

On the elements of syntactic variation

LUIGI RIZZI

2.1 Introduction

How to properly characterize syntactic invariance and variation is the core question of theoretical and comparative syntax. The parametric approach introduced a novel technical language to address this issue, which inspired much descriptive and theoretical work in syntax, as well as a new way to study language acquisition. The approach also raised questions and controversy, both within generative grammar and in the larger setting of the study of language as a cognitive capacity. This chapter offers a personal view on the debate raised by the theory of parameters, based on my own research experience and current work, and with no ambition of a systematic coverage of the relevant issues.* In the first part of the chapter, I will briefly describe the origins of the parametric approach, the context in which it was introduced and the impact that the idea had on syntactic and acquisition studies. In the second part, I will discuss the way in which parameters can be integrated in a minimalist grammar, and nourished by the results of cartographic studies. I will address some critical appraisals which question the restrictiveness and deductive richness of the approach, and will try to respond to such critiques. In the third part, I will broaden the picture to the larger debate between 'language faculty' and 'cultural' approaches to language diversity and language acquisition: I will address the question of how the study of acquisition could bear on this conceptual divide, and review some experimental results which are naturally expected within an approach based on a biologically determined language faculty consisting of principles and parameters.

* A preliminary version of this chapter was presented at the Workshop *Linguistic Variation in the Minimalist Framework*, January 14–15, 2010 UAB, Casa de Convalescència, Hospital de Sant Pau, Barcelona, and appeared in *STiL—Studies in Linguistics*, CISCL Working Papers, 4, 2011, University of Siena.

2.2 Origins of the parametric approach

The problem of syntactic variation was explicitly addressed in the early 1970s in the context of the discussion on 'explanatory adequacy'. This is the level of empirical adequacy that is attained by an analysis which comes with a reasonable account of how the property under investigation is acquired by the language learner (Chomsky 1964). The crucial model at the time was the Extended Standard Theory (EST), based on the following notions:

- (1) EST Model (i.e. Chomsky 1973, 1976, 1977):
 - Particular grammar: a system of language-specific, construction-specific rules, expressing the adult speaker's linguistic competence.
 - Universal grammar: a grammatical metatheory specifying a broad format for rules and some general principles on rule application.
 - Acquisition: rule induction

Much as in earlier formulations of generative grammar, the theory was really focused on the notion of particular grammars as systems of formal rules specific to a particular language, and construction-specific: language-specific phrase structure rules were assumed for major phrases, as well as construction-related transformational rules like passive, question formation, relative clause formation, etc. Such a rule system would constitute the grammar of English, and similar systems were postulated for Italian, Chinese, etc. Universal Grammar (UG) was thought of as a kind of general metatheory of grammatical properties specifying the format for rules and expressing certain general constraints on rule application such as the A over A principle and, after Ross' (1967) thesis, the island constraints. This system presupposed a particular conception of language acquisition. Acquisition would be rule induction: the child would act like a 'small linguist', unconsciously formulating and testing hypotheses in order to figure out what the rules of her particular grammar were on the basis of the format provided by Universal Grammar and of the empirical evidence presented to her.

There were some obvious problems with this way of looking at things. One critical problem was that a system based on language specific rules was not suitable for comparing languages: one would build a rule system for language A, and then start from scratch and build another rule system for language B, etc. Such rule systems would obviously bear some kind of family resemblance, but the architecture of the model would not favour a clear identification of the primitive properties that remained uniform and of those that varied, a rather frustrating state of affairs. Comparative syntax was not really feasible on that basis because the fundamental invariant and variable elements could not be isolated in a sufficiently transparent manner.¹

¹ Part of the problem was related to the construction-dependent definition of grammatical rules (relative clause formation, passive, etc.). The problem is that constructions are molecular entities, organized structures made of finer atomic ingredients: while variation is overwhelming at the construction level, it is

Another serious problem was that this system could not successfully address the problem of acquisition because ideas were not precise enough about how rule induction could work so that the analyses did not attain the level of 'explanatory adequacy' in the technical sense defined by Chomsky (1964). It was clear at that time that one could hope to successfully address this problem only by radically restricting the options offered by UG, i.e. by making the rule systems among which the child was assumed to choose more and more constrained, in order to make the selection of the 'right' rule system feasible within the limits of time and access to the data that characterize language acquisition. Constraining the expressive power of UG was a major success of EST around the mid 1970s, thus making the goal of reaching explanatory adequacy more feasible. Still, the theory needed an appropriate technical language to address linguistic variation, and a suitable mechanism for the acquisition of language-specific properties.

Things changed around the second half of the 1970s. Recently, I came across the passage in (2) in Chomsky's 'Conditions on Rules of Grammar'; which contains, as far as I can tell, the first mention of the term 'parameter':

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

N. Chomsky (1976). 'Conditions on Rules of Grammar', republished in Chomsky (1977: 175).

The passage considers the possibility that certain principles or rules could be parameterized and that could account for certain aspects of variation. The idea was purely abstract at that time but the first concrete instantiation came up a few years later through the study of extraction from *wh*-islands. It turned out that in some languages it is possible to extract an element from an indirect question as in (3) in Italian, while in other languages this option is marginal or absent (Rizzi 1978, reprinted in Rizzi 1982, chapter 2):

- (3) Ecco un incarico [_{CP} che [_{IP} non so proprio [_{CP} chi [_{IP} potremmo affidare]]]]

'Here is a task that I really don't know to whom we could entrust'

only at a finer level of granularity that invariant properties fully emerge. Constructions can be looked at, in current models, as structural molecules consisting of elementary operations such as merge and move, and the featural specifications in the functional lexicon triggering them, the latter specifications being the natural locus of variation in an otherwise invariant system (see section 4 for discussion). An important antecedent to a 'compositional' approach to constructions providing a basis for comparative syntax is to be found in Bach (1965).

If we take the word by word equivalent in other languages, e.g. German (modulo word order and other properties), we obtain a deviant structure (on complexity and variation in English, the language originally compared to Italian, see the discussion in Grimshaw 1986):

- (4) *Das ist eine Aufgabe, [_{CP} die [_{IP} ich wirklich nicht weiss [_{CP} wem [_{IP} wir ____ anvertrauen koennten]]]].
'Here is a task that I really don't know to whom we could entrust'

It seemed too radical to assume that the relevant locality principle deemed to be responsible for (4), Subjacency, would not be operative at all in languages like Italian: somewhat more complex examples showed that Italian is sensitive to locality effects reasonably amenable to Subjacency. For instance, while extraction from an indirect question is normally possible, extraction from an indirect question which in turn is embedded under another indirect question (a double wh-island) is clearly degraded:

- (5) *Ecco un incarico [_{CP} che [_{IP} non so proprio [_{CP} a chi [_{IP} si domandino [_{CP} se [_{IP} potremmo affidare ____]]]]]].
'Here is a task that I really don't know to whom they wonder if we could entrust'

So, the idea was proposed that Subjacency is operative in both language types, banning movement across two bounding nodes; but the set of bounding nodes could be parameterized in a way that would yield the difference: i.e. by taking CP as the clausal bounding node for Italian, and IP for more restrictive languages (in fact, S' and S in the original notation). So that two BN (two occurrences of IP) would be crossed in (4), but only one BN (CP) would be crossed in (3); two CPs would be crossed in the double wh-island (5), thus accounting for the deviance of the structure in Italian.²

In retrospect, this turned out to be a rather peripheral kind of variation. Judgements are complex, graded, affected by many factors, difficult to compare across languages,³ and in fact this kind of variation is not easily amenable to the general format of parameters to be discussed later on. Nevertheless, it was soon realized that this kind of mechanism could be successfully employed to express major, crystal-clear

² Certain varieties of German are very restrictive on wh-extraction, banning extraction even from embedded declaratives and permitting the expression of questions like 'Who do you think we should meet?' through other techniques, such as 'partial movement' (Felser 2004). The strong restrictions on extraction in such varieties have sometimes been treated in terms of the parametrization of bounding nodes, e.g. in Freidin and Quicoli (1989). Other varieties, spoken e.g. in Southern Germany and Austria, straightforwardly permit extraction from declaratives. More liberal varieties may also marginally permit certain kinds of extractions from wh-islands, showing the asymmetries referred to in the following footnote (Grewendorf 2012).

³ Relevant factors involve structural and interpretive properties of the extracted element, such as Discourse-linking. In fact, a very stable cross-linguistic pattern emerged from this line of research: the existence of asymmetries between elements at least marginally extractable from indirect questions, and elements which strongly resist extraction. This led to the discovery of weak islands and to much theoretical work to capture the asymmetries. See Szabolcsi (2005) for a general overview.

cross-linguistic differences, and one could entertain the ambitious programme of dealing with the core of cross-linguistic variation in terms of a system of parametric choices; the postulation of a set of language specific rules could thus be disposed with entirely.

Parametric theory introduced a powerful technical language for doing comparative syntax, one which permitted a transparent identification of invariant and variable properties. So it is not surprising that comparative syntax flourished as soon as the new 'principles and parameters' approach was introduced (Chomsky 1981). I believe it would not be difficult for a historian of our field to gather massive evidence in scholarly journals, proceedings of conferences, and book series documenting a rather dramatic shift: in very few years, comparative generative grammar grew from very sparse attempts to a substantial body of scholarly work on dozens or hundreds of languages analysed in a comparative perspective in terms of the parametric model. Moreover, the theory of principles and parameters provided a promising model of the acquisition of syntax *qua* parameter setting, a much more appealing conception than one based on an obscure notion of rule induction (see Hyams (1986) for a proposal which inspired much acquisition research, and Rizzi (2000), Introduction, for a general assessment). Reaching the level of explanatory adequacy thus became a feasible enterprise, even if by no means an obvious one, due to the ambiguity that the primary data may leave open about distinct arrays of parametric values (Gibson and Wexler 1994).

2.3 Some problems with the initial parametric model

A theory of parameters should address the questions of the format (what is a possible parameter?) and of the locus (where are parameters expressed?) of such entities. Initially, not much theoretical reflection was devoted to the format of parameters, but one clear idea on the locus was explored. As the first concrete proposal on parametrization looked like an option specified on a principle, perhaps that was the locus of parameters in general. So the hypothesis was entertained that parameters would be expressed in the set of UG principles:

- (6) Parameters expressed in principles: each UG principle specifies one (or a small number of) parameter(s), a choice point to be fixed on a certain value for the principle to become operative.

This had certain consequences. For instance, it gave a rough estimate of the size of the set of parameters: as there were few principles in the modular structure assumed by the government-binding approach (X-bar theory, Case theory, Theta theory, the theory of binding, the ECP, Subjacency, and a few others), one would expect relatively few parameters. The approach also gave rise to the so-called switchboard metaphor (an image originally due to James Higginbotham): the child is confronted with a little switchboard with principles specifying parameters, and the acquisition process

consists essentially in setting the switches on the basis of experience; once this is done, the syntax of the language is acquired.

Not much attention was paid initially to the format of parameters, that is to say, to what a possible parameter is. So that virtually every property was proposed as a potential target of parametrization. In (7), I give a little list of parameters that were identified around the late 1970s or in the early or mid-1980s:

- (7) - the bounding nodes are ... (Rizzi 1978, Sportiche 1981, ...)
 - null subjects are licit, (Taraldsen 1978, Rizzi 1982, ...)
 - *believe* type verbs select an IP (English vs. Romance: Chomsky 1981)
 - P assigns structural/inherent Case (P-stranding, ... Kayne 1983a)
 - the head precedes/follows the complement
 - V moves to I (Emonds 1978, Pollock 1989)
 - V moves to C (V-2 Germanic: den Besten 1977/1983)
 - N incorporates into V (Baker 1988a)
 - Nominative is assigned under agreement (SVO) or under government (VSO) (Koopman and Sportiche 1991)
 - there are long-distance anaphors (English vs. Icelandic, etc.: Manzini and Wexler 1987)
 - wh-movement is overt or covert ... (English vs. Chinese, etc.: Huang 1982)
 - the language is non-configurational (Hale 1983)

The list thus includes properties of locality, the licensing of empty elements, selectional properties of special verb classes like *believe*-type epistemic verbs, movement properties of various sorts, linear order, and also very general statements about global properties of a language like Ken Hale's proposal that there is a configurationality parameter. Some languages are configurational, based on hierarchically organized structures, others are non-configurational, involving flat (or flatter) structures, and that affects in a very deep way the whole structure of the language; first and foremost this property is responsible for the freedom in word order.

It became clear pretty soon, already in the early 80s, that this approach had to face serious problems. One was the unnatural character of the list in (7), and then, there were other problems, some of which are indicated in (8):

- (8) Some problems with the model of 'parameters as specifications on principles':
 a. the arbitrary-looking character of the list of the first parameters;
 b. some principles don't seem to require/allow any parametrization;
 c. some parameters do not express global properties of the particular language, but appear to be directly keyed to the presence of particular items in the lexicon of the language;
 d. certain global parameters like non-configurationality can be advantageously reanalysed as conglomerates of more elementary parameters.

Point a. is self-evident on inspection of list (7). As for point b., one can consider, for instance, the hierarchical properties of X-bar theory—always the same across languages, presumably: structures are built by heads projecting and taking complements and specifiers, following a binary branching organization (Kayne 1983a); the core of the Theta Criterion also seems to be exceptionless; e.g. no known language seems to admit structures like '*My friends seem that John likes Mary', '*Bill happens that John left early', leaving a DP in argument position not integrated into a thematic nucleus; certain aspects of the Binding theory, such as the core case of Principle C, appear to be exceptionless: whenever a pronoun c-commands a DP, a referential dependency is uniformly banned, as in 'He thinks that John will win' and its equivalent across languages, modulo linear order and other language-specific peculiarities.

Point c., perhaps more important, is that some parameters appeared to be directly related to the presence of a particular lexical item in the language. Take for instance long-distance anaphora. It is very clear that we cannot say that long-distance anaphora is a global parameter concerning the binding theory in one language because it depends on the presence in the lexicon of that language of a particular item that functions as a long-distance anaphor, like *sig* in Icelandic for instance, which has such type of binding properties. So, clearly, long-distance anaphora is not a global property of binding in a particular language, it is a property of a particular lexical item, and cross-linguistic variation depends on the presence or absence of that particular item in the lexicon of the language.

As for point d., it turned out that certain global parameters like non-configurationality could be advantageously reanalysed as conglomerates of smaller parameters. On the one hand, detailed work on extreme cases of 'non-configurationality' led to the observation of many manifestations of a strictly configurational hierarchical structure: c-command effects, positive response to constituency tests, strict orders in special environments, etc. (see the discussion in section 7). On the other hand, it became progressively clear that languages manifest distinct degrees of 'non-configurational' properties such as freedom of word order, a gradation that is not expected under the view of a single 'all or nothing' configurationality parameter. For instance, it is clear that null subject languages are, in an intuitive sense, less configurational than non-null subject languages because they manifest a higher level of freedom in the position of the overt subject (with subject inversion, subject dislocation, and the like). Scrambling languages are also more non-configurational than non-scrambling languages as they admit a number of alternative orderings (but if the analysis is refined, one particular order generally emerges as the fundamental one, as shown in much detailed work on scrambling in German and Japanese over the years: see Grewendorf and Sternefeld (1991) and references quoted there). There are more surface ordering options in languages in which it is possible to split the DP than in languages which necessarily preserve the DP integrity, a property plausibly related to the nature of the D system (Bošković 2009), etc. So, one observes a gradation of non-configurationality, not a

continuum in the technical sense of course, but a number of discrete degrees that are better accounted for in terms of smaller parameters. The extreme cases of this spectrum (say, English and Warlpiri) look like radically different systems, but many intermediate cases are attested, which again suggests the necessity of breaking up a very radical macroparameter into a set of parameters independent from one another and more restricted in scope.

2.4 From 'parameters expressed in principles' to 'parameters in the functional lexicon'

A significant shift with respect to the initial assumption on the locus of parameters, directly suggested by problem (8)c, can be expressed as follows:

- (9) Parameters are specified in the functional lexicon of particular grammars.

This hypothesis was in fact formulated very early on, and is clearly expressed in the following quote taken from Hagit Borer's work:

- (10) 'The inventory of inflectional rules and of grammatical formatives in any given language is idiosyncratic and learned on the basis of input data. If all interlanguage variation is attributable to that system, the burden of learning is placed exactly on that component of grammar for which there is strong evidence of learning: the vocabulary and its idiosyncratic properties.'

Borer (1984: 29)

This view was widely accepted from its initial formulation, but, in much work in comparative syntax, it coexisted for a long time with the view of 'parameters expressed in principles', and only more recently (9) started being quite generally assumed as the exclusive characterization of the locus of parameters.

I will basically adhere to this conception in the rest of this chapter, but a preliminary caveat is in order. The idea of restricting the expression of parameters to the functional lexicon is clearly motivated by the desire of constraining the parametric space as much as possible. But it is not entirely obvious that all the properties that we want to consider parametric are exclusively associated to functional elements, at least if we assume a simple-minded, traditional view of the functional-contentive divide. Take, for instance, the familiar, sharp difference in syntactic behaviour between the infinitival complements of epistemic verbs like *believe* in English and Romance as in (11) and (12):

- (11) English:
 a. I believe [John to know the answer]
 b. *I believe [PRO to know the answer]
 c. John was believed [____ to know the answer]

- (12) Italian (Romance):
 a. *Credo [(di) [Gianni sapere la risposta]]
 b. Credo [di [PRO sapere la risposta]]
 c. *Gianni era creduto [(di) [____ sapere la risposta]]

In English, *believe* type verbs take infinitival complements which manifest exceptional Case marking, no control, and the possibility of licensing subject to subject raising with the passive voice, as in (11). In Romance, one observes the mirror image of these properties: no exceptional Case marking, control, and impossibility of raising, as in (12). Now, these properties seem somehow to be keyed to the selectional properties of *believe* vs. the equivalent in Romance languages: in classical GB terms, we have a lexical parameter differentiating the categorial selectional properties of epistemic verbs in Romance (which uniformly select a CP as a clausal complement, with non-finite C overtly realized as Italian *di*, or null, as in French) and English (which apparently directly selects an infinitival IP, with the whole CP layer truncated); these seem to be parametric properties associated to (classes of) lexical verbs, at least if the divide between lexical and functional verbs is maintained in a traditional form. The fact that these systematic properties affect whole classes of verbs, rather than single items, suggests a possible analysis consistent with the assumption that the parametrization is limited to the functional lexicon, as Frédérique Berthelot (p.c.) points out. Thinking of the decomposition of verbs into v and root components (Harley 2011; Marantz 2013; Ramchand 2008, and references quoted there), the class could be characterized by the presence of a specially 'flavoured' v, say v_{epist} , a functional element whose featural specification could be responsible for the c-selectional properties of the complex $v_{\text{epist}} + \text{root}$.⁴

Other problematic cases come to mind, e.g. the cross-linguistically (and language internally) variable c-selection of DP vs. PP complements (*écouter la radio* vs. *listen to the radio*; *entrer dans la chambre* vs. *enter the room*), etc., and all the item-particular cases in which categorial selection seems to depart from the Canonical Structural Realization of semantic selection (Grimshaw 1979; Pesetsky 1982) in language-specific, and item-specific ways. A possible solution here may be provided if 'selected' prepositions are reanalysed as being part of the functional structure associated to the verb (again, to specially flavoured v items), much as in Kayne (2000a).

In the remainder I will continue to make the assumption that the locus for the expression of parameters is the functional lexicon, but it is important to bear in mind the problems just mentioned, which may require a rethinking of the traditional divide between functional and substantive lexicon (on this issue see Kayne 2005b, Cinque and Rizzi 2010 for discussion).

⁴ Things are further complicated by the fact that the class does not behave in a fully homogeneous manner (Postal 1974: *allege* differs somewhat from *believe*, etc.), which may require further refinements of the decomposition $v + \text{root}$.

2.5 The theory of parameters in current models

What does a parameter look like in current syntactic theorizing? The drastic simplification of the structure of UG assumed by Minimalism makes hypothesis (9) on the locus of parameters particularly congenial to minimalist thinking. Capitalizing on the emphasis put by Minimalism on the elementary ingredients of linguistic computations (Chomsky 1995, 2000), and building on some suggestions in Rizzi (2009a), I would like to propose the following informal characterization:

- (13) A parameter is an instruction to perform a certain syntactic action expressed as a feature on an item of the functional lexicon, and made operative when the item enters syntax as a head.

So, when an item is selected from the functional lexicon and enters syntax, it will contain certain formal featural specifications which will instruct syntax by triggering certain syntactic actions, first and foremost Merge itself.

More precisely, I would like to propose the following extremely simple format for parameters, which is accompanied by the specification of the locus:

- (14) **Format:** H has F {yes, no}
Locus: parameters are specified in the functional lexicon

Where H is an item of the functional lexicon, and F is a morphosyntactic feature. In order to make the system properly restrictive, we must now specify the range of F more precisely. Features express properties of various kinds: of sounds, of meanings, of morphological characteristics, etc. Most of such properties do not affect syntax in any way, so that they are not relevant here. I will make the rather standard assumption that in the set of possible linguistic features there is a well-defined subset of morphosyntactic features which are capable of triggering the basic syntactic actions.⁵ If we assume

⁵ Boeckx (this volume) underscores the importance of a proper understanding of how morphosyntactic features are assigned to heads. This is indeed an important issue, not only for the proposed approach to linguistic variation, but for the study of morphosyntax in general. A proper account should provide a characterization 1) of what a possible morphosyntactic feature is; 2) of what a possible functional head is; 3) of how a morphosyntactic feature can be associated to a functional head. As for the first question, see Cinque's (2013) discussion on the fact that only a small subset of cognitively relevant features are 'grammaticalized' and used by natural language syntax: here a fruitful integration is possible of much work on grammaticalization pursued within the typological tradition (e.g. Heine and Kuteva 2002), Cinque points out. As for the second question, one guideline of cartographic studies is the assumption that each (interpretable) morphosyntactic feature defines a functional head, both in the inflectional space and in the left periphery (Cinque and Rizzi 2010); complex conglomerates of features could thus only be derived syntactically, through head movement; if this guideline is on the right track, the third question on feature assignment to heads becomes trivial. Whether this strong position can be fully upheld (see also much work in Nanosyntax, Starke 2009), or it must be weakened in some form, the problem of the assignment of morphosyntactic features to functional heads is clearly and narrowly defined: there is no risk of combinatorial explosion or uncontrolled complexity, and the restrictiveness of the proposed approach to parametrization is ensured by the highly restrictive character of minimalist syntax (see section 6).

a highly restrictive theory of possible syntactic actions, as in minimalist syntax, parametric features will be restricted to the features triggering the elementary operations of Merge, Move, and Spell-out. So, in a nutshell, we have the following basic typology of parameters:

- (15) A typology of parameters:
1. Merge parameters;
 2. Movement parameters;
 3. Spell-out parameters.

Merge parameters may primarily express cases in which the head's categorial selection (c-selection) does not immediately reflect principles of canonical structural realization (Grimshaw 1979; Pesetsky 1982): e.g. the cases of 'truncated' clausal selection of English epistemic verbs referred to in the previous section. Other cases may involve the permissible cross-linguistic variation in functional hierarchies: a Negative Phrase which can be very high (in the CP zone), or in the high, intermediate, or low IP zone (Zanuttini 1997; Cinque 1999; Moscati 2007); types of Agreement (or agreement-bearing) heads, which can vary significantly from language to language (Haegeman 1992; Cinque 1999; Belletti 2001); single or recursive Topic in the left periphery, presence or absence of a Top position in the lower left periphery (Rizzi 1997; Bianchi and Frascarelli 2010; Haegeman 2012), presence or absence of Focus projections in the CP and/or in the vP periphery specialized to new information or contrast (Rizzi 1997, 2004a; Belletti 2004a, 2009; Cruschina 2006), etc.

Move parameters express the ability that a head has of attracting another head (incorporation), or a phrase to its specifier position (the latter case being uncontroversial and subsuming the former in some approaches). Parametric properties involving the movement of the verb to an inflectional head (Emonds 1978; Pollock 1989; Cinque 1999; Holmberg and Platzack 1995; Roberts and Holmberg 2005; Holmberg and Roberts 2013), and of the inflected verb to the C-system are expressed here, as well as all the parametric variation involved in movement to a Spec position (wh-movement languages vs. wh-*in-situ* languages, etc.); I omit here the further refinements required by the assumption that movement is search + (internal) merge (as in Chomsky 2000), which could lead to distinct possible parametrizations on the search operation, and on internal merge.

The head-complement ordering parameters may be seen as merge parameters in more traditional approaches (merge the complement to the left/right of the head), or as movement parameters in antisymmetric approaches (Kayne 1994: move the complement to a higher Spec position); or else as a spell-out parameter if ordering is a property confined to externalization (Berwick and Chomsky 2008: spell out the complement before/after the head). Whatever the exact nature of this property, the ordering parameters can be seen as particular instances of format (14), with the crucial feature specified on the functional categories assigning the categorial status to lexical

roots (i.e. v, n, a, p, etc.), and with the greenbergian tendency to uniformity across categories (Greenberg 1963) expressed grammatically (Biberauer et al. 2008) or extra-grammatically (Newmeyer 2005).

A straightforward spell-out parameter has to do with whether or not a given functional head is pronounced: so, a Top head is pronounced in Gungbe (Aboh 2004), but not in English; and with the licensing of a null specifier: Top has this property in Topic Drop languages (perhaps derivatively from the capacity that a given node may have to constitute the 'root' of the structure: Rizzi 2006a); and some inflectional heads have the capacity to license a null pronominal subject and/or a null pronominal object in some languages (Rizzi 1982, 1986), etc. Various kinds of ellipsis may also be head-driven (Merchant 2001), and functional nouns may have null variants in particular structural environments (Kayne 2000a).

In a sense, this view leads us back to a version of the switchboard model, except that the switches are now expressed in the lexical items: each item of the functional lexicon has a small number of switches, corresponding to the typology in (15); acquiring the syntactic properties of the lexical item amounts to setting its switches on the basis of the linguistic data the learner is confronted with. So, a given head may c-select a particular category (departing from the canonical structural realization of its s-selectional properties), attract another head or a specifier, be spelled out or not, and govern the spell-out properties of its dependents.

2.6 On the numerosity of parameters

The view that the functional lexicon is the locus of parameters affects the expectations on the number of parameters. We will have many more parameters than it was initially assumed if the size of the set of parameters is related to the size of the functional lexicon: clearly, there are many more opportunities for parametric specifications than in the assumption that the locus is the small set of UG principles. Moreover, if cartographic studies are on the right track (Cinque 2002; Belletti 2004b; Rizzi 2004b; Cinque and Rizzi 2010), the functional lexicon is much richer than in more traditional approaches, so the number of potential parametric specifications is even greater.

Such assumptions on the numerosity of parameters, a natural, and in fact virtually inescapable consequence of the conceptual shift reported in section 3 and of the view on the format in (14), are sometimes taken as a kind of *reductio ad absurdum* of the core idea of parametric syntax, the idea that syntactic diversity is amenable to a finite set of binary options open to all languages. If the options offered by the system are so numerous, why continue to call them parameters? Doesn't the term improperly suggest a highly restrictive space of variation? So, the current conception is sometimes seen as an undeclared retreat to the EST conception of grammar as a system of language-specific rules (see, e.g., Newmeyer 2004, 2005): if there are so many possible

parameters, how is this conception different from one treating variation through language specific rules?

This argument does not take into account the fundamental distinction between the locus and the format of parameters. Under the current conception, the loci of parameters are quite numerous and diverse, but the format is extremely restrictive, as determined by the restrictiveness of minimalist syntax and its mechanisms. The syntactic actions that a featural specification on a head can trigger are very few, restricted to the very basic and general operations of merge, move, and spell-out: the parametric space is thus radically more restricted than the space of possible language-specific rules of arbitrary complexity in EST models.

We are thus very far from the explosion of possibilities determined by an unconstrained notion of language specific rule system. Therefore, the problem of restrictiveness which hampered the explanatory capacity of pre-parametric models does not arise here. Assimilating current views on variation to EST models thus overlooks the genuine and substantial progress in the identification of the basic ingredients of linguistic computations over more than thirty years of syntactic research.⁶

The assumption we are now making on the size of the set of parameters has other consequences. If the system had only few parameters, sparse and relatively isolated in their consequences, one could expect that a single parameter could control a complex set of properties varying across languages: this was a natural expectation in a system based on few parameters expressed on principles (consider, e.g., Chomsky 1981, Rizzi 1982 on the Null Subject Parameter, which I go back to shortly). On the other hand, many parameters imply many intricate interactions: if parameters are so numerous, and ubiquitously expressed in the functional lexicon, it is very unlikely that a single parameter may fully control complex sets of properties. Again, this is sometimes taken as a major drawback of current parametric models, as evidence that parameters in the current view only have local consequences and a parameter-based system with many parameters has no deductive depth, hence, ultimately it is not a particularly revealing model of language variation. Quite the contrary is true, in my opinion. Parameters undoubtedly express local properties, encoding how a particular item interacts with

⁶ Of course, the choice of a particular terminology is largely an arbitrary decision. So, one may decide not to use the term 'parameters' for the devices referred to by (14), (15) and call them 'language-particular rules', or the like, without changing in any way the structure of the approach. Nevertheless, using the term 'rule' in connection to such theoretical entities would be misleading. First, because the term 'rule' evokes the complex phrase structure and transformational rules of pre-parametric models, which are quite different from the highly restrictive devices expressed in (14), (15). Second, because the shift of the locus for parameters from UG principles to the functional lexicon took place already around the mid-1980s, is a development largely (if not unanimously) accepted by the scientific community of comparative syntacticians, and major work in comparative syntax over the last quarter of a century has consistently used the parametric terminology to refer to such concepts and tools both in the pre-minimalist and minimalist era (see, e.g., Kayne 2000a, and many contributions in Cinque and Kayne 2005). In the absence of a new conceptual or formal shift, I think it would be misleading to introduce a new terminology, or go back to a highly connotated old terminology.

its immediate structural environment; but the deductive structure of the system is tight and rich, so that even a small difference at a particular point may well have systemic consequences, through the interaction with principles, computational mechanisms, and other parametric choices.

Consider an analogy with the structure of DNA and its role in determining the development, shape, and functioning of the organism. The action of a single gene may be local—perhaps limited to turning on or off another gene, but this local action may have cascading effects with pervasive consequences for the structure of the phenotype. At the same time, it is very unlikely that a single gene may autonomously control a complex isolable component of the body, say the shape and organization of a complex organ: this will be typically done by many interacting genes. This state of affairs could hardly be advocated as pointing to inadequacies of DNA-based models of the shape and growth of the body.

The action of parameters is very local, particularly if the format is something like (14). But then some of these local actions may happen to be performed in structural positions close to certain crucial ganglia or crossroads of the system, hence give rise to systemic repercussions through interactions with other subsystems. For instance, the licensing of a null subject pronoun tightly interacts with various special properties of subjects: the obligatoriness of the subject position in the clausal structure (or the 'EPP' in traditional GB terms), the constraints on subject extraction and ECP effects (two properties that may well be closely related: Rizzi 2006b; Rizzi and Shlonsky 2007), properties of the Case-agreement system, interpretive properties associated with the subject position, etc. So we observe that null subject languages may have automatic and systematic access to null expletives permitting the formal satisfaction of the EPP property, hence endowing the system with more freedom on the surface distribution of the subject, with the options of subject inversion and free subject extraction (no manifestation of that-trace effects).⁷ Should we then expect a perfect correlation between such properties? In a system with few sparse parameters, this was a reasonable expectation, but in a system based on (14) we cannot expect such correlations to hold perfectly in general, simply because some other microparametric property of the language may affect the general pattern. For instance, the language might disallow extraction from a tensed clause altogether, hence make the presence of a potential 'skipping' device irrelevant. Analogously, we cannot expect non-Null Subject Languages to systematically manifest that-trace effects because other parametric options (such as a morphologically null version of the French *que* → *qui* rule) might create an independent skipping device, as presumably happens in the varieties of English not sensitive to that-trace,

⁷ In the terms of Rizzi and Shlonsky (2007), the null expletive offers a free skipping device from the freezing effects of the Subject Criterion: the expletive formally satisfies the criterion, and the thematic subject can be extracted from a lower position, thus skipping the freezing position). On the null subject parameter as a minimalist system see the discussion in Biberauer et al. (2010a).

Sobin (2002), in Norwegian, Taraldsen (1986), etc. Along similar lines, 'free subject inversion' in Romance null subject languages has been reanalysed as a device permitting focalization of the subject in a low, clause final position (Belletti 2004a). This property capitalizes on the availability of null subjects, but also requires an independent parametric option, the activation of the low focal position. So, certain Bantu null subject languages (Lingala etc.) do not have this option, hence they do not manifest the 'subject inversion' characteristic of Romance null subject languages.

In conclusion, there are very intricate cross-linguistic patterns of interactions which parametric theory can capture and elucidate, but, under current assumptions on the numerosity of parameters, there is no reason to expect that a single parameter could autonomously determine a complex cluster of properties. Of course, complex interactions are harder to fully spell out in a system with many parameters, but the deductive structure and its explanatory potential is intact, it only requires accuracy and ingenuity to be charted.

Fully elucidating the deductive consequences of a system of parametric values may be a difficult enterprise in general, but there are two kinds of privileged situations from which the endeavour is more immediately feasible. One is the case of very close grammatical systems. Consider the abstract situation in which, literally, 'all other things are equal', i.e. in the case of two systems differing for only one parameter, thus avoiding a priori the potential interfering effects of other parametric differences. Of course, such an extreme case never arises in practice; but reasonable approximations may be found through the microcomparison of historically very close grammatical systems, i.e. in the cases provided by dialectological studies. This is the microcomparative perspective, the closest approximation to a controlled experiment in comparative syntax, as Richard Kayne pointed out (see Kayne 2000a, 2013 for discussion).⁸

Another possible way to chart deep deductive consequences of parametric values is to identify a domain which, because of its very nature and properties, is relatively insulated from too many parametric interactions. Consider for instance Cinque's (2005) revisitation of Greenberg's (1963) Universal 20: roughly, when N is final, the only attested order of demonstrative, numeral, and adjective is the one illustrated by English—*These three nice books*—, while when N is in non-final position many orders are possible, some very frequent and some rare. Cinque is able to explain this non-trivial pattern of variation in terms of general properties of syntactic computations interacting with a handful of possible parametric choices: the basic order, determined by selection and scope properties, is Dem Num Adj N; it may be altered by stepwise N movement (as a head, or as a nominal projection) with or without pied-piping, determining different possible orders (with the relative frequency of the order correlated to the complexity of the required pied-piping operation); but if N does not move and

⁸ An extremely minimal case is the comparison of two registers of the same language, which may also be amenable to a parametric analysis: Haegeman (2013).

remains in final position, there is no other possible source of movement to reshuffle the elements, and the basic order always surfaces. The basic ingredients of Cinque's explanation thus are External Merge (determining the uniform basic ordering) and Internal Merge, or movement, of N. The parametrization on movement (if it takes place, if it involves pied-piping, what kind of pied-piping is involved) determines the observed cross-linguistic variation in the derived order. The unattested orders are simply underivable by such restrictive mechanisms.

Why is it that clear cross-linguistic patterns emerge in the nominal domain, while things are more opaque in the verbal or clausal domain? Presumably, this is due to the fact that the nominal system is relatively insulated, in the sense that major movement processes determining word order in clauses, e.g. movement to express scope-discourse properties of operator scope, focus, topicality, etc., pervasive at the clausal level, typically do not apply (or apply in a reduced form) in the nominal domain, thus excluding important interacting factors which blur the word order picture (but see Cinque 2013 for an attempt to adopt the same logic for the clausal domain).

In conclusion, parametric choices give rise to deep deductive chains with systemic effects. In order to fully capture this deductive structure, blurred in other cases by the numerous interactions with other parametric values, one may conduct microparametric analyses of very close systems, which are differentiated by relatively few parameters (as in much current Romance and Germanic dialectology), or macroparametric analyses of relatively insulated subsystems, like the nominal system. In general, both micro- and macro-comparative dimensions are needed: microparametric studies offer optimal conditions for identifying the irreducible parameters of the system; macroparametric studies (on which see Baker 2001, 2013) allow us not to lose track of the grand picture of language variation.

2.7 Broadening the picture: Language faculty vs. language as culture

The critiques addressed in the previous section target specific aspects of the parametric models, but do not question the general view that cross-linguistic variation is tightly constrained by a dedicated language faculty, part of the biological endowment of our species. This view, though, is not uncontroversial. More radical critical appraisals focusing on language diversity have questioned the existence of a prevailing cross-linguistic invariance: in the words of a contribution which recently attracted much attention (Evans and Levinson 2009), linguistic universals are a 'myth' which does not withstand empirical scrutiny, a view which echoes a dominant analytic tradition in American structuralism (Joos 1957 etc.), later marginalized under the impact of the ideas and discoveries of generative grammar. In this view, language is a cultural product, and virtually no aspect of the structure of language is immune from variability.

I will not address such radical critiques of the programme of generative grammar here (see many short responses to Evans and Levinson 2009 in the same issue of *Brain*

and *Behavioral Sciences*; I have expressed my own view in more detail in Rizzi 2010, and many other responses are available in the current literature). I will simply hint at aspects of this debate which are directly related to the previous discussion of parameters, and, in the final part, I will discuss some acquisition research which bears on the general issue.

Granting the importance of carefully studying language diversity, I believe that there are serious reasons to doubt the validity of the conclusion that, in Joos' (1957) words 'languages can vary without assignable limits...'. It simply is not the case that 'anything goes' in cross-linguistic variation. The general architecture of language is constant, there is a structured system of strict language universals (e.g. the ubiquitous role of hierarchical structure illustrated by the pervasive c-command effects that are observed in language after language), and also in domains in which variation seems to be the dominant factor, precise patterns of exclusion emerge (as in the Greenberg-Cinque discussion of word order properties already referred to). Clearly, though, abstract properties may take very different superficial forms because of the complex interactions between parametric choices, as we have seen, so that the underlying uniformity may need a lot of detailed analytical work to be detected. Consider, for instance, the issue of non-configurationality, already hinted at in section 3. Languages with very free word order (literary Latin, Australian Aboriginal languages, etc.) seem to defy configurational laws, not respect hierarchical constituent structure, and, in short, use representations different from Merge-based hierarchical structures. In fact, for some time the hypothesis was seriously entertained and explored, also within mainstream generative grammar, that such languages may differ from configurational languages in substantive ways. Later, careful analytic work stemming from the same tradition on extreme cases of 'free word order' languages, such as the Australian language Warlpiri, showed the pervasive presence of configurational effects such as the role of c-command, symptoms of hierarchically articulated constituent structure, rigid word order in certain environments, and various exclusion patterns: we are far from an 'anything goes' situation in the syntax of such languages and we find clear signs of a configurational organization, blurred in part on the surface by certain language specific properties (Legate 2002, 2008). Analogously, properties of classical free word order languages turned out to be analysable with revealing results through the configurational tools developed for unquestionably configurational languages (e.g. Salvi 2005, Dankaert 2012 on Latin using the tools introduced for the study of Romance left periphery in Rizzi 1997 etc.). In conclusion, assessing the configurational (or other architectural) properties of language requires much detailed analytical work on individual languages: the simple scrutiny of superficial properties will not allow us to reach firm comparative conclusions, such as the proper assessment of hypotheses on the universal structure of language. As soon as a detailed analytical work is undertaken, much as in the cases just quoted, a rich invariant structure always emerges from the variability of surface arrangements. This is true for configurationality and other word order

properties, for properties of binding and case-agreement systems, for movement and locality, etc. (see many contributions in Rizzi 2013b for relevant discussion of these points).

2.8 On the early acquisition of abstract grammatical properties

Very different kinds of evidence can be brought to bear on the broad divide between 'language faculty' and 'language as culture' approaches. In this final part, I would like to briefly review some evidence coming from the study of language acquisition. The timing of the acquisition process matters here. The 'language faculty' approach naturally leads to the expectation of a fast acquisition of cross-linguistically variable properties. In this approach, the problem that the language learner is confronted with is very well defined and narrowly circumscribed: as far as syntax is concerned, it is a matter of fixing the parameters of (the functional lexicon of) the language, and the learner is guided by task-specific cognitive resources which allow her to quickly converge to the correct parametric values. The 'language as culture' approach, all other things being equal, leads to the expectation of a slower acquisition process, basically in line with other aspects of the development of general problem-solving capacities and the acquisition of cultural skills. So, one would expect a certain degree of correspondence between the acquisition of variable properties of language and the acquisition of culturally-driven technical abilities of various sorts.

Let us address the question of the time course in connection with the acquisition of a fundamental cross-linguistically variable property: word order, and in particular the VO vs. OV order of the language. How early is this property acquired by the language learner? Corpus studies are unambiguous on this point: already in the first syntactically relevant productions, in the two-word stage, the child conforms to the target order: so the two-year-old learning English will typically say 'eat cake', and the two-year-old learning Japanese will say 'cake eat' (modulo morphophonological and lexical choices).

This is acknowledged by everyone, but the interpretations given by the two camps are very different. The language faculty approach typically assumes that the child has from very early on the abstract grammatical knowledge 'my language is VO', 'my language is OV', as a consequence of the early fixation of an ordering parameter (whether this is a merge, move, or spell-out parameter, as per our previous discussion, is not crucial here).

On the contrary, the 'language as culture' approach, represented here by the 'constructivist' or 'item-based' acquisition hypothesis proposed by Michael Tomasello and his associates in a number of papers (Tomasello 2000, 2003; Akhtar and Tomasello 1997; Tomasello et al. 1997), assumes that the child initially memorizes fragments she hears, and stores in memory individual items with the associated syntactic environments. There is no abstract generalization initially; there is only memorization of

fragments, individual items with the syntactic structures in which they are found. The hypothesis is that for a while the child stores this item-based knowledge and retrieves and reproduces it in her early productions; only much later on does the child generalize the item-based knowledge to abstract and general grammatical statements like 'my language is OV (or VO)' through a domain-general capacity for analogical generalization.

So, both approaches are consistent with the corpus data; but they lead to clearly different expectations about the child's early capacity to generalize her knowledge to new items and structures: the parametric approach leads one to expect that there should be an immediate generalization to a novel item because the relevant knowledge is abstract from early on; on the contrary, the 'constructivist' approach expects that the young child should not be able to generalize because her initial knowledge is concrete, item-based (she hears and memorizes 'eat apples', and obediently reproduces 'eat apples'), hence initially she has no basis to generalize to new items.⁹

2.9 Some experimental evidence

Franck et al. (2013) recently provided experimental evidence bearing on this question. In order to test the abstract grammatical knowledge of 19-month-old infants exposed to French, these authors combined three ingredients:

1. The preferential looking paradigm: the infant sits on her caretaker's lap in front of two computer screens, and hears a sentence. The two screens reproduce short videos with two distinct actions, one matching and the other not matching the uttered sentence. The child looks preferentially (for a longer time) at the screen with the matching video (see Naigles 1990; Gertner et al. 2006; Hirsh-Pasek and Golinkoff 1996 for detailed discussion of this method).
2. The 'weird word order' paradigm: the uttered sentence is sometimes an NP V NP sequence (grammatical in French), and sometimes an ungrammatical NP NP V sequence (this method is borrowed from production experiments reported in Abbot-Smith et al. 2001; Akhtar and Tomasello 1997; Matthews et al. 2005, 2007, and adapted to comprehension).

⁹ To be fair, neither approach is structured enough to make a very precise prediction on the time course of the acquisition of such abstract properties; nevertheless, within the parametric approach the straightforward interpretation of the target-consistent ordering in the two-word stage (hence before the second birthday) is that the relevant parameter has already been correctly fixed at this point (much as in Wexler's 'very early parameter setting', see Wexler 1998), while constructivist approaches seem to assume that abstract knowledge will arise through analogical generalization only well after the third birthday (consider, e.g., the fact that Matthews et al., on which see later in the chapter, compare a younger group around age 2.9 and an older group around age 4 in view of showing the abstract character of linguistic knowledge in the second group). So, even though the two approaches do not generate sharp predictions about the exact time course of the acquisition of abstract knowledge, they clearly lead to quite distinct expectations about the earlier or later character of such acquisition.

3. Pseudo-verbs are used, morphophonologically possible items which are not listed in the French lexicon, so that we can be sure that the child has never heard them in her previous experience.

Concretely, there are two conditions: grammatical (NP V NP) and ungrammatical (NP NP V) sentence. In the grammatical condition the infant hears a sentence like *Le lion dase le chien* 'the lion dases the dog', *daser* a possible but non-existent French verb. One of the videos reproduces a transitive action (for instance, the lion puts a crown on the dog's head), and the other video a reflexive action (each one of the characters puts a crown on his own head). In the ungrammatical condition the infant hears an ungrammatical sentence like *L'âne le chat poune* 'the donkey the cat pounes', a sentence violating the SVO order of French, with *pouner* a possible but non-existent French verb. Attention is paid to assign a natural-sounding prosody to the ungrammatical sentence, so that no obvious prosodic cue will mark it as deviant. As before, one of the videos reproduces a transitive action (for instance, the donkey puts a crown on the cat's head), and the other video a reflexive action (each one of the characters puts a crown on his own head).

The two approaches make clearly distinct predictions here. The parametric approach predicts a preference for the transitive video in the grammatical NP V NP condition, and no preference in the ungrammatical condition: in this approach it is natural to expect that at 19 months, or 1.7 years, around or right before the onset of the two-word stage, the infant will already have the abstract knowledge 'my language is SVO'. So, as soon as she hears a sentence like *Le lion dase le chien*, even if she has never heard that particular verb, she will immediately recognize a transitive NP V NP, or 'agent – action – patient' sentence scheme and will look preferentially at the transitive video. On the other hand, the ungrammatical sentence *L'âne le chat poune* will not evoke any abstract grammatical scheme in French, so the sentence will not offer any guidance to the child to preferentially look at one or at the other video.

An item-based approach assuming no abstract grammatical knowledge in young children, on the other hand, does not predict any preference in either case. As in this approach the infant does not have any general grammatical scheme to build on, but only item-specific knowledge, she would have no good reason to prefer the transitive action only with the grammatical NP V NP order: both in the grammatical and ungrammatical order she has not previously heard the occurring verb, hence in neither case does she have previous item-based knowledge to build on. So, no preference for a particular video specifically linked to the grammatical word order is expected in either case.¹⁰

¹⁰ More precisely, the item-based approach would lead us to expect no preference *specifically linked to the grammatical word order*: it would be consistent with a grammatically unselective preference that children might have, e.g. a general preference for transitive videos over reflexive videos, irrespective of the grammatical or ungrammatical character of the sentence which is uttered. So, crucial evidence to disentangle the two approaches can be provided by the existence of a contrast (or lack thereof) between the grammatical and ungrammatical condition.

The experimental evidence clearly is in line with the expectations of the 'abstract grammar' approach: it is reported in Franck et al. (2013) that infants look at the transitive video significantly more than at the reflexive video in the grammatical NP V NP condition, while they show no preference between the two videos in the ungrammatical NP NP V condition (hence one cannot say that they prefer to look at transitive actions in general, regardless of the sentence they hear). So, the child acquiring French at 19 months appears to have abstract knowledge of the type 'my language is SVO'.¹¹

There is an apparent contradiction between these results and the conclusion reached, e.g. by Matthews et al. (2005, 2007) on the basis of production experiments. They elicited the repetition of sentences with pseudo-verbs which had been presented both in grammatical and weird word order; their claim is that older children (at 4 years) correct more weird word order sentences than younger children (at 2 years 9 months), who reproduce sentences in the weird word order more frequently than the older group. These authors thus claim that their production study supports the constructivist position: younger children at age 2.9 only have an item-based knowledge, which does not allow them to correct ungrammatical orders on the basis of an abstract grammatical scheme. This result clearly conflicts with our result in comprehension, which shows abstract grammatical knowledge already at age of 1.7. Should one postulate a major divide between production and comprehension systems with respect to the availability of abstract grammatical properties?

Franck et al. (2011) have redone the Matthews et al. (2005, 2007) experiments by introducing certain modifications in the methodology, in particular by improving the communicative situation; they found that younger children acquiring French at 2.11 were not distinguishable from older children at 3.11 in the repetition of grammatical and weird word order sentences, showing as much abstract grammatical knowledge as the older group: both groups were found to match the grammatical word order significantly more often than ungrammatical word orders, also with pseudo-verbs they had not heard before. Moreover, both younger and older children's productions gave clear indications of morphosyntactic productivity in the grammatical NP V NP order, sometimes modifying the input to produce sentences like *La vache, elle a dasé le chien* 'the cow, it has dased the dog' with pronominalization, dislocation, the introduction of compound tenses, etc. In contrast, children in both groups failed to manifest any sign of productivity in the rare ungrammatical NP NP V sentences they produced: no compound tenses, no special inflectional properties on the verb, no pronouns, dislocations, or other manipulations in their ungrammatical NP NP V sentences, which were systematically produced with full NPs and verbs in the present tense exactly as they appeared in the input. Both groups of children therefore used their productive grammatical knowledge when they produced sentences in the grammatical order, while

¹¹ On the possible prosodic cues or statistical analysis which may guide the child to fix this fundamental word order property very early on, see Christophe et al. (2003); Gervain et al. (2008).

they just repeated the input string in the (rare) occasions in which they reproduced the ungrammatical NP NP V order. These authors therefore conclude that the younger group also shows grammatical knowledge of abstract word order properties: there is no basis for assuming an asymmetry between the two groups, nor between production and comprehension (except that, of course, production could not be tested in a reliable manner with children as young as 1;7, as they are just entering, or about to enter, the two-word stage). Franck et al. (2011) then conclude that when production is tested in plausible communicative conditions, children of the younger group show no less abstract knowledge than children of the older group. This is in line with the result of the comprehension experiment, and is what the language faculty approach would lead us to expect.¹²

2.10 Conclusions

Parameters of syntactic variation can be thought of as morphosyntactic features expressed on the items of the functional lexicon and acting as instructions for the basic syntactic actions: Merge, Move, Spell-out. Parameters are numerous because their locus of expression, the functional lexicon, is rich; nevertheless, the space of variation is severely constrained because the possible syntactic actions in a minimalist model are so limited. Combining the central idea of the principles and parameters approach with minimalist syntax thus yields a coherent, restrictive system for the study of language variation. The numerosity of parameters makes it unlikely that a single parameter may be able to fully control a complex cluster of properties, because there will inevitably be too many interactions with other parametric values (with the possible exception of Kayne's 'controlled experiments' in comparative syntax, the privileged cases arising from the micro-comparative analysis of very close varieties, and approximating the ideal of two systems differing for a single parametric value; and of the macro-comparative study of structural systems which are sufficiently insulated to limit parametric interactions). The complexity of the interactions does not mean that the system has a limited deductive structure and that each parameter only has local consequences. Quite the contrary is true: each parameter will enter into complex deductive interactions with principles and other parametric values, and disentangling and reassembling the elementary components of such interactions will continue to shed light on the observed, complex patterns of variation.

In the last part of the chapter, I have broadened the perspective to the general issue of the nature of cross-linguistic variation, and the plausibility of assuming dedicated

¹² See also Franck and Lassotta's (2012) detailed critical discussion of the methods, results, and argumentation of papers using the Weird Word Order paradigm in production (Akhtar 1999; Matthews et al. 2005, 2007, etc.). Under Franck and Lassotta's reanalysis, the data presented in these papers actually support the hypothesis that children have abstract grammatical knowledge of word order from early on.

cognitive resources constraining linguistic variability. Relevant evidence here can be gathered from comparative syntax, but also from the study of the timing and characteristics of language acquisition (and, in principle, from many other sources: pathology, brain imaging, etc.). I have focused on one particular case study: the rapidity of the acquisition of language-particular word order properties in the form of abstract and general grammatical knowledge is unexpected under views looking at language as a cultural object, with the acquisition of variable properties solely guided by general intelligence and general problem-solving skills, much as the acquisition of a simple technology of some kind; the evidence just reviewed is more readily consistent with a view in which the child is guided very early on to have certain expectations about structural properties of the language, and to quickly make well-defined choices of a rather abstract character, as in parametric models.