

# Rethinking Parameters

Edited by  
Luis Eguren,  
Olga Fernández-  
Soriano, and Amaya  
Mendikoetxea

## Rethinking Parameters

## OXFORD STUDIES IN COMPARATIVE SYNTAX

Richard Kayne, General Editor

*Movement and Silence*

Richard S. Kayne

*Restructuring and Functional Heads: The*

*Cartography of Syntactic Structures,*

*Volume 4*

Guglielmo Cinque

*Scrambling, Remnant Movement and*

*Restructuring in West Germanic*

Roland Hinterhölzl

*The Syntax of Ellipsis: Evidence from*

*Dutch Dialects*

Jeroen van Craenenbroeck

*Mapping the Left Periphery: The*

*Cartography of Syntactic Structures,*

*Volume 5*

Edited by Paola Benincà and Nicola

Munaro

*Mapping Spatial PPs: The Cartography of*

*Syntactic Structures, Volume 6*

Edited by Guglielmo Cinque and Luigi

Rizzi

*The Grammar of Q: Q-Particles, Wh-*

*Movement, and Pied-Piping*

Seth Cable

*Comparisons and Contrasts*

Richard S. Kayne

*Discourse-Related Features and*

*Functional Projections*

Silvio Cruschina

*Functional Heads: The Cartography of*

*Syntactic Structures, Volume 7*

Edited by Laura Brugé, Anna

Cardinaletti, Giuliana Giusti, Nicola

Munaro, Cecilia Poletto

*Adverbial Clauses, Main Clause*

*Phenomena and Composition of the Left*

*Periphery: The Cartography of Syntactic*

*Structures, Volume 8*

Liliane Haegeman

*Variation in Datives*

Edited by Beatriz Fernández and Ri-

cardo Etxepare

*Locality*

Edited by Ian Roberts and Enoch Aboh

*Aspects of Split Ergativity*

Jessica Coon

*A Comparative Grammar of Borgoma-*  
*nerese*

Christina Tortora

*Cross-Linguistic Studies of Imposters and*

*Pronominal Agreement*

Edited by Chris Collins

*Japanese Syntax in Comparative*

*Perspective*

Edited by Mamoru Saito

*Micro-Syntactic Variation in North*

*American English*

Edited by Raffaella Zanuttini and

Laurence R. Horn

*Functional Structure from Top to Toe: The*

*Cartography of Syntactic Structures,*

*Volume 9*

Edited by Peter Svenonius

*Chinese Syntax in a Cross-linguistic*

*Perspective*

Edited by Y.-H. Audrey Li,

Andrew Simpson, and W.-T. Dylan Tsai

*The Architecture of Determiners*

Thomas Leu

*Beyond Functional Sequence: The*

*Cartography of Syntactic Structures,*

*Volume 10*

Edited by Ur Shlonsky

*The Cartography of Chinese Syntax: The*

*Cartography of Syntactic Structures,*

*Volume 11*

Edited by Wei-Tien Dylan Tsai

*Argument Licensing and*

*Agreement*

Claire Halpert

*Rethinking Parameters*

Edited by Luis Eguren, Olga

Fernández-Soriano, and Amaya

Mendikoetxea

# RETHINKING PARAMETERS

Edited by Luis Eguren, Olga Fernández-Soriano,  
and Amaya Mendikoetxea

OXFORD  
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and certain other countries.

Published in the United States of America by Oxford University Press  
198 Madison Avenue, New York, NY 10016, United States of America.

© Oxford University Press 2016

First Edition published in 2016

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior permission in writing of Oxford University Press, or as expressly permitted by law, by license, or under terms agreed with the appropriate reproduction rights organization. Inquiries concerning reproduction outside the scope of the above should be sent to the Rights Department, Oxford University Press, at the address above.

You must not circulate this work in any other form  
and you must impose this same condition on any acquirer.

Library of Congress Cataloging-in-Publication Data  
Rethinking parameters / edited by Luis Eguren, Olga Fern?ndez-Soriano, and Amaya Mendikoetxea.

p. cm. — (Oxford Studies in Comparative Syntax)

Includes bibliographical references and index.

ISBN 978-0-19-046174-4 (pbk. : alk. paper) — ISBN 978-0-19-046173-7 (hardcover : alk. paper) — ISBN 978-0-19-046175-1 (ebook) — ISBN 978-0-19-046176-8 (online content)  
1. Principles and parameters (Linguistics) 2. Language and languages—Variation. 3. Language acquisition. 4. Grammar, Comparative and general—Syntax. 5. Generative grammar. 6. Reference (Linguistics) I. Eguren, Luis, editor. II. Fern?ndez Soriano, Olga, editor.

III. Mendikoetxea, Amaya, editor.

P295.5.R48 2016

415.01'822—dc23

2015021457

9 8 7 6 5 4 3 2 1

Printed by Sheridan Books, USA

# CONTENTS

*Contributors*      *vii*

*Introduction*      *1*

*Luis Eguren, Olga Fernández-Soriano, and Amaya Mendikoetxea*

## **PART I: The nature of variation and parameters**

1. Language variability in syntactic theory      49  
*David Adger*
2. Considerations pertaining to the nature of logodiversity      64  
*Cedric Boeckx*
3. Deriving linguistic variation from learnability conditions:  
The Chunking Procedure      105  
*Adriana Fasanella and Jordi Fortuny*
4. Lexical items and feature bundling: Consequences  
for microparametric approaches to variation      133  
*Ángel J. Gallego*
5. Some remarks on parameter hierarchies      170  
*Ian Roberts*

## **PART II: Parameters in the analysis of language variation: Case studies**

6. Parametric variation in nominal root compounding      203  
*Leah S. Bauke*
7. Macroparameters break down under the weight of evidence:  
The “NP/DP” Parameter as a case study      236  
*Youngmi Jeong*
8. The wh parameter and radical externalization      252  
*Éric Mathieu*
9. The Null Subject Parameter and the lexicon in minimalist  
syntax      291  
*Guido Mensching and Anja Weingart*

10. Rethinking parameter theory diachronically: A macrocomparative approach to the analyticization of the verbal tense-aspect systems of Brazilian Portuguese and Coptic Egyptian	329
<i>Chris H. Reintges and Sonia Cyrino</i>	

<i>Index</i>	387
--------------	-----

## CONTRIBUTORS

**David Adger** is Professor of Linguistics at Queen Mary University of London. He is author of *Core Syntax* (2003, Oxford University Press) and coauthor of *Mirrors and Microparameters* (2009, Cambridge University Press). His most recent book is *A Syntax of Substance* (2013, MIT Press). He is editor of Oxford Studies in Linguistic Theory, and edited the journal *Syntax* for eight years. His work focuses on the foundations of syntactic theory and the interfaces between syntax and other aspects of language.

**Leah S. Bauke** is a Junior professor in English Linguistics at the University of Wuppertal. She received her Ph.D. from the University of Frankfurt. Her main research interests are in morphosyntax and its interfaces with other components of grammar. She has worked on compounds, nominalization, and small clauses.

**Cedric Boeckx** is Research Professor at ICREA (The Catalan Institute for Advanced Studies) and a member of the Department of Linguistics at the University of Barcelona.

**Sonia Cyrino** is an Associate Professor in the Department of Linguistics, University of Campinas, Brazil. Her interests include syntactic theory, syntactic change, and Romance languages. She has articles published in books by Mouton de Gruyter, John Benjamins, Oxford University Press, and in journals such as the *Journal of Portuguese Linguistics*, the *Iberia-International Journal on Theoretical Linguistics*, and *Natural Language & Linguistic Theory*.

**Adriana Fasanella** defended her Ph.D. thesis, *On How Language Mechanisms Shape Natural Languages*, in the Universitat Autònoma de Barcelona in 2014. Her research has been focused on how morphophonological mechanisms of data analysis can predict morphosyntactic patterns of natural languages.

**Jordi Fortuny** Department of Catalan Philology. University of Barcelona.

**Ángel J. Gallego** is a Professor Agregat at the Departament de Filologia Espanyola of the Universitat Autònoma de Barcelona. He is member of the Centre de Lingüística Teòrica, a center of research on theoretical linguistics funded in



the early eighties. His main interests and publications concern the areas of syntax, comparative grammar, and parametric variation (especially within Romance languages).

**Youngmi Jeong** is doctor of Philosophy, University of Maryland.

**Eric Mathieu** is associate Professor at the Department of Linguistics of the University of Ottawa.

**Guido Mensching** is a professor of Romance linguistics and, since 2013, has held the chair of Romance linguistics at the Georg-August-Universität of Göttingen. His main fields of research are generative Romance syntax (see his *Infinitive Constructions with Specified Subjects: A Syntactic Analysis of the Romance Languages*, Oxford University Press, 2000) and historical linguistics and philology. The latter include the edition and analysis of medieval medico-botanical texts and medieval Judaeo-Romance.

**Chris H. Reintges** is a professor at CNRS and University Paris 7.

**Ian Roberts** is a professor at the Department of Linguistics of the University of Cambridge.

**Anja Weingart** is a researcher in Romance linguistics. After several assignments at the Freie Universität of Berlin, she currently works at the Georg-August-Universität of Göttingen. Her main fields of research are generative Romance syntax (in particular Spanish and Portuguese) and computing in the humanities.

## Rethinking Parameters



# Introduction

LUIS EGUREN, OLGA FERNÁNDEZ-SORIANO,  
AND AMAYA MENDIKOETXEA

## 1. EARLY PARAMETERS

The notion of parameter of syntactic variation is a key component of the theory of Principles and Parameters (P&P theory; Chomsky 1981, 1986; Chomsky and Lasnik 1993), a very productive research program in the history of generative grammar whose main objective is to account for the so-called logical problem of language acquisition, that is, how is it that the child learning a language comes to have a complex and subtle grammatical knowledge that goes far beyond the impoverished input in her linguistic environment? As is well-known, the solution P&P theory gives to the linguistic version of Plato's problem lies in devising a rich Universal Grammar (UG), the uniquely human genetic endowment for language. In a nutshell, the idea is that core aspects of the child's linguistic knowledge are given in advance by UG, and do not have to be learned at all.

Within the P&P framework, together with a pool of universal features, a series of invariant principles, and the specification of the general architecture, modules, and operations of grammar, UG provides a finite set of innately pre-determined choice points, or "parameters," each with a number of values (ideally two) (see also Chomsky, 1981: 11; 1986: 150):

"The theory of UG must meet two obvious conditions. On the one hand, it must be compatible with the diversity of existing (indeed, possible) grammars. At the same time, UG must be sufficiently constrained and restrictive in the options it permits so as to account for the fact that each of these grammars develops in the mind on the basis of quite limited evidence . . . What we expect to find, then, is

a highly structured theory of UG based on a number of fundamental principles that sharply restrict the class of attainable grammars and narrowly constrain their form, but with parameters that have to be fixed by experience.” (Chomsky 1981: 3–4)

In addition to the view that the domain of parametric options and the range of their values are restricted by UG, early parameters generally had two more defining traits. On the one hand, they were conceived as “grammatical” parameters, as Baker’s (2008a,b) dubs them,<sup>1</sup> which often concerned principles of grammar, but could also be related to a particular module (Theta Theory, Case Theory, etc.), or determine the choice of level (D-Structure, S-Structure, Logical Form) at which an operation of grammar applied. On the other hand, it was posited that parametric choices *may* have a great impact on the grammar of a language, giving rise to clusterings of syntactic properties:

“Each of the systems of (1) [subcomponents of the rule system of grammar] and (2) [subsystems of principles] is based on principles with certain possibilities of parametric variation . . . In a tightly integrated theory with fairly rich internal structure, change in a single parameter may have complex effects, with proliferating consequences in various parts of the grammar.” (Chomsky 1981: 6)

A good example of an early P&P theory parameter that was clearly endowed with these two attributes is Hale’s (1983) Configurationality Parameter (CP). The CP makes a division between configurational and non-configurational languages (e.g., English vs. Warlpiri). In the former, grammatical functions are linked to syntactic configurations; in the latter, syntactic configurations are, for the most part, lacking. This parameter was thought to have a cascade effect on the shape of grammars: the properties associated to non-configurationality include, according to Hale (1983: 5), free word order, the occurrence of discontinuous expressions, and the extensive use of “null anaphora” (i.e., non-overt representation of arguments), among many others. Moreover, this parameter was treated as a case of variation in grammar as a whole. Hale (1983) proposes, in particular, that configurationality is to be stated in terms of the Projection Principle (Chomsky 1981: 29), whereby the subcategorization properties of lexical items are represented at each syntactic level (D-structure, S-structure, and LF):

- (1) *The Configurationality Parameter*
  - a. In configurational languages, the Projection Principle holds of the pair (LS, PS).
  - b. In non-configurational languages, the Projection Principle holds of LS alone.

The CP in (1) locates the typological difference between configurational and non-configurational languages in the manner expressions in Phrase Structure (PS) are related to argument positions in what Hale names “Lexical Structure” (LS). In configurational languages, the Projection Principle establishes a bi-unique and structurally isomorphic relation between LS arguments and PS nominal expressions. In non-configurational languages, the CP does not determine any connection at all between LS and PS. From this parametric choice, Hale (1983) points out, the non-configurational properties of languages like Warlpiri would follow.

However, the original definition of the locus and scope of syntactic parametrization was not completely uniform, since not all proposed parameters were seen as grammatical options or had clustering effects. Take, for instance, two classical parameters: the Null Subject Parameter (NSP) and the Head Parameter (HP). The NSP differentiates between so-called *pro-drop* languages, like Spanish or Italian, which allow null subjects in tensed clauses, and *non-pro-drop* ones, like English or French, in which the subject position must be phonetically realized:

- (2) a. Ø vendrá. (Spanish)  
       Ø will-come.3sg  
       b. \*Ø will come. (English)

As in the case of the CP, early parametric studies also attributed a clustering effect to the NSP, by specifying that a number of superficially unrelated phenomena tend to correlate with the null subject property in *pro-drop* languages, most notably, the lack of expletive pronominal subjects (3a), free subject inversion (4a), and the possibility of extracting *wh*-subjects across an overt complementizer (5a) (cf. Chomsky 1981: 240; Rizzi 1982, and the references therein)<sup>2</sup>:

- (3) a. Ø llueve.  
       b. It rains.
- (4) a. Vendrá Juan.  
       b. \*Will come John.
- (5) a. ¿Quién<sub>i</sub> crees que t<sub>i</sub> vendrá?  
       b. \*Who<sub>i</sub> do you think that t<sub>i</sub> will come?

Chomsky (1981: §4.4) stated the NSP as a grammatical parameter, linking it to the Empty Category Principle, which constrained the occurrence of empty categories by positing that they must be properly governed. Under this view,

languages would vary as regards the governing properties of verbal inflection: in languages like Spanish (or Italian) verbal inflection properly governs the subject NP, whereas in languages like English (or French) verbal inflection is not a proper governor. However, Chomsky (1981: 241) himself also suggests that “... there is some abstract property of AGR, correlated more or less with overt morphology, which distinguishes pro-drop from non-pro-drop languages, from which the clusterings of properties follow.” In this line, Rizzi (1982: 130–131) claims that “... the characteristic property of null subject languages (NSLs) is that their verbal inflections have (clitic-like) pronominal properties” and adds that “this intuition can be straightforwardly implemented by assuming that INFL in NSLs is specified with the feature [+pronoun].”<sup>3</sup> Arguably, the NSP was thus formulated as a “lexical” parameter, again in Baker’s (2008a,b) terms, and not as a grammatical one (see fn. 1), almost from the start.<sup>4</sup>

The Head Parameter distinguishes head-initial languages (like Spanish), where heads uniformly precede their complements, from head-final ones (like Basque), systematically showing complement-head order:

- (6) a. [leer el libro]<sub>VP</sub> (Spanish)  
       read the book  
       b. [fotos de Juan]<sub>NP</sub>  
           pictures of John  
       c. [apropiado para la casa]<sub>AP</sub>  
           appropriate for the house  
       d. [sin dinero]<sub>PP</sub>  
           without money
- (7) a. [liburua-a irakurtzea]<sub>VP</sub> (Basque)  
       book-the read  
       b. [Jon-en argazkiak]<sub>NP</sub>  
           John-GEN pictures  
       c. [etxe-rako egokia]<sub>AP</sub>  
           house-for appropriate  
       d. [diru-rik gabe]<sub>PP</sub>  
           money-PART without

The HP is characterized in grammatical terms in Chomsky (1986: 82), who takes it to be a parameter of X-bar theory, which could be formulated as in (8) (cf. Biberauer 2008: 19)<sup>5</sup>:

- (8) *A parameterized principle in the phrase structure module*  
       a. Principle:  $X' \rightarrow X$  ; Complement (where ; signifies an unordered pair).  
       b. Parameter: Heads X precede/follow their complements.

This parameter certainly has a great impact on the shape of grammars, since once fixed in a given direction, it affects all combinations of heads and complements in a language, but, strictly speaking, it does not have a clustering effect: clusters of properties have been conceived as sets of *different* grammatical properties which are all meant to derive from one abstract property, whereas, given X-bar theory, the fact that verbs, nouns, adjectives, and prepositions precede, for instance, their complements is just one and the same property.

Putting all the observations in this section together, the following general picture obtains as regards the initial characterization of parameters of syntactic variation: first, parameters (and their values) were attributed to UG; second, they were *mostly* formulated as grammatical parameters, affecting principles, operations, or levels of grammar; and third, they *often* had a rich deductive structure, with sets of different formal properties clustering together.

This view on parametrization has changed significantly in the last three decades of intensive work on parametric syntax. On the one hand, it has been argued that parameters are not provided by UG. On the other hand, the locus of syntactic variation has also been shifted, with all parametric options now being placed in the (functional) lexicon and/or at the syntax-PF interface. And some authors have even claimed that the clusterings of properties associated with particular parameters do not really hold across languages.

In the next section, we will review the long-standing and lively debates on the nature, locus, and scope of parameters, and we will also address another currently debated topic in parametric syntax: the (non)-existence of parameter hierarchies. Throughout the discussion we will incorporate the core ideas in the papers collected in this volume with respect to these issues at suitable points in our exposition, with the aim of emphasizing the coherence of the different contributions regarding parameters and parametric variation.

## 2. CURRENT ISSUES IN PARAMETRIC SYNTAX

### 2.1. The status of parameters

The original P&P conception of parameters as part of UG can still be traced in the work by Baker (2005, 2008a,b) on macroparametric variation, where he supports an overdeterministic view of UG as regards some macroparameters, like the Agreement Parameters, which account for a number of agreement properties in Indo-European and Bantu languages (see section 2.3). Since their values, Baker argues, cannot be fixed using general learning devices, these macroparameters should be attributed to UG.

In contrast to Baker's stance on the nature of (at least some) parameters, there is a growing consensus in current minimalist theorizing building on the



idea that parametric options are not primitives of Universal Grammar, and are rather to be seen as emergent properties of grammars stemming from underspecified aspects of UG, such as linear ordering or the selection and bundling of formal features (Roberts and Holmberg 2005, 2010; Richards 2008a; Holmberg 2010; Boeckx 2011, this volume; Roberts 2012).

Roberts (this volume) extends this conception of the nature of parameters to parameter hierarchies (see section 2.4), which, in his view, are not prespecified by UG, but emerge from the interaction of (i) the fact that the formal features of certain heads are underspecified by UG (in Robert's view, UG makes available a certain set of features, but does not indicate how or whether they are deployed in a particular language), (ii) the Primary Linguistic Data, and (iii) general markedness conditions.

The insight that parameters do not belong to UG, but represent points of underspecification instead, clearly goes hand-in-hand with the attempt in the Minimalist Program (MP) to substantially reduce the content of the genetic endowment for language, which is now thought to basically consist of a closed inventory of lexical features plus the unique structure-building operation Merge (Chomsky 2000: 100; 2004: 107–108; 2005: 4; 2007: 5–6). As has been noted, this move is partially motivated by the minimalist search for a principled explanation of the properties of mental grammars relying on language-independent “third-factor” conditions, in particular, principles of structural architecture (i.e., interface conditions imposed on the linguistic cognitive system by the sensory-motor and conceptual-intentional systems with which it interacts), and conditions of efficient computation holding of computational systems such as I-languages (Chomsky 2004, 2005, 2007, 2008, 2010; Berwick and Chomsky 2011): only if UG is devoid of as much content as possible, including both principles and parameters, so the reasoning goes, core properties of the language faculty can be externally explained in third-factor terms.

Together with this theory-internal argument, the minimalist idea of an underspecified UG receives additional support from considerations on the evolution of language (Chomsky 2005, 2007, 2008, 2013; Boeckx 2012), by which a minimized UG is indispensable in order to offer a plausible explanation for the sudden and recent emergence of the particular properties that characterize human language. Putting it in Chomsky's (2005: 8) words, “. . . evidently, the more varied and intricate the conditions specific to language, the less hope there is for a reasonable account of the evolutionary origins of UG.”

The shift from a complex and intricate UG to an underspecified system whose properties can be externally determined is also observed in the theory of language acquisition. As mentioned in the previous section, a central idea in P&P theory is that the linguistic knowledge needed to make sense of the partial, limited, and degenerate linguistic input is basically

innate; it is a Universal Grammar (UG), which comprises a set of universal principles that do not have to be learned, and options left open by UG (or parameters), which must be set on the basis of sufficient linguistic experience. Moreover, if the number of parameters is finite, and relatively small, there is only a finite number of grammars in the child's learning space. As Lorenzo and Longa (2009: 1302) phrase it: "the resulting notion of 'parameter setting' is thus a paradigmatic instantiation of the idea of 'selective learning' (Piattelli-Palmarini 1989), that is, the contention that experience is but a trigger of pre-existing resources of organisms instead of an instructor of their development."

The point to be addressed is then whether Minimalism has altered in any way this view of language acquisition, since, as explained above, the advent of the Minimalist Program has weakened the strong geneticism of P&P theory mostly through the introduction of third-factor principles, but also due to the role attributed to UG and the nature of the principles which conform it. To mention just some views on this issue, Yang and Roeper (2011: 552) argue that the answer to this question is both "no" and "yes." It is negative in the sense that Minimalism has not managed to provide the basic P&P architecture for the task of language acquisition, and positive in the sense that the new conception on the language faculty has led to new conceptions on learning which may lead to a more complete explanation of the mechanisms of language acquisition. A more negative approach is presented in Longa and Lorenzo (2008), who claim that the shift in how we judge the explanatory adequacy of the principles of the language faculty, as optimal solutions for the needs of the cognitive systems that interact with it, has been mostly ignored by researchers in child language acquisition. In a later paper, these authors (Lorenzo and Longa 2009: 1308) claim that the MP must shift its perspective in two crucial aspects in order to come close to a rigorous approach to language acquisition: (i) the recognition that the contribution of the environment to language acquisition should not be restricted to those aspects in which languages differ (see Tomasello 2003 for usage-based approaches to acquisition), and (ii) the acceptance that stages in child language acquisition are a source of information and play a role in the development of mature language systems.

One of the papers in this volume (Fasanella and Fortuny's) approaches the issue of language acquisition from a minimalist perspective, emphasizing that a parametric model must not only account for linguistic variation but must also provide the elements which guide the child in developing her linguistic knowledge. These authors explicitly reject the idea that parameters are coded in UG, and propose instead that the clusters of syntactic properties associated to traditional parameters can be derived from the conjunction of third-factor procedures of data analysis and bootstrapping mechanisms operating in the process of language acquisition.

## 2.2. The locus of parametric variation

A second major debate in parametric syntax, which is related to the discussion above on the (non)-UG-based status of parameters, has to do with the determination of the components of language with which parameters are associated. Three main lines of thought can be identified in generative research on variation in this respect: (a) what might be called the “Grammatical Parametrization Hypothesis” (GPH) of early P&P theory, whereby parameters are linked to general properties of grammars (i.e., principles, operations, modules, and levels); (b) the subsequent Functional Parametrization Hypothesis (FPH), also first proposed within P&P theory and then widely adopted in the first stages of the Minimalist Program, which identifies the functional lexicon as the locus of parametric options; and (c) the Externalization Hypothesis (EH), currently being advocated within the MP framework, which reduces parameters to cross-linguistic differences arising at the PF branch of grammar (see, e.g., Chomsky 2010, Berwick and Chomsky 2011, Berwick et al. 2013).<sup>6</sup>

The GPH has been described and illustrated in section 1. We will now outline the motivations and scope of the other two approaches to the location of parameters, comparing them both with each other and with the GPH, and focusing on certain issues that are still under discussion.

The hypothesis that parameters find their source in the functional lexicon was first explicitly advocated by Borer (1984) and further developed in Fukui (1986, 1988), Ouhalla (1991), and Webelhuth (1992), among many others. Fukui (2006: 108 [1995]) formulates this idea as in (9)<sup>7</sup>:

(9) *The Functional Parametrization Hypothesis*

Only [+F] elements in the lexicon are subject to parametric variation.

As shown by the following remarks by Chomsky in the introduction to *The Minimalist Program* (see also Chomsky 2001: 2), the FPH was soon incorporated into the MP<sup>8</sup>:

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. One proposal is that parameters are restricted to *formal features* with no interpretation at the interface. A still stronger one is that they are restricted to formal features of functional categories (see Borer 1984, Fukui 1986, 1988) . . . I will assume that something of this sort is correct . . .” (Chomsky 1995: 6)

In compliance with the FPH, and reinforcing a trend initiated in P&P theory, most work on syntactic variation within the MP framework has thus reformulated classical parameters, and defined new ones, as properties of functional

heads. Let us illustrate this fact with two well-known parameters: the Head Parameter (HP) and the *Wh*-Parameter.

As mentioned in section 1, the HP was initially viewed as a grammatical parameter in Chomsky (1986), and also in Koopman (1984)/Travis (1984), who linked it to the principles of X-bar Theory and the thematic and Case modules, respectively. Early proponents of the FPH, like Borer (1984) and Fukui (1988), left the HP outside the scope of this hypothesis, since, at that point, they could not envisage how to reduce cross-linguistic variation in the linearization of heads and complements to properties of lexical items. A way to do so has been articulated, however, within the MP building on Kayne's (1994) Antisymmetry of Syntax, which implies that all languages have an underlying head-complement order, so that surface head-final orderings must be the consequence of leftward movement of complements to the specifier of some functional head. Within this framework, the difference between head-initial and head-final languages would be, in accordance with the FPH, that the lexicon of the latter contains a set of functional categories, which are lacking in the former, bearing a movement-triggering feature which attracts the complement to the left of the head (see, e.g., Bouchard 2003: 4; Biberauer 2008: 25; Kayne 2011: 2).

The *Wh*-Parameter sets apart languages where *wh*-phrases in simple *wh*-questions are fronted (English or Spanish) from those in which they stay *in situ* (Chinese or Japanese). This parameter was first formulated in grammatical terms by Huang (1982a,b), who put forward the idea that these two types of languages differ in the level of representation at which *wh*-movement operates: in languages like English, *wh*-phrases move overtly at S-structure, whereas in languages like Chinese *wh*-movement applies covertly at Logical Form.

Various characterizations of the *Wh*-Parameter conforming to the FPH have been delineated in the minimalist literature as well. Chomsky (1995: 232, 289, 291) claims, for instance, that in languages like English *wh*-movement is motivated by the presence of a “strong” Q-feature on the interrogative complementizer, which must be checked before Spell-Out, whereas in languages with a “weak” Q-feature on C, like Chinese, the *wh*-phrase remains *in situ*, and is interpreted via unselective binding. Having discarded featural strength as an analytical tool, Chomsky (2000: 109) further attributes the difference in the structural position of *wh*-phrases between languages like English and Chinese to the presence or absence of an uninterpretable EPP-feature on C, requiring [Spec, CP] to be filled: under this view, languages with an EPP-feature on C will require *wh*-movement, and languages without it will show *wh-in-situ*. Still another analysis of the *Wh*-Parameter is developed by Cheng and Rooryck (2000). These authors, elaborating on the proposal in Cheng (1991), correlate the availability of *wh*-particles in languages like Chinese or Japanese with *wh-in-situ*, assuming that, in this type of languages, the

*wh*-particle itself checks the Q-feature of C, rendering *wh*-movement unnecessary (and therefore impossible); in languages without *wh*-particles, like English, the Q-feature of C could then only be checked by moving a *wh*-phrase to [Spec, CP] instead.

The FPH has been argued to have a number of advantages in comparison to the GPH (see, e.g., Roberts and Holmberg 2010: 32–33; Roberts, this volume), all following the simplicity requirement of Methodological Minimalism, which might explain why it has been standardly assimilated into the MP. This hypothesis first imposes both a strong limit on what can vary and a restriction on the form of parameters; second, confining syntactic variation to the (functional) lexicon is in any case needed to account for “microvariation” phenomena (see section 2.3); and third, as initially pointed out by Borer (1984: 29), “associating parameter values with lexical entries reduces them to the one part of a language which clearly must be learned anyway: the lexicon.” To these arguments in favor of the FPH, we can add its compatibility with the minimalist view of an underdeterministic UG. Once parametric variation is associated to the properties of lexical items, and not to supposedly innate global aspects of grammars (as in the GPH), we no longer need to ascribe it to UG: UG would now only provide a set of universal features, and the (functional) lexicon of a language would be conformed by selecting certain features from this universal pool and assembling them into language-particular items (Chomsky 2000: 101), the selection and assembling of features being left underspecified by UG.

As already mentioned, a third line of thought on the locus of parameters is currently being developed in minimalist theorizing confining most (if not all) syntactic variation to the process of externalization (i.e., “the mapping from internal linguistic representations to their ordered output form, either spoken or manually gestured”; Berwick et al. 2013: 89), which takes place at the PF branch of grammar (i.e., the post-syntactic morphological and phonological components) (see, e.g., Kandybowicz 2009, Chomsky 2010, Berwick and Chomsky 2011, Berwick et al. 2013).<sup>9</sup> Berwick and Chomsky (2011) present this view as follows (see Chomsky 2010: 60, and Berwick et al. 2013: 92, for similar remarks):

“Externalization is not a simple task. It has to relate two quite distinct systems: one is a sensory-motor system that appears to have been basically intact for hundreds of thousands of years; the second is a newly emerged computational system for thought, which is perfect insofar as the strong minimalist thesis is correct. We would expect, then, that morphology and phonology—the linguistic processes that convert internal syntactic objects to the entities accessible to the sensory-motor system—might turn out to be quite intricate, varied, and subject to accidental historical events. Parametrization and diversity, then, would be mostly—possibly entirely—restricted to externalization” (Berwick and Chomsky 2011: 37).

Boeckx (this volume), following the insights in Boeckx (2011), fully agrees with the Externalization Hypothesis (EH) above. In line with Chomsky's (2001: 2) Uniformity Principle in (10), he proposes the Strong Uniformity Thesis (SUT) in (11), which strengthens Berwick and Chomsky's position, by eliminating the "mostly" and "possibly" from their formulation of the EH:

(10) *Uniformity Principle*

In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances.

(11) *Strong Uniformity Thesis*

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

Under the SUT, which, as Boeckx indicates, blocks the possibility of an indirect parametrization of syntax through the elimination of pre-syntactic lexical parameters such as bundling parameters, variation would be *entirely* restricted to externalization, leading to the following statement on the options for parametrization:

(12) *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 component).

Like the FPH, the EH is fully compatible with the minimalist reduction of the genetic endowment for language (UG), given that the process of externalizing internal computation can well be claimed not to be part of what Chomsky and his colleagues call "the faculty of language in the narrow sense" (FLN), which is meant to comprise just properties that are both unique to humans, and to language itself (Hauser, Chomsky, and Fitch 2002; Fitch, Hauser, and Chomsky 2005). In Berwick and Chomsky's (2011: 38) words, "we have no reason to suppose that solving the externalization problem involved an evolutionary change—that is, genomic change . . . it follows that externalization may not have evolved at all; rather, it might have been a process of problem solving using existing cognitive capacities."

The difference between these two hypotheses on the location of parameters thus lies in their scope: that is, the EH covers a second dimension of variation, which the FPH does not contemplate. Together with variation in the morpho-syntactic features of lexical items, as in the FPH, which is now (mostly) relegated to a post-syntactic morphological component like the one proposed in Distributed Morphology (DM),<sup>10</sup> the EH postulates that there also exist core

cases of cross-linguistic variation at the level of phonological representation. These include, among other phenomena, linearization, overt or null realization of lexical items, (non)-affixal requirements of syntactic heads, and the pronunciation of higher or lower copies in movement chains (see fn. 9). In what follows, we will again show how this kind of variation has been expressed by reviewing some recent PF-accounts of both the Head Parameter and the *Wh*-Parameter.

As an alternative to the definition of the Head Parameter in terms of the presence or absence of a movement-triggering feature on functional heads, which has been criticized as being an *ad hoc* solution (see, e.g., Bouchard 2003: 7), it has been argued that variation in the relative ordering of heads and complements is a PF-mapping strategy pertaining to the externalization process. This strategy would ultimately result from a third-factor condition, the fact that the physics of speech demand that linguistic units must be pronounced sequentially in time, giving rise, in this case, to just two options: the head either precedes or follows its complement (see Bouchard 2003; Richards 2004, 2008a,b; Baker 2005; Holmberg 2010; Chomsky 2010; Berwick and Chomsky 2011; Berwick et al. 2013).

Regarding the *Wh*-Parameter, Mathieu (this volume), inspired by Richards's (2010) PF-based approach to *wh-in-situ* versus *wh*-movement, claims that this parameter is reducible to differences in prosodic properties between languages, thus supporting what he calls "radical externalization" of previous "strength" parameters (see above). Mathieu's insight is that *wh-in-situ* languages tend to be languages that mark focus prosodically in a demarcative way, while *wh*-movement languages tend to be languages that express focus culminatively. Under this proposal, cross-linguistic variation in the location of *wh*-phrases would be a syntax-external phenomenon constrained by the phonology of the language.

As shown by these PF reformulations of the Head Parameter and the *Wh*-Parameter, the EH offers a new perspective on time-honored parameters. The Externalization Hypothesis is not, however, unanimously accepted: a number of proposals can be found in the generative literature claiming that parametrization not only obtains at the externalization process, but can also affect other components or aspects of language, including both the computational and the semantic components, the pre-syntactic lexicon, and even third-factor conditions. To close this section, we will go through some of these proposals, examining whether they are a real challenge to the EH, or whether the facts they apply to are amenable to a different analysis instead.

Parametric options have been argued to arise within core syntax or in relation to conditions of efficient computation by Saito and Fukui (1998) and Baker and Collins (2006), respectively. Saito and Fukui (1998: 452) incorporate the effects of the Head Parameter under a parameterized



definition of Merge specifying whether there is projection of the left or the right element in an ordered pair. Baker and Collins (2006: 333) hold that the Minimal Link Condition (MLC), which is taken to be a third-factor principle of computational efficiency in current Minimalism, can be parametrized in order to capture the fact that in some languages, like Kinande, any nominal constituent of the VP can move to [Spec, Linker Phrase] (a *v*P-internal functional projection), whereas only the highest DP can move in other languages, like Hoan.

As indicated above, a better-founded PF-based characterization of the HP is, nevertheless, now available, which can replace Saito and Fukui's proposal. As for Baker and Collins' MLC-Parameter, both Richards (2008a) and Boeckx (2011) consider that it is rather implausible that language-independent third-factor conditions can be subject to variation. In addition to this general criticism, a different explanation for the linker data Baker and Collins analyze is offered in Jeong (2006), who accounts for the free ordering of internal arguments that motivated the parametrization of the MLC by resorting to the applicative typology in the context of linkers, so that high (VP-external) applicative structures provide the source of freer word ordering, as in Kinande, whereas low applicative structures impose a strict ordering, as in Hoan.

Semantic parameters have been advocated in work by Chierchia (1998) and Snyder (2012), among others.<sup>11</sup> Chierchia develops a hypothesis, which he calls the "Nominal Mapping Parameter" (NMP), whereby languages can vary with respect to the predicative or argumental nature of their nouns. Under the NMP, some languages (e.g., Romance languages) only have predicative nouns (denoting properties), which must obligatorily combine with a determiner in argumental positions, whereas in other languages, like Chinese, nouns are argumental (names of kinds) and can thus function as arguments on their own (i.e., without a DP projection).<sup>12</sup> Chierchia's semantic parameter thus allows either NPs or DPs to be argumental. However, building on Longobardi's (1994) proposal that a nominal expression is an argument only if it is introduced by a category D, an alternative to the NMP can be envisaged, in accordance with the EH, whereby argumental nominal expressions project a DP in all languages, and the cross-linguistic difference just lies in whether the D position has phonetic content or is filled by an empty category.<sup>13</sup>

Another semantic parameter has been recently proposed by Snyder (2012) as the latest version of the Compounding Parameter (Snyder 1995, 2001), which differentiates languages that have productive, recursive, and compositional nominal root compounding, like English, from those that do not, like Spanish.<sup>14</sup> Snyder now considers that this parameter is to be understood as the availability in English-like languages of a specific rule of semantic composition operating at the syntax-semantics interface, which



he terms “Generalized Modification” and which is unavailable in Spanish-like languages:

(13) *Generalized Modification*

If  $\alpha$  and  $\beta$  are syntactic sisters under the node  $\gamma$ , where  $\alpha$  is the head of  $\gamma$  and if  $\alpha$  denotes a kind, then interpret  $\gamma$  semantically as a subtype of  $\alpha$ 's kind that stands in a pragmatically suitable relation to the denotation of  $\beta$ .

Discussing Snyder's (2012) semantic characterization of the Compounding Parameter, Boeckx (this volume) critically points out that “in a more restrictive semantic framework such as the one put forth in Pietroski (2005) and Hinzen (2007), the limited repertoire of interpretive operations makes it impossible for languages to ‘deactivate’ some of them,” giving rise to variation in the semantic component. Boeckx thus concludes that “it is difficult to see how a language would be able to completely do away without as general a rule as Generalized Modification.” Moreover, once again, alternative lexical analyses of this parameter have also been provided in the generative literature, conforming (at least) to the FPH (see fn. 18 later). And recursivity of compounds has even been argued to be determined by word-stress location (an externalization mechanism) in Tokizaki (2010).

Finally, the pre-syntactic lexicon has also been identified as the locus of a particular type of parametrization in Gallego (2011). This author argues that lexical variation can have two different sources: (i) the way the outputs of syntactic computation are spelled out (variation *after* syntax, conforming to the EH), and (ii) the way features provided by UG are assembled into language-particular lexical units (variation *before* syntax, which, in his view, is needed to account for clustering effects).

Gallego (this volume) retakes the idea that some lexical variation patterns (i.e., those resulting from feature-bundling) do not fit in well with the EH. He mentions in this respect a number of syntactic phenomena, like the lack of VSO sentences in Catalan, the fact that Serbo-Croatian has multiple *wh*-movement, or the insular status of indicative dependents in Polish (as opposed to Spanish in all three cases), which can be argued to depend on the (pre-syntactic) selection and bundling of particular features in particular languages. A different view can be found in Boeckx (this volume), who, in compliance with the Strong Uniformity Thesis in (11), suggests that “instead of talking about pre-syntactic lexical bundles, we can just as easily talk about post-syntactic morpho-phonological bundles.”

As can be inferred from the discussion in the last part of this section, a pending task for parametric inquiries is to determine whether *all* cases of parametrization originate in the externalization process, being reducible to properties of the morphophonological component, or whether there are also

points of variation in other components (or aspects) of language, which ought to be analyzed in a different way.

### 2.3. Macroparameters and microparameters

Research on syntactic variation within chomskyan linguistics has followed two different (and coexistent) paths. On the one hand, a number of scholars have concentrated on discovering parameters with widespread consequences on grammars (i.e., “macroparameters” and standard medium-sized parameters, or “medioparameters,” as Baker 2008b calls them). On the other hand, a great deal of work has also been dedicated to the study of “microparameters.” In this section, we will review these two trends in parametric syntax, as well as the ongoing debate on the existence of (macro)parameters<sup>15</sup> and their interaction with microparameters.

Three main criteria have been invoked in the literature on parametric variation to differentiate between macroparameters, medium-sized parameters, and microparameters: the extent of variation, the methodology of comparison, and the locus of variation (see Baker 2008b). With respect to the extent of variation, macroparameters have been characterized as having a global effect on the shape of grammars, whereas medium-sized parameters are not as pervasive in their influence as macroparameters are taken to be, but still have “a fairly significant effect on the overall feel of a language” (Baker 2008b: 352), and microparameters constitute small-scale differences amongst grammars. Under the second criterion, it is generally assumed that macroparameters result from the comparison of historically unrelated languages, while medium-sized parameters apply to languages of the same family (or to families of languages), and microparameters are tied to specific constructions in very closely related languages or dialects of the same language (cf. Kayne 2005: 8–10). Finally, according to Baker (2008b), in particular, macroparameters are to be distinguished from microparameters (and also medioparameters, like the NSP; see Baker 2008b: fn. 1) in that they concern principles of grammar, and do not have a lexical source, whereas microparameters (and medioparameters) can be formulated as properties of the (functional) lexicon.

Appealing to the extent of variation to isolate the three types of parametric options above can be useful as a descriptive tool. However, as has often been pointed out, this criterion meets a serious problem: no independent measure has ever been given which unequivocally defines how big (or small) the (clustering) effect of a particular parameter has to be in order for it to be considered an instance of macro-, medio-, or microvariation. The difference between macroparameters, medium-sized parameters, and microparameters thus basically lies in the methodology of comparison that is adopted, and, for some authors, also in how the different patterns of variation are explained.<sup>16</sup>

Macroparametric studies have as their major exponent the proposals by Mark Baker (1996, 2008a,b). In a monograph that lays the foundations of this area of research, Baker (1996) first posits that, alongside the Head Parameter, one more macroparameter can be identified that isolates polysynthetic languages, in which a single verb, built up of many parts, conveys the information expressed by a whole sentence in non-polysynthetic languages. He informally formulates this parameter as in (14), where the “morphemes” in the definition are either agreement morphemes or incorporated roots (Baker 1996: 11):

(14) *The Polysynthesis Parameter*

Every argument of a head element must be related to a morpheme in the word containing that head.

In an attempt to develop (14) into a precise principle, Baker further proposes the condition on  $\theta$ -role assignment in (15) as the distinctive property of polysynthetic languages (Baker 1996: 17):

(15) *The Morphological Visibility Condition*

A phrase X is visible for  $\theta$ -role assignment from a head Y only if it is coindexed with a morpheme in the word containing Y via:

- (i) an agreement relationship, or
- (ii) a movement relationship

Yes: Mohawk, Nahuatl, Mayali . . .

No: English, French, Chichewa . . .

To support his view that the Polysynthesis Parameter is a macroparameter, Baker argues that the condition in (15) underlies a wide range of properties which are typical of polysynthetic languages, such as syntactic noun-incorporation, rich object and subject agreement, free *pro-drop* of all arguments, and free word order, among many others (see Baker 1996: 498–499).<sup>17</sup>

Besides the Head Parameter and the Polysynthesis Parameter, Baker (2008a,b) proposes that two new macroparameters emerge from the comparison of Niger-Congo languages and Indo-European languages:

(16) a. *The Direction of Agreement Parameter*

F agrees with DP/NP only if DP/NP asymmetrically c-commands F.  
(Yes: Niger-Congo languages; No: Indo-European languages)

b. *The Case-Dependency Parameter*

F agrees with DP/NP only if F values the Case feature of DP/NP or vice versa.

(No: Niger-Congo languages; Yes: Indo-European languages)

The idea behind these parameters is that agreement-bearing functional heads behave differently from one language to another with respect to the conditions under which they agree: in Niger-Congo languages, the agreed-with NP must be higher in the clause than the agreeing head, whereas in Indo-European languages agreement is subject to the condition that the two elements match in Case. Baker again takes these two agreement parameters to be macroparameters, since they seem to hold of all potential agreeing functional categories (C, T, *v*, P, D, etc.) in these two families of languages.

Finally, another parameter that might well qualify as a macroparameter is the High Analyticity Parameter delineated in Huang (2006, 2010), who observes that an array of different properties appear to cluster together in Modern Chinese, as compared to English and other languages, and proposes that these properties are all manifestations of a single macroparameter, stating that Chinese lexical items are highly analytic at the lexical, functional, and argument structure levels.

Three well-known medium-sized parameters, to which some of the papers included in this volume pay attention (see below in this section), are the Null Subject Parameter, Snyder's Compounding Parameter, and Bošković's NP/DP Parameter. The NSP has already been presented in section 1. We will now focus on the Compounding Parameter and the NP/DP Parameter.

As indicated in our discussion on semantic parameters in the previous section, the Compounding Parameter (CP) (Snyder 1995, 2001, 2012; Roeper, Snyder, and Hiramatsu 2002; Roeper and Snyder 2005) reflects the fact that languages differ in whether they allow endocentric, nominal root compounding as a fully creative process or not (e.g., English: *university lab space committee* versus Spanish: *\*comité espacio laboratorio universidad*).<sup>18</sup> Snyder and his colleagues have always thought that this parameter has a solid clustering effect. Snyder (1995, 2001) argues, in this respect, that the positive setting of the CP in a language strongly correlates with the presence of the verb-particle construction (*Mary pulled the top off*), the adjectival-resultative construction (*John wiped the table clean*), and the double-object construction (*Alice sent Sue the letter*). Roeper and Snyder (2005) further point out that there is also a connection between the property which gives rise to recursive root compounding (the availability, in their view, of an Abstract Clitic Position as the complement to a lexical category) and *make*-causatives (*make John buy the book*), bare-V/N idioms (*pay attention*), middles (*This book reads easily*), and null-P constructions (*jump (over) the fence*).

The NP/DP Parameter (Bošković 2005, 2008, 2009; Bošković and Gajewski 2011) establishes that Traditional Noun Phrases (TNPs) may correspond to either a DP or an NP, so that languages with articles (English, Bulgarian, Macedonian . . . ) project a DP, whereas languages without articles (Serbo-Croatian, Russian, Polish . . . ) project an NP. This parameter is again endowed with a rich deductive structure. Bošković (2008) claims, in particular, that a single

difference between the two types of languages (the presence or absence of the definite article) lies behind a good number of cross-linguistic generalizations concerning very different grammatical phenomena: in his view, (i) only languages without articles may allow left-branch extraction; (ii) only languages without articles may allow adjunct extraction out of TNPs; (iii) only languages without articles may allow scrambling; (iv) negative raising is disallowed in languages without articles, while those with articles allow it; (v) languages without articles do not show superiority effects in multiple *wh*-fronting; (vi) only languages with articles may allow clitic doubling; (vii) languages without articles do not allow transitive nominals with two genitives; (viii) only languages with articles allow the majority superlative reading; (ix) head-internal relatives display island sensitivity in languages without articles, but not in those with articles; and (x) polysynthetic languages do not have articles.<sup>19</sup>

The second, very productive, direction in parametric syntax, the microparametric approach, focuses on small differences between related languages (and dialects), and explains these differences in terms of lexical properties (see Black and Motapanyane 1996; Kayne 2000, 2005, 2013; Barbiers 2009; among many others). These localized points of variation are typically tied to particular constructions, as in the work by Barbiers (2009) on microvariation in Dutch dialects with respect to complementizer drop, ONE-insertion, strong reflexives, and doubling in *wh*-chains. Research on microparameters may even involve a single lexical item, as illustrated by Kayne's (2005) fine-grained analysis of the syntactic properties of quantity words in English and French. As Richard Kayne, one of the central figures in microcomparative research, points out, a major advantage of this kind of study is that, by examining closely related languages or dialects, most properties of the grammars involved are kept constant, and we can safely determine whether a change in a specific property parametrically corresponds to the change in another one, or not (Kayne 2000: 5–6; 2005: 8). This way, so Kayne argues, we would be closer to discover what the minimal units of syntactic variation are: putting it in his own words, “microparametric syntax is a powerful tool, whose growth is perhaps to be compared with the development of the earliest microscopes, that allows us to probe questions concerning the most primitive units of syntactic variation” (Kayne 2000: 9).

As mentioned in the beginning of this section, the role (macro)parameters play in accounting for syntactic variation is currently under discussion. The debate is centered on two core issues: (a) do (macro)parametric effects really exist?, and (b) can (macro)parametric differences between languages be ultimately reduced to microparametric ones? We will deal with these two questions in turn.

The controversy on the existence of (macro)parameters has been instigated by Newmeyer's (2004, 2005) claim that the clustering effects associated to

these parametric options fail to hold when a wider variety of languages is taken into consideration (see also Haspelmath 2008, Boeckx 2011). In support of this view, Newmeyer brings up the case of the Null Subject Parameter (NSP). Relying on Gilligan's (1987) study, which tested the correlations put forward by Rizzi (1982) against a 100-language sample, he holds that the cluster of properties that was supposed to follow from the NSP (i.e., the possibility of null thematic subjects in tensed clauses, null expletive subjects, subject inversion in simple sentences, and *that*-trace violations; see the examples in (2)–(5) in section 1) has been shown to be an illusion. Newmeyer thus advocates abandoning parameters altogether, and replacing them by a model in which languages particular differences are captured by differences in language-particular rules.

The ideas in Newmeyer (2004, 2005) have been contested by a number of linguists (see, e.g., Roberts and Holmberg 2005, Biberauer 2008, Holmberg 2010, Roberts and Holmberg 2010), who first criticize the use of language-specific rules to express cross-linguistic differences as a “retrograde step” (“a return to the rule-based system of the 1960s and 1970s”), which represents a retreat from explanatory adequacy, and also predicts that languages may vary at random, contrary to the facts. All these authors therefore consider that the notion of parameter, which sets limits to syntactic variation by grouping together complex cross-linguistic differences, is to be maintained, although with major refinements and provisos, like the following: (a) parameters must be defined in a more adequate and precise way; (b) weaker clusterings and correlations could be envisaged; (c) a more articulated view of the structure of a parameter, including both sub-parameters and microparameters, has to be adopted; and (d) when coming up to what appears to be a counterexample to the expected clustering effect of a parameter, it should always be kept in mind that parameters can interact, so that the clustering at hand can be distorted by the interference of other parametrical differences between the languages being compared.

A good example of this new, more sophisticated, look at (macro)parametric variation can be found in Roberts and Holmberg's (2010) revision of Gilligan's (1987) study on the NSP, from which they conclude that Newmeyer's claim that results such as Gilligan's invalidate the parametric cluster proposed in Rizzi (1982) is not warranted. Roberts and Holmberg come to this conclusion basically for three reasons (see Roberts and Holmberg 2010: 23). First, we do not know enough about a good number of problematic languages in order to be sure that they are genuine counterexamples. Second, as Gilligan himself acknowledges, the non-obvious implication that languages with free inversion allow complementizer-trace violations remains. And third, a more modest implicational hierarchy can be designed on the basis of Gilligan's results showing that Rizzi's cluster still holds to some extent (i.e., Free Inversion  $\rightarrow$  (allow *that*-trace violations  $\rightarrow$  expletive null subjects)), which defines three types of

languages: type I (Italian, Spanish) would have all these properties, type II (Cape Verdean, Berbice Dutch) allows complementizer-trace violations and hence allows expletive null subjects, and type III (Haitian, basilectal Jamaican) only allows expletive null subjects. To these considerations Roberts and Holmberg (2010: 19) add that, in analyzing the behavior of a particular language with respect to the NSP, it must be taken into account that independent cross-linguistic differences can disguise its effects. They illustrate this fact with the case of the Celtic languages Welsh and Irish. Both languages are null subject languages with rich agreement inflection on the finite verb. However, the status of the correlation with the absence of complementizer-trace effects is difficult to evaluate in this case, since subjects are never adjacent to finite complementizers in VSO languages, like Welsh or Irish. VSO order thus neutralizes this property, and we do not have here a true counterexample to the proposed cluster, so Roberts and Holmberg argue.

Turning our attention now to the debate on the relation between (macro)parameters and microparameters, it seems clear that (macro)parametric effects must be complemented with microparametric ones in any event. Only in this way can we capture the complex empirical facts that pervasively show up when the correlations associated to a particular (macro)parameter are examined in different languages. The Head Parameter (HP), for instance, can certainly be set in a uniform way for all categories, developing harmonic languages. However, a number of “mixed” languages are also attested, like German, in which the head precedes its complements in certain phrases and follows them in others, and there also appear to be languages which exhibit head-directionality of one kind in the clausal domain and the opposite in nominals, and even cases arise where individual lexical items determine which one of the two parametric options is chosen (see the discussion and examples in Biberauer 2008: 10–12). To account for all these “deviant” patterns of variation, we must surely resort to microparameters. Though there seems to be no discussion on this point, the debate lies somewhere else. It has to do with a far-reaching question: can (macro)parameters be fully reduced to microparameters?

This question has also been answered in different ways. On the one hand, Kayne (2005) holds that apparent macroparametric differences are always the result of the cumulative effect of microparametric ones, and hence conjectures that it might turn out that “every parameter is a microparameter” (Kayne 2005: 10; see also Kayne 2013: fn. 23). On the other hand, Baker (2008b) supports the idea that macroparameters exist alongside microparameters. Baker admits that macroparameters could well be formulated as a set of microparameters: instead of stating, for instance, that heads precede their complements in a particular language (the HP), we could equally say that verbs precede their complements, nouns precede their complements, and so on for each category (or for each item of each category). However, in his view, the relevant issue is whether the large-scale correlations associated



to macroparameters can always arise from the addition of smaller-scale differences between languages or not. He claims, at this respect, that “a system in which all syntactic variation is ascribed solely to microparametric differences in the make-up of functional heads, as assumed under the so-called Borer-Chomsky Conjecture [see fn. 8], cannot account for actually attested patterns of variation” (Baker 2008b: 351). To justify this statement, Baker focuses on the Head Parameter, as specified below.

Baker (2008b) argues that if there were only head-directionality microparameters, which are totally independent from each other, there should be many mixed languages and relatively few consistently head-initial or head-final languages, contrary to what we actually find: there seems to be more pure head-initial or head-final languages than there are languages with a mixture of the two orders. Baker illustrates this fact with the statistics on the internal order of verb phrases and adpositional phrases across languages, showing that there are many more consistent languages (V-O and P-NP order, or O-V and NP-P order) than inconsistent ones (V-O and NP-P order, or O-V and P-NP order), and further offers more statistical information manifesting that homogeneous languages are much more common than non-homogeneous ones also in the case of the Polysynthesis Parameter and the Agreement Parameters. This, Baker says, is the scenario that is expected under a view that includes macroparameters as well as microparameters: the setting of macroparameters would result in the prevalence of harmonic systems, and microparameters could then conceal the effect of a macroparameter, giving rise to intermediate “noisy” cases, which must be relatively few, because, in his words, “it takes a whole series of microparametric choices all set in a certain way to override or disguise the effects of a single macroparametric choice.” Mark Baker then concludes, *contra* Kayne, that together with microparameters, there are also global macroparameters that regulate languages as a whole, which cannot be reduced to microparametric choices in the featural content of individual lexical items, as required by the Borer-Chomsky Conjecture.

Ian Roberts and Anders Holmberg (Roberts and Holmberg 2010; Roberts 2012, this volume) develop a proposal that reconciles Baker’s notion of macroparameters with the BCC. They agree with Baker in that macroparametric correlations do exist, but conceive macroparameters as aggregates of microparametric settings affecting formal features of functional categories. Their basic insight is that macroparametric effects obtain when a group of functional heads are specified for the same property. As a central part of their proposal, they further argue that this aggregate behavior is determined not by UG, but by the learning strategy in (17):

(17) *Generalization of the Input*

If acquirers assign a marked value to H, they will assign the same value to all comparable heads.



Roberts and Holmberg remark that the markedness convention in (17) is not a grammatical principle, but an acquisition strategy motivated by the overall conservatism of the learner, who tries to set parameters as efficiently as possible. An idea along these lines is independently explored by Boeckx (2011, this volume), who considers that some parametric clusters, which he takes to be tendencies, as Newmeyer (2004, 2005) does, are due to a bias in the learning process, the Superset Bias in (18), a third-factor economy principle, by which the child acquiring a language seeks to maximize the similarity across parameter settings, harmonizing their values.

(18) *Superset Bias*

Strive for parametric-value consistency among similar parameters.

Some of the papers included in this volume are concerned with the formulation and scope of particular parameters, directly bearing on the issues discussed in this section: that is, the (non)-existence of (macro)parametric clusterings, and the (non)-reduction of (macro)parametric options to microparametric ones. These papers deal with certain aspects of four parameters that have already been presented in this introduction (the Analyticity Parameter, the Null Subject Parameter, the Compounding Parameter, and the NP/DP Parameter), and come to different conclusions, thus showing how central (and controversial) the aforementioned topics are in current generative research on syntactic variation.

Reintges and Cyrino approach parameters from a diachronic perspective in their study on the analyticization of the verbal tense systems in Brazilian Portuguese and Coptic Egyptian, and claim that the results of their work are consistent with Baker's idea that both macroparameters and microparameters play a role in accounting for cross-linguistic differences. In their view, there does exist a cluster of properties associated to the change towards analyticity in the temporal morphological structure of languages, which has to be explained in (syntax-based) macroparametric terms, whereas the synthetic residue in analytic temporal systems can be derived from (lexicon-based) microparameters.

Mensching and Weingart examine the consistency and origin of the clustering of properties associated to the Null Subject Parameter, concentrating on two of these properties (the existence of free inversion, and the non-existence of overt expletives), and ask how they are related to the core property of the NSP: the existence of *pro*. In this fine-grained analysis of (parts of) the NSP, these authors suggest that the non-existence of overt expletives universally co-occurs with the existence of *pro* due to general properties of the lexicon, such as economy-driven lexical blocking effects: their idea is that a language with an empty pronominal will also use this lexical unit as a covert expletive, thus blocking the existence of an overt one. As for free inversion

(which is to be viewed as the fact that the subject can remain in its base position), they hold that this property does not universally follow from the existence of *pro* because this correlation can be blurred by factors which are independent of *pro* (such as the position in the sentence in which *pro<sub>expl</sub>* is merged). Mensching and Weingart thus conclude that “some basic insights of the classical *pro*-drop theory originally proposed by Rizzi (1982) can be maintained.”

In their papers on Snyder’s Compounding Parameter and Bošković’s NP/DP Parameter, both Bauke and Jeong are critical, however, of parametric clusterings.<sup>20</sup> Bauke brings up a number of phenomena that challenge the Compounding Parameter, like the existence of the verb-particle construction in Italian under certain circumstances (e.g., *Gianni è corso via* ‘lit. Gianni is run away’), or the Spanish clitic doubling construction (e.g., *Juan le entregó la carta a Pedro* ‘lit. John CL handed the letter to Peter’), which has been argued to behave like the English double object construction; Bauke points out that this scenario is unexpected, since, under Snyder’s parameter, both constructions should be disallowed in languages (like Italian or Spanish) that lack productive, recursive, and compositional nominal root compounding. To this, Bauke adds that the parameter does not capture the fact that Romance has a productive pattern of phrasal compounding (e.g., *tasse à café* ‘lit. cup of coffee’), and Germanic has a range of compound forms that are non-compositional (e.g., *Kindbett* ‘lit. child bed’). A main point in Bauke’s work is therefore that all these cross-linguistic differences call for a microparametric approach to syntactic variation.

To end up this section, it is clear from the title of Jeong’s paper that, for this author, “macroparameters break down under the weight of evidence.” To support this view, she focuses on the NP/DP parameter, and claims that it has some serious conceptual and empirical problems. As for the conceptual issues, she points out that the NP/DP parameter does not really qualify as a true categorical parameter that facilitates the process of language acquisition, since, as Bošković himself acknowledges, the generalizations associated to this parameter are “one-way correlations, where the lack of articles is a prerequisite, but not necessarily the only requirement for the operations in question,” which “could turn out to be strong tendencies.” Jeong further presents a number of empirical counterexamples to those generalizations, like the fact that, contrary to the claim that only languages with articles allow clitic doubling, several languages without a definite article also do so.

## 2.4. Parameter hierarchies

The idea that parameters do not come in an unordered list but are hierarchically organized was first fully developed by Baker (2001). Baker notes that

some parameters have much greater impact on the form of languages than others. It is the case, also, that the particular choice of a parameter can render another parameter irrelevant (it is known, for example, that polysynthetic languages always allow free word order and generalized *pro-drop*). In other words, depending on the option taken with respect to one parameter, other parameters will not be applicable, since some properties will not be accessible. Baker thus claims that parameters are ranked by their power to affect one another and proposes to approach parameters in terms of their placement in this hierarchy: in particular, “parameter X ranks higher than parameter Y if and only if Y produces a difference in one type of language defined by X, but not in the other” (Baker 2001: 163).

The point of departure of the parameter hierarchy is thus a list of parameters. A scheme is obtained by situating the most highly ranked parameter at the top; in the next line the value of the chosen parameter is specified. Then the parameter affected by those options (values) occupies the immediate lower position. The hierarchy obtained goes as follows: the highest ranked parameter is the Polysynthesis Parameter. Languages with the positive value of this parameter are not affected by the next one, the Head Directionality Parameter (since internal arguments in polysynthetic languages are dislocated), which is therefore situated below. The parameter which determines the Subject Side (of the sentence) comes next, and the hierarchy moves down until the lowest ranked parameters, such as the Null Subject Parameter (we refer the reader to Baker (2001: 183) for the full description and formalization of his parameter hierarchy).

An important prediction of Baker’s proposal is that only some types of languages are attested, since only some conceivable typological combinations are possible. Within this framework, the notion of markedness relates to the number of choices that have to be made to arrive at a particular typological feature in the hierarchy (a language will be more uncommon if it has to make more choices or “decisions”). For example, Baker points out that the relatively lower frequency of VSO languages, like Welsh, with respect to SVO languages would derive from the fact that two more parametric choices are needed for the characterization of verb-initial types.

Baker’s parameter hierarchy has been questioned on empirical grounds. Newmeyer (2004, 2005) observes that the correlation between the number of decisions about parameter setting and the oddity of a language type is not clear. For example, there are many more non-polysynthetic languages than polysynthetic ones, despite the fact that this is the highest ranked parameter, with only one choice required to be set. This author also notes that, given that the Null Subject Parameter is very low in the hierarchy, null subject languages should be rare, which is also contradicted by the fact that the majority of languages of the world are null subject. These and other observations regarding the location of parameters in Baker’s hierarchy lead Newmeyer to conclude

that this model of parametric choice has to be abandoned, and that language-particular differences are to be captured by differences in language-particular rules. In his view, parameters would therefore ultimately amount to (particular) rules (see section 2.3).

In Roberts and Holmberg (2005, 2010) it is claimed, however, that the empirical problems noted by Newmeyer do not question the concept of the parameter hierarchy itself, but can be reduced to just difficulties of formulation. What is needed, they claim, is a more fine-grained idea of parameter, as well as more structured parameter systems. As the authors put it, what should be aimed at is “a theory of parameters which places substantive restrictions on their form and function while maintaining their descriptive power” (Roberts and Holmberg 2010: 32).

Such a theory is aimed at in the work by Longobardi and his collaborators on “parameter schemata” (Longobardi 2005, Gianollo, Guardiano, and Longobardi 2008). As in Baker (2001), these authors explore the widespread interdependence amongst parameters (i.e., the existence of parametric grids), and conclude that these grids can be reduced to a short list of general schemata. Longobardi (2005) proposes that there exist four abstract schemata, which are reproduced in (19), restricting the form of possible parameters:

- (19) a. Is F, F a functional feature, grammaticalized?
- b. Is F, F a grammaticalized feature, checked by X, X a lexical category?
- c. Is F, F a grammaticalized feature, spread on Y, Y a lexical category?
- d. Is F, F a grammaticalized feature checked by X, strong (i.e., does it overtly attract X)?

As Longobardi tells us, the question in (19a) is meant to capture the fact that a particular feature must occur in a certain structure in some languages, but not in others; (19b) asks whether a feature acts as a *probe* searching for a *goal* (in Chomsky’s 2001 terminology); (19c) asks if a feature which is interpreted in a certain structural position has uninterpretable counterparts in other categories which depend on it for valuation; and (19d) asks whether the dependency in (19b) involves overt movement of X, or not.<sup>21</sup>

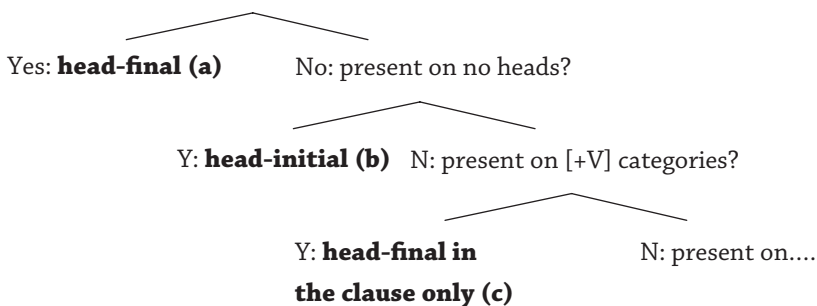
If this approach is on the right track, Longobardi goes on arguing, we no longer need to suppose that UG provides specific parameters, but only a limited number of parameter schemata, “which combine with the appropriate elements of the lexicon (features and categories) under the relevant triggers in the primary data to both yield the necessary parameters and set their value for each language” (Longobardi 2005: 412). The view in Longobardi (2005) on the content of UG is represented in (20) (his (8)):

- (20) Principles & Schemata model: UG = principles and parameter schemata. Parameter schemata at  $S_0$ , closed parameters at  $S_s$

A restrictive theory of parameter hierarchies (or networks) is also developed, on different grounds, in Roberts and Holmberg (2010) and Roberts (2012). Roberts (2012), in particular, building on the insights in Roberts and Holmberg (2010), takes on the idea that micro- and macroparameters are both necessary and should be combined (see section 2.3). His proposal is that the restrictions on variability and on the form of parameters, as well as on the set of possible grammars, should be preserved, as claimed by microparametric analyses. Nevertheless, at the same time, the number of parameters should be reduced on a principled basis, as aimed for by macroparametric work. To do so, a new conception of macro- (and micro)parameters is presented: macroparameters can be viewed as the result of a group of functional heads being specified for the same properties. Macroparameters are thus considered as the result of aggregates of microparameters which act as a single one giving rise to parameter hierarchies.

Parameter hierarchies are set up in relation to different axes of variation (word order, null arguments, word structure . . .). Macroparameters sit at the top of the hierarchy and systems become more marked as we move down, to microparameters: “The options move from subsets of the set of formal features  $F$  to singleton features of heads  $f \in F$ , to increasingly context-sensitive environments, ultimately perhaps to single lexical items” (Roberts 2012: 321). As an example, let us take Roberts’ hierarchy of word order, which is represented as in (21):

(21) Is the head-final feature present on all heads?



(from Roberts 2012: 321)

In the hierarchy in (21), (harmonically) head-final languages such as Japanese and Korean occupy the top left branch (a), while the right branch covers (harmonically) head-initial languages like Celtic and Romance (b). The lower branch features languages like German and Dutch, which are basically head-initial but show head-final TP, vP, and VP. The most embedded (right) branch

is the one that subdivides in increasingly specific microparameters which behave in a more marked and differentiated fashion.

One crucial point in this approach, as already mentioned in section 2.3, is that this behavior of parameters is determined by learning strategies: parameter hierarchies are understood as defining learning paths with the higher options being computationally less complex and thus chosen by learners unless data indicate otherwise. In this way, language acquisition is conceived as moving down the hierarchies, from a more simple to the next-most-complex stage until there is no disconfirming Primary Linguistic Data (PLD). Markedness conditions would then be of the form of the Generalization of the Input Condition in (17) above, repeated in (22)<sup>22</sup>:

(22) *Generalization of the Input*

If acquirers assign a marked value to H, they will assign the same value to all comparable heads.

The degree of specificity of both the grammatical categories and the grammatical operations involved in a parameter also determine its position in the hierarchy: the more specific, the more complex and hence the more microparametric. Micro- versus macroparametric differences thus derive from this notion of markedness, which is claimed to be formulated in terms of third-factor properties in the sense of Chomsky (2005). Roberts therefore concludes that “the form of parameters is thus not specified by UG, but is an emergent property of the interaction of UG, the acquirer and the data. In this way, parametric variation in fact arises from all three of the factors Chomsky (2005) discusses as contributing to language design: UG (underspecification), PLD, and the computational conservatism of the learner, which underlies [the Generalization of the Input]” (Roberts 2012: 334).

Roberts (this volume) works out the conception of macro- and microparameters and their combination by means of parameter hierarchies sketched in Roberts and Holmberg (2010) and Roberts (2012), now focusing on the nature of the features governing parametric variation. He speculates on this point that only N and V features are universal and only N is invariant, and further suggests that formal features associated with phase heads (being left underspecified by UG) are all subject to variation. The issue of parameter hierarchies is also addressed in this volume in the paper by Boeckx, who shares the idea that learning tendencies, such as the Generalization of the Input in (21), play a central role in deriving the harmonic patterns observed in (at least some) macroparameters (see section 2.3). Boeckx’s vision on parameter hierarchies departs from the one in Roberts and Holmberg (2010) and Roberts (2012), however, in that he takes the aggregation of microparameters to result in the emergence of parameter hierarchies in a bottom-up fashion, whereas Roberts and Holmberg seem to think that learning paths go the other way

round, with the learner moving down the hierarchies in the process of language acquisition (see above).

## 2.5. Conclusions

In this section, we have reviewed a number of debated topics within the Parametric Theory of syntactic variation: (a) Are parameters (and their values) provided by UG, or should they be better viewed as emergent properties of grammars arising from points of underspecification instead?; (b) which is the locus of parametrization?, can parameters be entirely confined, in particular, to the externalization process holding at the syntax-PF interface?; (c) do the clustering effects that have been associated to certain parameters actually hold across languages?; (d) do (macro)parameters exist alongside microparameters, or can (macro)parametric differences between languages be ultimately reduced to microparametric ones?; and (e) are there parameter hierarchies?, and if this is the case, which is their form and source?

As can be deduced from our discussion, all these questions are still waiting for a definite answer. Against this background, the aim of the present volume is to offer a representative sample of current theorizing and work on parameters that thoroughly address some (or all) of the aforementioned topics.

The rest of this introduction summarizes the general ideas and specific proposals of the papers included in the volume, which is organized into two parts according to their main orientation and goals: Part I, The nature of variation and parameters, and Part II, Parameters in the analysis of language variation: Case studies.

## 3. THE NATURE OF VARIATION AND PARAMETERS

Part I in this volume brings together those papers whose main objective is to discuss global issues related to parameters (or variation more generally). It begins with the work by **Adger**, which deals with the nature of variation in a broad sense, asking, in particular, whether a categorical theory of syntax is compatible with apparently probabilistic distributions of variation in the syntactic data *within a single speaker's* grammar. The central question in this paper is thus how generative syntactic theory can model intrapersonal variability. Adger argues at this respect that current minimalist theorizing is compatible with (some kinds of) grammatical intrapersonal variability if this is taken to derive from the way that interpretable and uninterpretable features combine.

To support this view, examples are analyzed of certain varieties of English which involve agreement systems (inflection), choice of functional elements



(relative clauses, *do* support), and word order, illustrating cases where operations of agreement, movement, and deletion give rise to different options available to the same speaker (in the same utterance). These cases pose a problem for what is called *Variation in Exponence* view, which lies at the heart of a variationist sociolinguistic approach: that is, how to determine the equivalence between two syntactic forms and a single semantic interpretation (the question of whether there are synonyms in human languages). Within the *Combinatorial Variability* model proposed by Adger, it is expected that variation within a single grammar is possible just in the context of feature checking operations (such as agreement); that is, only when uninterpretable features are involved. The reason for this claim is that the presence of uninterpretable features does not impact on the semantic interpretation. Therefore, in Adger's words, "a choice of lexical items A and B will be available to agree with C when either of A or B bear uninterpretable features that can match with C, but are distinct from each other." The grammar then produces a "Pool of Variants," each of which is a distinct feature complex with the same semantic interpretation and different phonological forms. In this model, the systems of use are conceived as a choice function on the pool of variants, given a context of utterance. This function is sensitive to phonology, sociolinguistic connotations, and also to frequency with respect to the speaker, as well as to preferences of particular speakers for particular words.

Boeckx's paper starts with an overview of the problems posed by the classical notion of parameter put forth in Chomsky (1981), and the more recent proposals, inspired in Borer (1984), restricting the notion of variation to lexical parameters. It is then argued against what the author dubs "constructive parameters," such as Snyder's Compounding Parameter and "syntactic parameters" such as Baker's Polysynthesis or Direction of Agreement Parameters: all of them fail to hold once we look at a wide range of languages. Boeckx considers that the ultimate aim of all those proposals is mainly to capture results from typological research (what he calls Greenberg's problem) rather than to determine the possibilities open in order to successfully acquire a specific language, which is closer to Plato's problem, the real focus of biolinguistic inquiry.

Based on Boeckx (2011), in order to account for described "cascade" or clustering effects associated to (sub)parameters and parameter hierarchies, a "Strong Uniformity Thesis" (SUT) (see (11) in section 2.2) is further proposed, which eliminates any parametrization of syntax and restricts linguistic diversity to the processes of externalization. Under SUT, parameters are *points of variation* that should be confined to the margins of narrow syntax, that is, to the morphophonological component. Two facts constitute the point of departure of the reasoning: (i) structures generated by (universal) syntax have to be externalized and (ii) this is done with the tools provided by morphophonology. It is the combination of (i) and (ii) that results in a pool of variants



available to language learners. Parameters are thus reformulated in terms of realizational properties, and parametric values are defined as “constructional idioms.” Within this framework, the learning mechanism consists of “constructing one’s own native (grammatical) vocabulary.”

Boeckx further builds on Fasanella and Fortuny’s (this volume) basic dimensions of variation—the bound/unbound distinction (phonological dependence) and the analytic/synthetic distinction (morphological complexity)—but reduces the four logical possibilities of the typology obtained to three, since the bound/unbound distinction only applies to analytical elements. This asymmetrical typology is then compared with known parametric hierarchies to conclude that the language faculty favors the alignment of parametric values for related parameters. Rather than by principles this is done by means of a bias, the “Superset Bias” (SB) (see (18) in section 2.3). SB shares with Roberts and Holmberg’s Generalization of the Input (see (22)) the possibility to predict more macro-patterns of variation than a purely microparametric approach, but a bias does this while allowing for exceptions. Boeckx finally argues that SB can be conceived as a third-factor principle (Chomsky 2005), a regularization bias that, in his words, is “a principle of (self)organization that does not require pre-specification to give rise to macroscopic patterns.” From classical parameters and the switchboard metaphor Boeckx proposes to move towards a “pointillist painting” conception of variation, “one that takes variation to be patterned, but not pre-patterned.”

The paper by **Fasanella and Fortuny** (F&F) explores the relation between parametric theory and learnability conditions. F&F critically review a representative selection of macro- and microparametric proposals in order to show that they do not satisfy certain intuitive learnability conditions, thus failing to account for Plato’s problem. Regarding macroparameters (e.g., Baker’s Polysynthesis Parameter), F&F argue that, though they are elegant in systematizing linguistic variation, they fail as formal models for language acquisition since they exhibit what these authors call the Locality Problem: they are defined on the basis of highly general properties which are scattered and spread across different components and constructions of the language so that, to fix their value, learners have to analyze data globally. This is highly implausible as a learnability condition. Rather, parameters should be atomic, they cannot be clusters of properties (Atomicity Condition). Microparametric schemata, such as Roberts and Holmberg’s (2010) approach to the Null Subject Parameter, are more plausible as expressions of efficient learnability conditions, but according to F&F, they often rely on highly abstract syntactic notions (probe-goal, Agree, and so on), facing the so-called Linking Problem: parameters are defined over abstract linguistic entities, and the child must face the task of linking these abstract mental representations to physical properties of the speech signal. Though microparameters satisfy the Atomicity Condition, they fail to satisfy the Accessibility Condition. A third problem, affecting both macro- and