On the format and locus of parameters: The role of morphosyntactic features.

1. Introduction

The idea that linguistic variability could be reduced to a limited number of parameters introduced a new technical language for comparative formal syntax, and was a decisive factor in the enormous growth of comparative studies over the last thirty five years. The notion of parameter of the early parametric models has undergone significant changes under the pressure of theoretical advances and empirical evidence: the way in which we can conceive of parameters in current minimalist models is clearly different from early formulations in Chomsky’s Lectures on government and binding and in the literature in the early 1980’s. Nevertheless, in my opinion there is a fundamental conceptual continuity between early and more recent approaches to language variation which justifies the use of the same term to refer to the irreducible choice points of the system, and would make a terminological change misleading.

In this article, I will start by a quick survey of the historical development by referring to the two key notions of the format and locus of parameters across different successive models of comparative generative grammar. Then I will address the question of how and where parameters are expressed within current minimalist models. I will make explicit a conception of parameters as morphosyntactic features expressed on functional heads and providing explicit instructions for the triggering of syntactic operations. This will lead to introducing a typology of features, hence of parameters, on the basis of the operations they trigger: there are merge parameters, operating on structure building, movement parameters, triggering different kinds of movement, and spell-out parameters, giving instructions on the syntactic positions that are or are not pronounced.

Under the Borer – Chomsky conjecture on the locus of parameters, the size of the set of parameters is determined by the size of the functional lexicon, hence a parametric system based on these assumptions has many parameters. In spite of that, the space of variation is drastically limited by the tight constraints operating on minimalist syntax: only very few operations are possible in syntactic computations. So, the large number of parameters notwithstanding, the system is radically more constrained than approaches to linguistic variation based on an unlimited set of language particular rules, interacting with a much less constrained notion of Universal Grammar, as in pre-parametric models of the Extended Standard Theory.

In the final part of the article I will address the issue of the elements of variation in structural maps emerging from cartographic work. Possible forms of parametrization arising in functional sequences will be illustrated through the case study of the uniqueness or multiplicity of topics in the left periphery of the clause.
2. Some elements of the history of parameters.

The first twenty years of the history of generative grammar mainly focused on linguistic uniformity. This was due in part to a contingent reason: the available data base consisted largely, if not uniquely, of a body of empirical analyses of English, so that little if any formally analyzed evidence was available about language variation. Nevertheless, a crucial assumption was that grammatical fragments of a single language would directly bear on issues of universality through poverty of stimulus arguments: if from the analysis of a single language certain properties emerge which, plausibly, are not inductively determinable from the data available to the language learner, such properties have to be deductively connected to some general property of the language faculty, hence tell us something on the structure of Universal Grammar (UG). The Standard Theory and the Extended Standard Theory of the 1960’s and 1970’s of course also assumed devices to express language variation: individual languages were assumed to be generated by particular grammars, consisting of rules specific to the particular language, and respecting general constraints on rule format and application defined by UG. One major problem with this conception came from the insufficiently constrained character of the possible language-particular rules: the search space available to the language learner remained much too vast to permit a reasonable account of language acquisition. Therefore, reaching the level of “explanatory adequacy” (Chomsky 1964; see Rizzi 2015a for a recent discussion of this notion) remained a distant goal.

Things changed suddenly around the end of the 1970’s, with the introduction of the Principles and Parameters model (systematized in Chomsky 1981 on the basis of much research conducted in the previous years). If the possibility of a parametrization of individual principles and rules had been abstractly envisaged before, the first empirical motivation came from the observation that certain island properties, such as the (non-) extractability from indirect questions and from subject noun phrases turned out to manifest a certain level of variability (Rizzi 1978, 1982, ch. II). Poverty of stimulus considerations made it unlikely that the relevant locality principle could be figured out from scratch by the language learner on the basis of the primary data available to him/her; moreover the observed variability clearly appeared restricted to a narrow range of options, so it seemed necessary to maintain the hypothesis that some UG principle was involved. On the other hand, the assumption of a strong universal UG based component had to be reconciled with the observed variation. So, the idea was explored that the relevant locality principle, Subjacency (first introduced in Chomsky 1973) could contain a parameter, a choice point determining a narrow variation in the operation of the principle: in particular, the choice of the bounding nodes, counting for the computation of locality, could be language-specific to some extent (in the original terminology, either S or S’, currently IP or CP, could be the clausal bounding node).

In hindsight, this kind of variability turned out to be rather marginal (with variation within the same language: Grimshaw 1986), and not even naturally amenable to more modern conceptions of parametrization, such as those discussed below. Nevertheless, the S/S’ discussion had the merit of offering the first concrete case of a formal device specifically intended to address cross-linguistic variation in a UG-based system, an idea which immediately showed a vast potential for the analysis of language diversity. The idea that language variation could be reduced to the fixation of a finite number of binary parameters introduced a precise theoretical language to address invariance and variation, and this simple innovation immediately showed a great heuristic capacity, and made formal comparative syntax possible, and attractive. In very few years the data base of languages analyzed through
generative tools grew enormously, and for the first time the comparative dimension got to center stage in generative grammar.

In discussing properties of different conceptions of parametrization it is useful to isolate two questions:

(1) What is the *format* of parameters? i.e., what constraints should we expect on possible forms of parametrization?

(2) What is the *locus* of parameters? i.e., where are parameters expressed in the general grammatical architecture?

In the early years of parametric theory, not much attention was devoted to the issue of the format: more or less anything could be a possible parameter, and proposals varied from word-order properties (e.g., OV vs VO), to the licensing of null elements of various kinds (subjects, objects, etc.) to properties of the triggering of movement (V to T, V2, etc.), to the choice in the component in which a given operation took place (e.g., wh-movement in overt syntax or in the syntax of LF in *in situ* languages, affixation in syntax or in the PF component, etc.), to much more global properties of grammatical systems, such as the configurationality of the language (Hale 1983).

As for the locus of parameters, as the S/S’ parameter looked like a specification on a given principle, Subjacency, it seemed natural to generalize this case, and assume that UG principles could be, in general, the place in the global grammatical architecture in which parameters were expressed.

But this conception of the locus of parameters quickly turned out to be dubious. On the one hand, certain UG principles did not show any plausible kind of parametrization (e.g., the Theta Criterion); on the other hand, and more importantly, some properties of variation turned out to be linked to the presence of particular items in the functional lexicon, rather than being global properties of certain principles or modules of grammar. So, for instance, long distance anaphora, a much studied element of variation in the functioning of the binding theory (Manzini & Wexler 1987), seemed to depend on the presence of particular items in the functional lexicon, *sig* in Icelandic, *sé* in Italian, etc., rather than being a global property of binding in a given language (i.e., other items, such as *se stesso* in Italian, do not manifest the long distance binding option: Giorgi 1984); the prepositional or postpositional character of an adposition depended in part on the lexical specifications (arguably even in largely uniform languages like English: *after three weeks, three weeks ago*), etc..

This kind of consideration lent support to a different view on the locus of parameters, often called the Borer – Chomsky conjecture (see Borer 1983):

(3) Parametric values are expressed in the functional lexicon.

This conception leads to the expectations that parameters are more numerous than in the initial conception of parameters expressed on principles: principles are few in a grammatical architecture such as the Government-Binding framework (and even fewer in a minimalist grammar), whereas the functional lexicon is a rich inventory, all the more so if cartographic studies are on the right track (see below). Under the Borer-Chomsky conjecture, there are many more opportunities for parametrisation and the size of the functional lexicon offers a
reasonable basis for an estimate of the numerosity of parameters, which then shifts from the order of tens to the order of hundreds (or more). Correspondingly, in a system with numerous parameters each parameter is expected to be more local and limited in scope than in a system with few parameters. In particular, it is less likely that a single parameter will autonomously control complex typological clusters, because it will interact with many other parameters which will make the deductive connection between a single parameter and large arrays of properties more complex and indirect.

Evidence for such a more local and diffuse effect of individual parameters came from the observation that certain parameters assumed initially as directly and autonomously controlling several properties turned out to interact with other parameters which limit their scope. For instance, the Null Subject Parameter (Rizzi 1982), initially assumed to be directly responsible for so-called “subject inversion” in languages like Italian, interacts with another parametric property, at least partially independent, having to do with the activation of a vP peripheral focus position which hosts postverbal subjects in languages like Italian (Belletti 2004, and, for general considerations on the Null Subject Parameter in a system analyzing subject inversion as focalization, Nicolis 2005: see below). So, some languages licensing null pronominal subjects have clause-final subject focalization, and others do not.

More radically, it was proposed that certain global properties initially assumed to be controlled by large parameters actually dissolved into the operation of smaller parameters. So, freedom in word order is not controlled by a single global “non configurationality” parameter, but rather dissolves into the combination of smaller properties: null subject and null object languages are typically freer in the positioning of the subject and object DP’s, respectively, than other languages; scrambling languages are freer than non-scrambling languages; languages permitting the disintegration of nominal structures (possibly as a consequence of properties of the determiner system: Boskovic 2009) are freer than languages keeping DP’s intact; etc. so, there are in fact distinct degrees of freedom of word order as a function of the fixation of more local parameters, rather than a unique binary opposition between configurational and non-configurational languages. This is the state of affairs expected under the microparametric view advocated by Kayne 2000, 2005, and directly connected to the Borer – Chomsky conjecture (now largely assumed, but not completely uncontroversial: see Baker 2001, 2008 for discussion).

This conception of parametric theory has sometimes been criticized as involving an undeclared retreat to the idea of language particular rules (Newmeyer 2004, 2005): if there are so many parameters, how is parametric theory different from a theory permitting language-specific rules, as in EST models?

To address this point, the distinction between format and locus of parameters becomes important: a system with parameters expressed in the functional lexicon permits many parametric specifications, but may still define an extremely restrictive framework for linguistic variation: it all depends on the format for parameters that the theory assumes. This leads us to address the format of parameters in current grammatical models.

3. On the format of parameters.

What is a possible parameter in a minimalist grammar? I would like to adopt the following definition, adapted from Rizzi (2014):
(4) A parameter is an instruction for the triggering of a syntactic operations, expressed as a morphosyntactic feature associated to a functional head.

So, when a functional element enters syntax becoming a functional head, it will trigger certain syntactic operations on the structure which is being built, on the bases of the featural instructions associated to it. What kinds of operations can be triggered? The list of possible elementary operations permitted by a minimalist grammar is highly restricted. Keeping the discussion at a rather informal level, we can identify:

(5) a. Merge
    b. Move
    c. Spell out

Correspondingly, we may envisage merge parameters, controlling properties of structure building, movement parameters, controlling the various movement properties, and spell-out parameters, determining the pronunciation of a given head and of its immediate dependents. Each parametric property is strictly local, in that the triggered operation can only affect the immediate structural environment of the relevant head. Given the highly limited nature of the possible syntactic operation and the locality of their structural consequences, the system is radically more restricted than an EST type system based on an unlimited set of possible language specific rules. The search space within which the language learner must determine the specific properties of the language s/he is exposed to is restricted accordingly, so that the learnability problems that an EST type grammar had to face are kept under control (see the discussion in Rizzi 2014, which also addresses experimental evidence bearing on the early fixation of parametric properties).

Here I will adopt the view, central in minimalism, that each operation is triggered by a morphosyntactic feature. So, we may envisage the following general definition of the format for parameters:

(6) X has F

in which X is an element of the functional lexicon, and F is a morphosyntactic feature triggering syntactic operations of merge, move and spell-out. X may have F in one language, and not in another language, a binary choice.

The next question is: what is F, the class of relevant morphosyntactic features, hence of parametric properties? In order to structure the approach, let us now look in some detail at the morphosyntactic features which trigger syntactic operations. I will consider three broad classes of features, hence of parametric properties, along the lines of the typology of operations in (5).

3.1 Merge parameters.

They deal with all the properties of structure building which may vary across languages. While properties of semantic selection (s-selection, in the sense of Grimshaw 1979, Pesetsky 1982) are presumably invariant (apart from the possibility of interface parameters, as
advocated in Chierchia 1998), properties of categorial selection (c-selection) not directly
deducible from s-selection may vary. In the terminology and classification adopted here, they
are merge parameters (or, more precisely, external merge parameters, if we understand
movement as internal merge, see below). One familiar example may be the selection of a
reduced clausal complement by certain verb classes, e.g., epistemic verbs. Believe-type verbs
in English select a reduced clausal complement, permitting exceptional case marking and
inconsistent with control; in Romance they select a full clausal control infinitive introduced
by a complementizer system (with an overt prepositional complementizer in Italian and a null
C in French):

(7)a John believes [ Bill to be a nice guy ]
   b * John believes [ PRO to be a nice guy ]

(8)a * Gianni crede [ Piero essere una brava persona ]
   b Gianni crede [ di [ PRO essere una brava persona ] ]

(9)a * Jean croit [ Pierre être un type bien]
   b Jean croit [ [ PRO être un type bien] ]

The fact that these properties appear to affect verb classes, rather than individual verbs, is
consistent with the idea that the parametrization is linked to elements of the functional
lexicon, in this case an appropriately flavored instance of v, along the lines of Harley (2011),
Ramchand (2008) and related work.

Analogously, with perception verbs, Italian and French permit pseudo-relatives (Cinque 1990,
Guasti 1988, Casalicchio 2013) while English does not, and French and English permit
participial-gerundival complement clauses, while Italian does not ((11)b is fine in Italian, but
with the gerundival clause functioning as a clausal adjunct controlled by the subject “…while
I was coming back home”, not as a complement of the perception verb with Gianni as
subject):

(10)a * I saw John that was coming back home
   b I saw John coming back home

(11)a Ho visto Gianni che tornava a casa
   b * Ho visto Gianni tornando a casa

(12)a J’ai vu Jean qui rentrait à la maison
   b J’ai vu Jean rentrant à la maison

Examples of this sort could easily multiply. Merge parameters also include certain elements
of variation in functional sequences, such as the number and type of topic positions permitted
in the left periphery of a given language, the position of negation and agreement markers in
the functional sequence of IP etc. We will address such cartographic parameters in a separate
section.

3.2. Move parameters.

In order to illustrate this kind of parametrization, certain assumptions on movement must be
made explicit. Following current guidelines, I will assume the following:
(13)a. Move is a complex operation (Chomsky 2000) involving
   I. the establishment of a probe-goal search relation followed by
   II. (internal) merge of the goal.

   b. Movement may involve a head or a phrase.

If movement is a complex operation, then both components, search and internal merge, are a
priori parametrizable; and if head movement exists as a distinct operation from phrasal
movement (assumption (13)b), we need a way to differentiate the two, and express the fact
that they both admit parametrization. In Rizzi (2015c) I proposed to modify Bare Phrase
Structure by using the feature lex, distinguishing heads from phrases, a device which I will
assume here.

Let us start then from head movement parameters. Consider the difference between English
and French in terms of head movement. Both English and French have verbs agreeing with
subjects, but in French lexical verbs are attracted by the functional head bearing Phi features
(indicated as Phi in (14), the kind of position designated by the label AgrS in previous work),
while in English lexical verbs do not move, as is shown by the respective ordering with
adverbials (Emonds 1978, Pollock 1989):

(14)a  John \( \text{Phi} \) often sees Mary

   b  Jean voit+\( \text{Phi} \) souvent ___ Marie

So, the head indicated as Phi acts as a probe and launches a search operation in both
languages, looking for a verbal head with matching Phi features, the goal. Once the probe-
goal relation is established, the features on the goal are valued, and this determines the
agreeing morphology on the verb. At this point, the goal may be internally merged with the
probe in French-type languages, yielding the order in (14)b, or remains \textit{in situ} in English-type
languages.

Other languages, e.g. creole languages such as Jamaican (Durrleman 2008), do not show any
kind of morphological manifestation of agreement on the verb; they may involve a Phi-like
head attracting the subject and determining its (abstract) case, but they may not involve any
search from the Phi head for a particular verbal form, as V remains invariant. So, search may
or may not be triggered from Phi to V; if it is, V may or may not be moved to (internally
merged with) Phi. In short, we have the following parametrization for Phi involving its
syntactic (head – head) relation with V:

(15) \[ \begin{array}{c}
   \text{Search for V?} \\
   \text{Yes} \quad \text{No (Creole)}
   \end{array} \]

\[ \begin{array}{c}
   \text{Internal merge of V?} \\
   \text{Yes (French)} \quad \text{No (English)}
   \end{array} \]
Here I assume, as in Chomsky (2000), that search is a prerequisite for internal merge, and that this holds for head-movement as well. If a search relation is established, the language may trigger internal merge of the goal, or not. Therefore we have, in association with individual functional heads:

(16)a Search parameters: is a probe-goal search activated or not wrt a particular feature set?

b Internal merge parameters: once a probe-goal relation is established, is internal merge of the goal with the probe activated or not?

Consider now phrasal movement. Here I will continue to assume that internal merge presupposes the establishment of a probe-goal relation, followed by internal merge of the goal. As we are in the phrasal domain, the goal, a phrase, could not be internally merged with the probe, a head: heads may be complex, but they can only contain other heads, not phrases, a kind of extended structure preserving constraint called Lexical Uniformity in Rizzi (2015c). So, phrasal movement can only involve internal merge to the projection of the probe, rather than to the probe itself, which would violate Lexical Uniformity.

As in the case of head movement, once a search relation is established, the probe may or may not attract the goal to be internally merged with its projection. So, for instance, in wh-movement languages the probe, a Q complementizer, attracts the phrasal goal to internally merge with its projection, whereas in wh-*in situ* languages internal merge does not take place (or at least, not overtly). Both options are available in French main questions, so the following are both possible (the underscored elements have been connected by search and are in a probe-goal relation; I do not address here the question of whether qui undergoes covert movement in (17)a, and of what kind):

(17)a  Q  Tu as vu qui?
     ‘You have seen who?’

b  Qui  Q  tu as vu __  ?
     ‘Who have you seen ?

Analogously, in SVO languages we have a search from a Phi-type head to the closest nominal expression in its domain (again, the successful search is annotated by underscoring the probe and the goal in what follows), followed by internal merge of the goal to the projection of the probe Phi (whether or not Phi has also attracted V), whereas in VSO languages the search relation between Phi and S is established, so that agreement in Phi features is checked, but not followed by internal merge of the goal:

(18)  SVO:

\[
\begin{align*}
\text{Phi} & \quad S \quad V \quad O \quad \rightarrow \quad \text{Search} \\
\text{Phi} & \quad S \quad V \quad O \quad \rightarrow \quad \text{Internal merge} \\
S \quad \Phi & \quad _- \quad V \quad O
\end{align*}
\]

(19)  VSO:

\[
\begin{align*}
\text{Phi} & \quad S \quad V \quad O \quad \rightarrow \quad \text{Search}
\end{align*}
\]
(the correct ordering is derived in (19) once the probe Phi attracts the verbal goal; alternatively, Phi may host an auxiliary verb, in which case the order will be Aux S V O; the postverbal subject in VSO presumably has been moved from it first merge position in the vP, as is shown by the ordering V S Adv O, possibly for labeling reasons (Rizzi 2015b), but does not move as far as the Spec of Phi in this analysis of VSO).

In conclusion, a functional head acting as a trigger of movement may have two distinct pairs of features, responsible, respectively, for phrasal movement and head movement:

(20) For phrasal movement:
   a. A search feature at the phrasal level (SF).
   b. The corresponding internal merge feature at the phrasal level (IMF), what is traditionally called an EPP feature.

(21) For head movement:
   a. A search feature at the lex level (S_{lex}F)
   b. The corresponding internal merge feature, again at the lex level (IM_{lex}F)

So, the head carrying Phi in the clausal IP structure in a language like French carries all four features: a search feature for a noun phrase carrying matching Phi features, and the corresponding internal merge feature (what is traditionally referred to as an EPP feature) triggering internal merge of the goal (the subject) with the projection of the probe (not with the probe itself because we are here at the phrasal level, and a phrase cannot be internally merged with a head because of Lexical Uniformity); a search feature for a lex verbal element (a verb with matching phi features), and the corresponding internal merge feature triggering internal merge of the goal with the probe and yielding a complex head, a verb endowed with agreement features. So, we have a lexical representation like (22) for Phi in French:

(22) Phi: SF: Phi, D
     IMF: Phi, D goal
     S_{lex}F: Phi, v_{lex}
     IM_{lex}F: Phi, v_{lex} goal

The head is endowed with a search feature at the phrasal level (SF) establishing a search relation with a DP endowed with Phi features, and of the corresponding internal merge feature (IMF), instructing the system to internally merge the goal DP with the structure headed by Phi; the functional head Phi is also endowed with a search feature at the head level (S_{lex}F) establishing a search relation with a verbal head carrying Phi features, and of the corresponding internal merge feature (IM_{lex}F), instructing the system to internally merge the goal v with the probe Phi: here the goal can be merged with the probe head, under lexical uniformity, as both are heads (lex elements). This yields the following representation (omitting many details):
Phi in English would have the same specification, except that it would lack the IM\textsubscript{lex}F feature, thus no verb movement would be triggered (a finer distinction would be needed here to distinguish lexical from functional verbs, as the latter do raise to Phi; we will omit this refinement). At least for a first approximation, a VSO language would presumably involve Phi with the same specification as French, except that the IMF, triggering movement of the subject DP, would not be specified.

Let us now illustrate the system at the CP level. A functional head may have none of these features, in which case it is inert for movement, e.g., the complementizer \textit{that} of an embedded declarative clause; or it can have just one feature, e.g., the Q complementizer of a \textit{wh-in situ} language, triggering a search relation with an \textit{in situ} interrogative element; or two such features, e.g. a search feature for a \textit{wh}-phrase and an internal merge feature triggering merge of the goal with the projection of the probe, as in a \textit{wh} movement language without inversion (Hebrew and other Semitic languages, Brazilian Portuguese, various creole languages, for instance); or three or all four features. The last case, a language with a complementizer head attracting both a \textit{wh}-phrase and a head of the IP system, may be illustrated by the Q complementizer in English main clauses, with lexical representation (24), and yielding a structural configuration like (25) (I assume that in inversion structures the relevant complementizer Q attracts an occurrence of T endowed with Q):

\begin{equation}
(24) \quad \text{Q: Search:Q} \\
\quad \text{IM: Q goal} \\
\quad \text{Search\textsubscript{lex}} : T-Q\textsubscript{lex} \\
\quad \text{IM\textsubscript{lex}} : T-Q\textsubscript{lex} \text{ goal}
\end{equation}

\begin{equation}
(25) \quad \text{DP} \\
\quad \text{Which book} \\
\quad \text{you read} \\
\quad \text{did} \\
\quad \text{Q}
\end{equation}

In conclusion, move parameters involve two pairs of featural instructions, one involving search and internal merge for a phrase, and the other involving search and internal merge for a head (a lex element, in the formalism adopted here); phrasal (non-lex) internal merge involves merge of the goal phrase with the projection of the probe, while head (lex) internal merge involves merge of the goal head with the probe, in accordance with Lexical Uniformity. This
mechanism handles all major cases of phrasal (A and A’) and head movement as cases of triggered movement. If untriggered movement exists, i.e. a set of movement operations not depending on a previously established probe-goal relation, or even, more broadly, there are kinds of movement not depending on specifications of morphosyntactic features, extra options should be added, and it is not clear how possible parametrisations could be expressed in a way consistent with the Chomsky – Borer conjecture. I will not go into the many ramifications implied by such options, and, for the sake of this discussion, I will simply assume that untriggered movement does not exist.

Notice that in my illustration of movement to Phi and to Q I have kept the traditional assumption that the same head Phi attracts the inflected verb and the subject DP, as in the French case, and that in the C system, the same head attracts the T bearing element and the wh phrase. A conceivable alternative is that each individual functional head is endowed at most with a single pair of features triggering movement, so that an individual head is a trigger for either head movement or phrasal movement, but not for both simultaneously. If so, the relevant cases would have to be reanalyzed as involving two distinct heads, the lower one triggering head movement and the higher one triggering phrasal movement. The possible interpolation of various kinds of adjuncts in French between the subject and the inflected verb suggests that distinguishing two separate heads as triggers of the two types of movement is at least an option:

(26) Jean, à mon avis, trouvera la solution
    ‘Jean, in my opinion, will find the solution’

Similarly, a construction like complex inversion in French (Rizzi & Roberts 1989), with the lexical subject occurring between the wh-element and the tensed verb (which in turn precedes the subject clitic, suggesting that it has moved to the C-system) supports the splitting hypothesis:

(27) Où Jean est-il allé?
    ‘Where Jean did he go?’

So, the possibility of expressing attracting features for head and phrasal movement on separate heads clearly is an option. I will leave open whether this splitting is just an option, or the splitting is what happens in general in cases of double attraction.

3.3. Spell-out parameters.

Much early work on parametric theory revolved around the Null Subject Parameter: some languages license a phonetically null subject pronoun, and this apparently low-level property connects with other structural properties of the language. This parameter then is, for historical reasons, the most important representative of spell-out parameters, having to do with variation in the obligatory, optional or impossible pronunciation of certain heads and of their immediate dependents.

In the original work, a single parametric property was intended to capture a cluster of properties which clearly differentiate non-null subject languages like French and English from null subject languages of the Romance family, such as Italian and Romanian (Rizzi 1982, ch. IV). Italian and Romanian allow null pronominal subjects with referential and non-referential
interpretation (as in (28)a’-b’, (29)a-b, respectively); they also allow subject inversion, with
the overt subject in clause-final position, as in (30)a’-b’, alternating with the clause initial
position; and free violations of that-trace effects, as in (31)a-b. English and French do not
allow any of these properties (as in the c-d, c’-d’ examples).

(28)a  Io parlo italiano  a’ ___ parlo italiano
  b  Eu vorbesc italiană.  b’ ___ vorbesc italiană.
  c  Je parle l’italien  c’ * ___ parle l’italien
  d  I speak Italian  d’ * ___ speak Italian

(29)a  ___ piove
  b  ___ plouă.
  c  Il pleut
  d  It is raining

(30)a  Gianni ha telefonato  a’ ___ ha telefonato Gianni
  b  Ion a telefonat.  b’ ___ a telefonat Ion
  c  Jean a téléphoné  c’* ___ a téléphoné Jean
  d  John telephoned  d’* ___ telephoned John

(31)a  Chi credi che ___ verrà?
  b  Cine crezi că ___ va veni?
  c  * Qui crois-tu que ___ viendra?
  d  * Who do you think that ___ will come?

In the early days of the parametric approach, a systematic attempt was made to establish
deductive connections between these properties, tracing them back to a single irreducible
difference, the fixation of the NSParameter interacting with the general structure of Universal
Grammar. The adopted formulation of the parameter was along the following lines:

(32) **pro** is licensed by the verbal inflection {yes, no}

Relevant UG principles interacting with the parameter were the EPP, expressing the
obliterariness of the subject position of clauses, irrespective of the thematic properties of the
verb, and the ECP, banning traces from non-properly governed positions, such as the subject
position.

So, in English and French, (28)c’-d’, (30)c’-d’ would be excluded because there was no
appropriate null filler for the obligatory subject position (under the EPP), and (31)c-d would
be excluded as an ECP violation. In Italian and Romanian (28)a-b and (29)a-b would be fine,
with the legitimate null pronominal with a referential or expletive interpretation; and the
preverbal subject position could be filled by the expletive **pro** in (30)a-b and (31)a-b, thus
permitting more freedom in the positioning of the thematic subject; in particular, subject
extraction would be possible in (31)a-b because extraction could take place from a lower
properly governed position.

In the following thirty years or so, much theoretical and empirical work led to a rather
different global picture. Belletti (2004) showed that the inversion operative in (30) does not
involve an optional movement process, but is in fact a device for focalizing the subject in the
low vP periphery, an option which is governed by an independent parameter. So, some null subject languages, like Italian and Romanian take this independent option, while other null subject languages, like the Bantu language Lingala (Salulessa 2004) do not. The null subject parameter thus interacts with at least another parameter in determining the pattern in (28)-(31). Another important line of research was devoted to the comparative study of partial null subject languages (Hebrew, Brazilian Portuguese, Finnish,...) which permit expletive null subjects but show important limitations in permitting null pronouns with referential interpretation (Biberauer, Holmberg, Roberts, Sheehan 2010). On the theoretical dimension, the possibility was explored that the licensing of the null subject may in fact be the result of a deletion operation (Biberauer et al. 2010), and that the impossibility of extracting the subject may be a consequence of a general freezing principle, rather than of the ECP (Rizzi & Shlonsky 2007). So, a substantive body of theoretical and empirical work was devoted to the topic, ideas and analytic tools were refined and changed, and much progress was made. Nevertheless, it seems to me that the core of the original approach has remained intact: a complex cross-linguistic pattern can be deductively connected to a parameter on the licensing of a null pronounal, interacting with UG principles and other parameters. The numerosity of the parameters, under the Borer-Chomsky conjecture, makes it unlikely that a single parameter may govern a complex pattern in isolation from other parametric choices, in this and many other cases: there are just too many parameters, so that interactions will inevitably arise. And indeed we have seen that more than a single parameter determines patterns like (28)-(31). But the important point is that the deductive depth of the system remains intact: so, the conclusion remains valid that null subject languages typically are insensitive to that-trace effects because the availability of the null expletive offers a well-formed filler for the freezing subject position, thus allowing the thematic subject to circumvent the freezing effect and remain extractable (Rizzi & Shlonsky 2007). Interactions are more complex in a system with many parameters, but the deductive depth of the system is unaffected; in fact, if anything, deductive chains tracing back empirically observed patterns to principles and parameters become even more intricate, and interesting (see also Nicolis 2005 on the correlation between the availability of a null expletive and the lack of that-trace effects).

As pointed out at the outset, the null subject parameter is historically the main representative of the class of spell-out parameters, but the class goes well beyond this case. Other spell-out parameters may involve null pronouns in other positions: null objects (Rizzi 1986 and much subsequent work) and other kinds of complements, and null topics and other unpronounced left peripheral elements. All these cases have in common the fact that the relevant head licenses a null specifier. Another class of spell-out parameters may involve the very fact that heads may often be null: languages may permit null copulas, null determiners, null topic and focus markers (Kayne 2015 makes the radical proposal that all functional heads are null; if that approach is correct, the whole spell-out parametrization would involve the pronunciation of specifiers). If ellipsis phenomena are triggered by special functional heads, as in Merchant (2001), and there is variation across languages, we have here another class of spell-out parameters expressible in the format advocated here.

4. Word order properties.

A very salient form of cross-linguistic variation involves word order: how does one express the parametrization distinguishing VO and OV, prepositional and postpositional languages, etc.? The traditional approach involved a head - complement ordering parameter expressed in X-bar theory, an approach also consistent with a merge-based structure building device. But
how can headedness be expressed in a way consistent with the Borer-Chomsky conjecture, and with our assumptions on the locus and format of parameters? If we think of lexical categories such as V, N, A, etc., as resulting from the merger of unlabeled roots and functional elements like v, n, a, (Marantz 2013, Borer 2005), and providing labels to the complex head root+v, root+n, etc., we may think that ordering statements are encoded in the lexical representation of such functional elements, so that v encodes the property “precedes the complement” in English and “follows the complement” in Japanese, etc.. This approach has been criticized in much recent work. In particular, it has been observed that there are asymmetries between VO and OV languages which would not be expected under a fully symmetric approach based on the headedness parameter (Kayne 1994, Cinque 2013).

An alternative developed in Kayne (1994) and much related work inspired by his antisymmetric approach proposes that all word order variation is derived by movement: hence, if VO is the universal basic word order, OV can be derived by moving the object to a higher specifier position. Along these lines, all word order variation would be caused, in our terminology, by movement parameters modifying a universal underlying order, where both underlying and surface orders are determined by Kayne’s Linear Correspondence Axiom, determining linear order on the basis of hierarchical c-command relations. Again, the relevant parametric property (here a movement property) would be encoded in the categorizing functional heads v, n, etc., and in other heads of the functional system.

A partially different proposal, developed, e.g., in Berwick and Chomsky (2011), is to assume that core syntactic representations are purely hierarchical and unordered, a consequence of the assumption that merge is a set-theoretic operation not encoding order. If this is so, linearization is an extra-syntactic process, taking place when syntactic representations are transferred to the sound system. In this view, the relevant parametric properties would involve, in our terminology, spell-out parameters, which could also be associated to the functional entries of v, n, etc..

It should be noticed that, whatever option is ultimately adopted to express word order parametrization, it should be able to capture certain cross linguistic generalizations, including those emerging from much comparative work generated by Greenberg (1964), to the extent to which they are grammatical in nature (see Newmeyer 2005, Holmberg & Roberts 2005, 2013 for different points of view). Good candidates for a grammatical analysis are Biberauer et al.’s (2008) Final Over Final Constraint, capturing the fact that a V-final VP can be embedded under a Aux initial projection (yielding the order S Aux O V), but there are no convincing cases of a V initial VP embedded under an Aux final functional projection (S V O Aux). Similarly, impossible orderings excluded by Greenberg’s Universal 20 can be ruled out by the derivational mechanism postulated by Cinque (2005) (for an extension to the more complex ordering patterns at the clausal, see Cinque, this volume).

5. Parametrisations in structural maps: uniqueness or reiteration of topic structures.

Over the last two decades, much attention has been devoted to the study of the functional structures associated to lexical projections and defining the global configurations of sentences and phrases. Cartographic research has focused on the fine properties of such functional structures, bringing to light their complexity and richness (Cinque & Rizzi 2010, Rizzi & Cinque 2015, Shlonsky 2010). The comparative analysis has underscored a robust cross-
linguistic uniformity in functional sequences, and also the necessity of postulating certain parameters of variation. Cinque (1999) showed that many cases of language-particular orderings of adverbials not conforming to his general hierarchy are amenable to movements of verbal chunks in the clausal spine: so, language particular reorderings are a function of movement parameters. In this section I will discuss some elements of parametrization emerging from cartographic research on the left periphery of the clause, with special reference to topic structures.

The left periphery of the clause is the privileged place for the expression of scope-discourse properties: the scope of operators of various kinds, the discourse-related configurations of topicality and focus. The criterial approach to scope-discourse semantics, a cornerstone of the cartographic research on the left periphery (Rizzi 1997), assumes that the left periphery is populated by a sequence of “criterial” functional head such as Top and Foc; criterial heads attract to the left periphery phrases endowed with matching features, and guide the interface systems to interpret the configuration in the appropriate way (e.g., as Topic – Comment), and to assign the special intonational properties that go with these configurations. E.g., in a Clitic left dislocation construction in Italian (Cinque 1990), the initial constituent is interpreted as the topic, and the rest of the clause as the comment expressed about the topic:

(32) Il tuo libro Top Gianni lo darà a Maria domani
    Your book Gianni will give it to Maria tomorrow’
    TOPIC COMMENT

Under the criterial approach, a functional Top head attracts a phrase with matching topic features, and guides interface systems to assign the topic – comment interpretation, and the special prosodic properties associated with this structure (Bocci 2013). Straightforward evidence for the view that scope-discourse properties are structurally encoded in this way (what is sometimes called the “syntacticization of scope-discourse semantics”) is offered by the languages in which the criterial heads are actually pronounced. For instance, the West African language Gungbe, which has been studied in great detail in a cartographic perspective in Aboh (2004) marks both topics and foci with overt particles (yà and wè respectively) occurring in the left periphery:

(33) … dò Kòfì yà gânkpá mè wè kpònòn lé sù - i dó
    ‘…that Kofi Top PRISON IN Foc policemen Pl shut him there’ (Gungbe: Aboh 2004)

Under familiar uniformity guidelines, it has been assumed that all languages involve a system of left peripheral criterial heads for the expression of scope-discourse property, except that heads of this system may be pronounced or not in different languages, a case of low level spell-out parametrisation which we mentioned in the previous section.

There are also less trivial cross-linguistic differences, though. In Italian, and in other Romance languages, topics can be indefinitely iterated. For instance all three arguments in a ditransitive sentence can be topics in the Clitic Left Dislocation construction (both the direct and indirect object are resumed by a clitic, while the subject is resumed by the familiar null pronominal subject):

(34) Gianni, il tuo libro, a Maria, glielo darà domani
‘Gianni, your book, to Maria, (he) to him it will give tomorrow’

This is not a universal property of topic configurations: many languages specify a single topic position in the left periphery. For instance in Gungbe, the topic position cannot be reiterated, as in (35):

(35) *… dò Kòfì yà gàŋkpá mè yà kpònôn lé sú - ̀ i dó
‘…that Kofi Top in prison Top policemen Pl shut him there’ (Gungbe: Aboh 2004)

Whereas the corresponding Italian structure is fully acceptable:

(36) Kofi, in prigione, i poliziotti ce lo sbatteranno senz’altro
‘Kofi, in prison, the policemen will shut him there for sure’

Uniformity guidelines would lead us to assume a minimal difference between the two cases, i.e., the same general system of criteria, uniform in syntax and at the interfaces, except that some property may be recursive in one kind of language but not in the other.

What property? A priori, the parametrization could be expressed in at least two ways. It could be that the Top head is recursive in the left periphery of Italian, so that each topic is attracted by its own Top head, hence we have

(37)

\[
\text{Gianni} \quad \text{Top} \quad \text{il tuo libro} \quad \text{Top} \quad \text{a Maria} \quad \text{Top} \quad \ldots \ldots \ldots \ldots
\]

Or, in an approach to structure building permitting multiple specifiers, it could be that both languages admit a single Top head, but the Italian varieties admits recursive creation of multiple specifiers, while the Gungbe variety does not. In this case, the structure of (34) would look like the following:

(38)

\[
\text{Gianni} \quad \text{il tuo libro} \quad \text{a Maria} \quad \text{Top} \quad \ldots \ldots \ldots \ldots
\]
Italian does not seem to offer direct evidence to choose between (37) and (38) because, among other things, Italian differs from Gungbe in not pronouncing the Top head, hence it does not offer immediate evidence for or against the possibility of reiterating the head.

But relevant evidence comes from an enlarged comparative perspective. The West African language Abidji involves an overt Top head, like Gungbe, and allows a multiplicity of topics, like Italian. As Hager – Mboa (2014) points out, in case of topic reiteration each topic can be accompanied by its own Top particle ëké:

(39) a. kòfì ë pipjé ëkókò ë
Kofi MA peel.RES banana Def.
« Kofi peeled the banana. »

b. ëkókò ëi ëké kòfì ë pipjé nìi
banana Def. Top° Kofi MA peel.RES p.a.i
« The banana, Kofi peeled it. »

c. kòfì ëké ëkókò ë ëké tì ë pipjé nìj
Kofi Top° bananaj Def. Top° tì MA peel.RES p.a.j
« Kofi, the banana, he peeled it. »

d. ëkókò ë ëké kòfì ëké tì ë pipjé nìi
bananaj Def. Top° Kofij Top° tj MA peel.RES p.a.i
« The banana, Kofi, he peeled it. »

So, (39)c-d provide straightforward evidence for the conclusion that Top recursion of the kind illustrated in (37) is a possible UG option. What about (38)? At the moment we do not have direct comparative evidence in favor or against this possibility. Nevertheless, restrictive approaches to phrase structure, such as Kayne (1994) rule out such representations with multiple specifiers associated to a single head. So, we may tentatively conclude, on grounds of restrictiveness, that multiple topics always occur in a structural representation like (37) in Italian, a representation directly supported by comparative evidence.

The assumption that multiple occurrences of topics always involve topic recursion of the kind illustrated in (37) may well be a simplification. Frascarelli & Hinterhoelzl (2007), Bianchi & Frascarelli (2010) have shown that different occurrences of topics may bear detectably different interpretive properties (they identify, in particular, aboutness shift, contrastive and familiarity topics). So it could be that no recursion is involved in (34), but simply the co-occurrence of different types of topics.

Nevertheless, even if the possibility of distinct topic position is taken into account, the postulation of some recursion mechanism is inevitable. The generalization about languages of the Italian type is that topics are not restricted to two or three per clause: any number of topicalizable arguments and adjuncts can be topicalized. Here is an example with six topic-like elements

(40) Gianni, a Piero, la sua residenza, a Parigi, da Londra, l’anno prossimo, gliela farà spostare senz’altro
‘Gianni, to Piero, his residence, to Paris, from London, next year, he will make him move it for sure’

As topics can be added indefinitely, at least one kind of topic position must be recursive.

6. Conclusion.

Ever since the Borer-Chomsky conjecture, a widely held assumption is that the locus in which syntactic parameters are expressed in the grammatical architecture is the functional lexicon. As for the format of parameters in a minimalist grammar, under the assumption that syntactic operations are feature triggered, we may think that parameters simply consist in associating to functional heads morphosyntactic features which trigger operations of structure building (external merge), movement (search and internal merge) and spell-out. So, the size of the set of parameters may well be large: cartographic studies suggest that the functional lexicon is very rich, hence if the parametrization is associated to this component the system will specify many parameters. Nevertheless, what is critical for the restrictiveness of the system is the format, not the locus of parameters: and the format is extremely restrictive, as the elementary operations that can be performed in minimalist syntax are very few. The space of variation that the system leaves open for the language learner is thus severely limited, so that no learnability problem arises even if parameters are very numerous: the situation is therefore very different from an approach to variation based on a potentially unlimited set of language-particular rules, as in pre-parametric models. The numerosity of parameters has an impact on the likelihood that a single parameter may govern a complex cross-linguistic pattern. There are just too many parameters to expect that any of them will be able to operate in isolation: each parameter will inevitably interact in complex ways with many other parameters to determine cross-linguistic variation. The case of the Null Subject Parameter straightforwardly illustrates this point. Nevertheless, it is not true that a system with many parameters will have a shallow deductive structure. Quite the contrary is true: the system of natural language syntax is so tight that complex deductive paths connect individual parameters and observable empirical patterns, involving rich deductive interactions with other parameters and UG principles. A central goal of comparative syntax is to reconstruct and elucidate such complex paths.

In the last decades, cartographic studies have offered a powerful tool to comparative syntax, providing structural maps which can be immediately compared across languages, thus offering direct evidence for invariance and variation. Cartographic results thus have nourished the parametric approach to syntactic variation by offering novel types of evidence for the formal tools that a theory of variation must assume.

Acknowledgement.

This research was supported by the ERC Advanced Grant n. 340297 “SynCart”.

18
References.


