriving the same effects. Consider first their word order hierarchy, represented in Figure 1:

a. Is head-final present?

No: **head-initial**      b. Yes: Present on all heads?

Yes: **head-final**      No: present on [+V] heads?

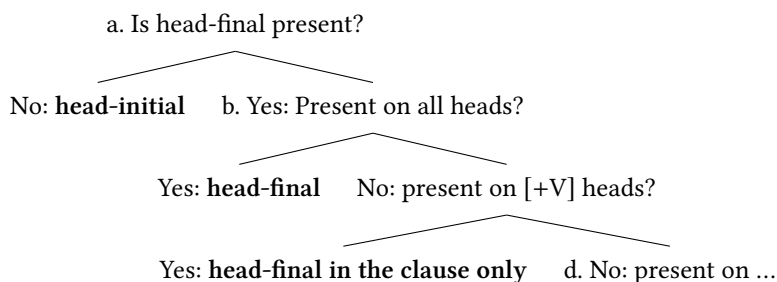Yes: **head-final in the clause only**      d. No: present on …

Figure 1: The Word Order Hierarchy of Biberauer et al. (2014: 110)

To illustrate with a made up example, let's say that a language is consistently head-initial except in NP, where the noun follows its complements. However, there is a definable class of nouns in this language that do precede their complements and a few nouns in this language behave idiosyncratically in terms of the positioning of their specifiers and complements (much like the English word *enough*, which is one of the few degree modifiers that follows the adjective). In their theory, the child will go through the fol-

lowing stages of acquisition, zeroing in step-by-step on the adult grammar. First it will assume that all phrases are head-initial, even noun phrases. Second, it will assume that all NPs are head-final. Third, it will learn the systematic class of exceptions to the latter generalization, and finally, it will learn the purely idiosyncratic exceptions.

The other hierarchies proposed in Biberauer et al. are more complex and depend on many assumptions about the feature content of particular categories. Consider for example their null argument hierarchy and the questions posed by the child in determining the status of such arguments in its grammar (Figure 2).

a. Are uφ-features present
on probes?

No:
**Radical pro-drop**

Yes:
Are uφ-features present
on all probes?

Yes:
**Pronominal arguments**

No:
Are uφ-features fully specified
on some probes?

No:
**Non-pro drop**

Yes:
Are uφ-features fully specified
on T?
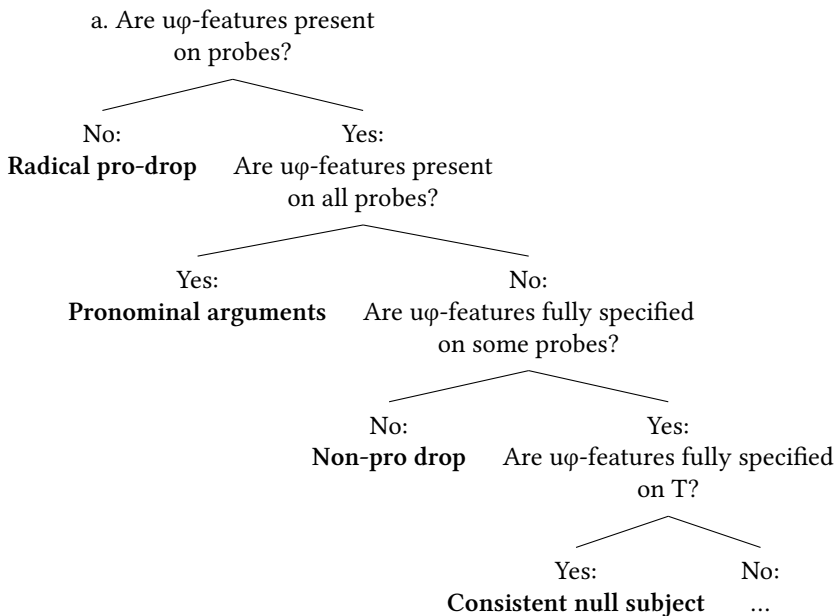
Yes:
**Consistent null subject**

No:
...

Figure 2: The Null Argument Hierarchy of Biberauer et al. (2014: 112)

There are many issues that one might raise about this acquisition scenario, the most important of which being whether children proceed from general to the particular, correcting themselves as they go, and gradually zero in on the correct grammar. Indeed, many acquisitionists argue for precisely the reverse set of steps, in which children learn idiosyncratic items before broad generalizations. Theresa Biberauer herself acknowledges (p.c., November 21, 2014) that these steps can be overridden in certain circumstances. For example, the early stages might correspond to a pre-production stage or possibly acquirers pass through certain stages very quickly if counterevidence to an earlier hypothesis is readily available. And finally, as they note, frequency effects can distort the smooth transition through the hierarchies.

One might also ask how much the child has to know *in advance* of the progression through the hierarchy. For example, given their scenario, the child has to know to ask

"Are unmarked phi-features fully specified on some probes?" That implies a lot of grammatical knowledge. Where, one might ask, does this knowledge come from and how does the child match this knowledge with the input?[7]

Despite these unresolved questions, the Biberauer et al. approach presents a view that preserves the major insights of parametric theory without positing UG-based parameters. As such, it needs to be taken very seriously.

## 7.3 Reductionist approaches and the need for language-particular rules

Reductionist approaches differ from epigenetic ones by reducing still further the role played by an innate UG in determining crosslinguistic variation. For example, returning to the FOFC, Trotzke, Bader & Frazier (2013) provide evidence that the best motivated account is to remove it entirely from the grammar, since, in their view, it can be explained in its entirety by systematic properties of performance systems. They also deconstruct the Head-Complement parameter in a similar fashion:

> [T]he physics of speech, that is, the nature of the articulatory and perceptual apparatus requires one of the two logical orders, since pronouncing or perceiving the head and the complement simultaneously is impossible. Thus, the head-complement parameter, according to this approach, is a third-factor effect. (Trotzke, Bader & Frazier 2013: 4)

Which option is chosen, of course, has to be built into the grammar of individual languages, presumably via its statement as a language-particular rule.

To take another example, Kayne (1994) provided an elaborate UG-based parametric explanation of why rightward movement is so restricted in language after language. But Ackema & Neeleman (2002) argue that the apparent ungrammaticality of certain "right-displaced" syntactic structures should not be accounted for by syntax proper (that is, by the theory of competence), but rather by the theory of performance. In a nutshell, such structures are difficult to process. A necessary consequence of their approach is that it is necessary to appeal to language-particular rules to account for the fact that languages differ from each other in the degree to which displacement to the right is permitted.

The microparametric approach to variation is well designed to capture the fact that even closely related speech varieties can vary from each other in many details. The question is why one would want to appeal to microparameters when the traditional term "rule" seems totally appropriate (see §6.2). The resistance to the idea of reviving the idea of (language-particular) rules is unsettling to some, perhaps because the idea of "rules" brings back the ghosts of pre-generative structuralism, where it was believed by

---

[7] An anonymous referee asks: "To be honest, I don't see how the approach sketched here is different from parameter setting. Perhaps it's my own ignorance of Biberauer et al.'s proposal, but if the hierarchy of questions that the learner must address is innate, how does this differ from parameters that are innately specified?" As I understand their proposal, the hierarchy of questions falls out from general learning principles, though I am hazy on the details of precisely how.

some that "languages could differ from each other without limit" (Joos 1957: 96), and the spectre of early transformational grammar, where grammars were essentially long lists of rules. But to call a language-particular statement a "rule" is not to imply that *anything* can be a rule. Possible rules are still constrained by UG. That of course raises the question of what is in UG. An obvious candidate is the Merge operation or something analogous, but surely there must be a lot more than that. For example, it is hard to see how the broad architecture of the grammar could be learned inductively. Consider the fact that syntactic operations have no access to the segmental phonology: There is no language in which displacement – Internal Merge, if you will – targets only those elements with front vowels. It seems probable that this state of affairs derives from UG.

However, if the general thrust of the work of John A. Hawkins is correct (see Hawkins 1994; 2004; 2014), the major constraints on the nature of rules derive from the exigencies of language processing. No language has a rule that lowers a filler exactly two clauses deep, leaving a gap in initial position. Such a rule, while theoretically possible, is so improbable (for processing reasons) that it will never occur. Norbert Hornstein's approach to variation, succinctly stated in the following passage, also stresses that it is not necessary to appeal to UG to explain why certain logically possible properties of grammars do not occur:

> There is no upper bound on the ways that languages might differ though there are still some things that grammars cannot do. A possible analogy for this conception of grammar is the variety of geometrical figures that can be drawn using a straight edge and compass. There is no upper bound on the number of possible different figures. However, there are many figures that cannot be drawn (e.g. there will be no triangles with 20 degree angles). Similarly, languages may contain arbitrarily many different kinds of rules depending on the PLD [ = primary linguistic data – FJN] they are trying to fit. However, none will involve binding relations in which antecedents are c-commanded by their anaphoric dependents or where questions are formed by lowering a *Wh*-element to a lower CP. Note that this view is not incompatible with languages differing from one another in various ways. (Hornstein 2009: 167)

In my view, the idea that a grammar is composed of language-particular rules constrained by both UG principles and third factor principles is an appealing vision that stands to inform research on crosslinguistic variation in the years to come.

## 8 Conclusion

Since the late 1970s, crosslinguistic variation has generally been handled by means of UG-specified parameters. On the positive side, thinking of variation in terms of parameterized principles unleashed an unprecedented amount of work in comparative syntax, leading to the discovery of heretofore unknown morphosyntactic phenomena and crosslinguistic generalizations pertaining to them. On the negative side, however,

both macroparameters and microparameters have proven themselves to be empirically inadequate and conceptually nonminimalist. Alternatives to parameters are grounded approaches, epigenetic approaches, and reductionist approaches, the last two of which seem both empirically and conceptually quite promising.

## Acknowledgements

## References

Ackema, Peter & Ad Neeleman. 2002. Effects of short term storage in processing rightward movement. In Sieb Noteboom, Fred Weerman & Frank Wijnen (eds.), *Storage and computation in the language faculty*, 219–256. Dordrecht: Kluwer.

Adger, David. 2003. *Core syntax: A minimalist approach*. Oxford: Oxford University Press.

Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Baker, Mark. 1996. *The polysynthesis parameter*. Oxford: Oxford University Press.

Baker, Mark. 2001. *The atoms of language: The mind's hidden rules of grammar*. New York: Basic Books.

Baker, Mark. 2008. The macroparameter in a microparametric world. In Teresa Biberauer (ed.), *The limits of syntactic variation*, 351–374. Amsterdam: John Benjamins.

Berwick, Robert C. & Noam Chomsky. 2011. The biolinguistic program: The current state of its development. In Anna-Maria Di Sciullo & Cedric Boeckx (eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 19–41. Oxford: Oxford University Press.

Besten, Hans den. 1977. On the interaction of root transformations and lexical deletion rules. In Werner Abraham (ed.), *On the formal syntax of the Westgermania*, 47–131. Amsterdam: John Benjamins.

Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2008. Linearising disharmonic word orders: The Final-over-Final Constraint. In Kyoung Ae Kim & Jong Yurl Yoon (eds.), *Perspectives on linguistics in the 21st century (Proceedings of ICLK 2007)*, 301–318. Seoul.

Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan. 2014. Complexity in comparative syntax: The view from modern parametric theory. In Laurel B. Preston & Frederick J. Newmeyer (eds.), *Measuring grammatical complexity*, 103–127. Oxford: Oxford University Press.

Boeckx, Cedric. 2011. Approaching parameters from below. In Cedric Boeckx & Anna-Maria Di Sciullo (eds.), *Biolinguistic approaches to language evolution and variation*, 205–221. Oxford: Oxford University Press.

Boeckx, Cedric. 2014. What principles and parameters got wrong. In M. Carme Picallo (ed.), *Linguistic variation in the minimalist framework*, 155–178. Oxford: Oxford University Press.

Boeckx, Cedric & Sandra Stjepanovic. 2001. Head-ing toward PF. *Linguistic Inquiry* 32(2). 345–355.

Borer, Hagit. 1984. *Parametric syntax: Case studies in Semitic and Romance languages.* Vol. 13 (Studies in Generative Grammar). Dordrecht: Foris.

Bošković, Željko & Jon Gajewski. 2011. Semantic correlates of the NP/DP parameter. *North-Eastern Linguistics Society* 39. 121–134.

Chan, Brian Hok-Shing. 2013. Sentence-final particles, complementizers, antisymmetry, and the Final-over-Final Constraint. In Theresa Biberauer & Michelle Sheehan (eds.), *Theoretical approaches to disharmonic word order*, 445–468. Oxford: Oxford University Press.

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

Chomsky, Noam. 1976. Conditions on rules of grammar. *Linguistic Analysis* 2. 303–351.

Chomsky, Noam. 1981. *Lectures on government and binding.* Vol. 9 (Studies in Generative Grammar). Dordrecht: Foris.

Chomsky, Noam. 1982. *The generative enterprise: A discussion with Riny Huybregts and Henk van Riemsdijk.* Dordrecht: Foris.

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

Chomsky, Noam. 1991. Linguistics and cognitive science: Problems and mysteries. In Asa Kasher (ed.), *The Chomskyan turn: Generative linguistics, philosophy, mathematics, and psychology*, 26–55. Oxford: Blackwell Publishing.

Chomsky, Noam. 1995. *The minimalist program.* Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

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

Chomsky, Noam. 2002. *On nature and language.* Cambridge: Cambridge University Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36. 1–22.

Comrie, Bernard. 1988. Linguistic typology. In Frederick Newmeyer (ed.), *Linguistics: The Cambridge survey, 1: Linguistic theory: Foundations*, 447–461. Cambridge: Cambridge University Press.

Derbyshire, Desmond C. 1985. *Hixkaryana and linguistic typology.* Arlington, TX: Summer Institute of Linguistics.

Emonds, Joseph E. 1978. The verbal complex V' – V in French. *Linguistic Inquiry* 9(2). 151–175.

Engdahl, Elisabet. 1985. Interpreting questions. In David R. Dowty, Lauri Karttunen & Arnold M. Zwicky (eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives*, 67–93. Cambridge: Cambridge University Press.

Enrico, John. 2003. *Haida syntax.* Lincoln: University of Nebraska Press.

Erteschik-Shir, Nomi. 2005. Sound patterns of syntax: Object shift. *Theoretical Linguistics* 31. 47–93.

Fodor, Janet D. 1978. Parsing strategies and constraints on transformations. *Linguistic Inquiry* 9. 427–473.

Fodor, Janet D. 2001. Setting syntactic parameters. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*, 730–767. Oxford: Blackwell Publishing.

Fukui, Naoki. 1988. Deriving the differences between English and Japanese: A case study in parametric syntax. *English Linguistics* 5. 249–270.

Fukui, Naoki & Yuji Takano. 1998. Symmetry in syntax: Merge and demerge. *Journal of East Asian Linguistics* 7. 27–86.

Gallego, Ángel J. 2011. Parameters. In Cedric Boeckx (ed.), *Oxford handbook of linguistic minimalism*, 523–550. Oxford: Oxford University Press.

Gianollo, Chiara, Cristina Guardiano & Giuseppe Longobardi. 2008. Three fundamental issues in parametric linguistics. In Teresa Biberauer (ed.), *The limits of syntactic variation*, 109–142. Amsterdam: John Benjamins.

Greenberg, Joseph H. 1963. Some universals of grammar with special reference to the order of meaningful elements. In Joseph Greenberg (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press.

Haspelmath, Martin. 2008. Parametric versus functional explanations of syntactic universals. In Teresa Biberauer (ed.), *The limits of syntactic variation*, 75–107. Amsterdam: John Benjamins.

Hauser, Marc D, Noam Chomsky & W Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298(5598). 1569–1579.

Hawkins, John A. 1994. *A performance theory of order and constituency*. Vol. 73 (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. Oxford: Oxford University Press.

Holmberg, Anders. 1999. Remarks on Holmberg's generalization. *Studia Linguistica* 53. 1–39.

Holmberg, Anders. 2000. Deriving OV order in Finnish. In Peter Svenonius (ed.), *The derivation of VO and OV*, 123–152. Amsterdam: John Benjamins.

Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and universal grammar*. Cambridge: Cambridge University Press.

Huang, C.-T. James. 2007. *The macro-history of Chinese syntax and the theory of language change*. Talk given at the University of Chicago.

Jacob, François. Darwinism reconsidered. *Le Monde* 1977-09-06.

Joos, Martin (ed.). 1957. *Readings in linguistics: The development of descriptive linguistics in America since 1925*. New York: American Council of Learned Societies.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kayne, Richard S. 2000. *Parameters and universals*. Oxford: Oxford University Press.

Kayne, Richard S. 2005. *Movement and silence*. Oxford: Oxford University Press.

Kayne, Richard S. 2008. Some preliminary comparative remarks on French and Italian definite articles. In Robert Freidin, Carlos Otero & Maria-Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 291–321. Cambridge, MA: MIT Press.

Kim, Wha-Chun M. 1976. *The theory of anaphora in Korean syntax*. Massachusetts Institute of Technology PhD thesis.

Lasnik, Howard. 1991. On the necessity of binding conditions. In Robert Freidin (ed.), *Principles and parameters in comparative grammar*, 7–28. Cambridge, MA: MIT Press.

Lightfoot, David W. 1999. *The development of language: Acquisition, change, and evolution*. Blackwell/Maryland Lectures in Language and Cognition, 1. Oxford: Blackwell Publishing.

Longobardi, Giuseppe & Cristina Guardiano. 2011. The biolinguistic program and historical reconstruction. In Cedric Boeckx & Anna-Maria Di Sciullo (eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 266–304. Oxford: Oxford University Press.

Manzini, Maria Rita & Kenneth Wexler. 1987. Parameters, binding, and learning theory. *Linguistic Inquiry* 18. 413–444.

Mobbs, Iain. 2014. *Minimalism and the design of the language faculty*. University of Cambridge PhD thesis.

Monod, Jacques. 1972. *Chance and necessity*. London: Collins.

Newmeyer, Frederick J. 2005. *Possible and probable languages: A generative perspective on linguistic typology*. Oxford: Oxford University Press.

Pinker, Steven. 1994. *The language instinct: How the mind creates language*. New York: Morrow.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.

Ramchand, Gillian & Peter Svenonius. 2008. Mapping a parochial lexicon onto a universal semantics. In Teresa Biberauer (ed.), *The limits of syntactic variation*, 219–245. Amsterdam: John Benjamins.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Vol. 11 (Studies in Generative Grammar). Dordrecht: Foris.

Rizzi, Luigi. 2014. On the elements of syntactic variation. In M. Carme Picallo (ed.), *Linguistic variation in the minimalist framework*, 13–35. Oxford: Oxford University Press.

Roberts, Ian. 2010. Introduction: Parameters in minimalist theory. In Teresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 1–57. Cambridge: Cambridge University Press.

Roberts, Ian & Anders Holmberg. 2005. On the role of parameters in Universal Grammar: A reply to Newmeyer. In Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz & Jan Koster (eds.), *Organizing grammar: Linguistic studies in honor of Henk van Riemsdijk*, 538–553. Berlin: Mouton de Gruyter.

Smith, Neil & Ann Law. 2009. On parametric (and non-parametric) variation. *Biolinguistics* 3(4). 332–343.

Snyder, William. 2001. On the nature of syntactic variation: Evidence from complex predicates and complex word-formation. *Language* 77. 324–342.

Starke, Michal. 2014. Towards elegant parameters: Language variation reduces to the size of lexically-stored trees. In M. Carme Picallo (ed.), *Linguistic variation in the minimalist framework*, 140–152. Oxford: Oxford University Press.

Stowell, Timothy A. *Origins of phrase structure*. Cambridge, MA: Massachusetts Institute of Technology PhD thesis.

Takano, Yuji. 1996. *Movement and parametric variation in syntax*. University of California Irvine PhD thesis.

Thornton, Rosalind & Stephen Crain. 2013. Parameters: The pluses and the minuses. In Marcel den Dikken (ed.), *The Cambridge handbook of generative syntax*, 927–970. Cambridge: Cambridge University Press.

Travis, Lisa. 1989. Parameters of phrase structure. In Mark Baltin (ed.), *Alternative conceptions of phrase structure*, 263–279. Chicago: University of Chicago Press.

Trotzke, Andreas, Markus Bader & Lyn Frazier. 2013. Third factors and the performance interface in language design. *Biolinguistics* 7. 1–34.

Trudgill, Peter. 2003. *The Norfolk dialect*. Cromer, UK: Poppyland Publishing.

Uriagereka, Juan. 1999. Multiple spell-out. In Norbert Hornstein & Samuel Epstein (eds.), *Working minimalism*, 251–282. Cambridge, MA: MIT Press.

Walkden, George. 2009. Deriving the Final-over-Final Constraint from third factor considerations. 5. 67–72.

Yang, Charles D. 2011. Three factors in language variation. In Theresa Biberauer, Cedric Boeckx & Anna-Maria Di Sciullo (eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 180–204. Oxford: Oxford University Press.