On the elements of syntactic variation*

LUIGI RIZZI
University of Siena
rizzil@unisi.it

In the first part of this paper, I would like to go back to the origins of the notion of parameters, briefly present the first steps of parametric theory some 30 years ago, and discuss an important conceptual change that took place very early on in the development of parametric syntax: the identification of the locus of parameters not in the structure of principles, but in the functional lexicon. I will then try to spell out a conception of the format of parameters as elementary instructions for syntactic actions; I will argue that this view is fully consistent with basic tenets of minimalist models, and in fact crucially capitalizes on the restrictive character of minimalist syntax. In the second part, I would like to turn to language acquisition and syntactic variation: how early is abstract grammatical knowledge of cross-linguistically variable properties acquired by the child? I will focus on the acquisition of a fundamental word order property, the VO or OV order, and discuss experiment evidence bearing on this issue. The available experimental evidence supports the view that the child possesses abstract knowledge of word order properties from very early on; this is more naturally compatible with the assumption that a dedicated language faculty, in the form of a parametrized system, constrains linguistic variation.

0. Introduction.

This talk is divided into two independent parts. The first part is conceptual and historical. The hope is that a historical perspective may somehow help us evaluate where we stand now in the theoretical study of syntactic variation, and possibly suggest where we want to go. I would like to go back to the origins of the notion of parameters, briefly present the first steps of parametric theory some 30 years ago, and discuss an important conceptual change that took place very early on in the development of parametric syntax. I will then try to spell out a view on parametrisation which, in my opinion, is implicitly assumed in much current work in comparative syntax. It seems to me that this view is fully consistent with basic tenets of minimalist models, and in fact crucially capitalizes on the restrictive character of minimalist syntax.

In the second part, I would like to turn to language acquisition and syntactic variation. The context is a broader issue now on focus in the debate in the cognitive neurosciences: is linguistic variation constrained by the human language faculty? Or, is it just a particular case of cultural variation that is solely constrained by general intelligence and multipurpose problem solving skills? One central empirical contribution that acquisition studies can give here

* This paper was presented at the Workshop *Linguistic Variation in the Minimalist Framework.* January 14-15, 2010 UAB. Casa de Convalescència. Hospital de Sant Pau. Barcelona. A revised version will be published in the proceedings.

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

is to provide evidence on the time course of the acquisition process: how early is abstract grammatical knowledge of cross-linguistically variable properties acquired by the child? I will focus on the acquisition of a fundamental word order property, the VO or OV order, and discuss an experiment bearing on this issue. I will suggest that the available experimental evidence is more naturally compatible with the assumption that a dedicated language faculty, in the form of a parametrized system, constrains linguistic variation.

Part I: Parameters as the elements of syntactic variation.

1. Origins.

So, let's start from the beginning and in the beginning (at least for the linguists of my generation, formed in the early 1970's) there was the Extended Standard Theory with the structure expressed in (1).

(1) EST Models (i.e., Chomsky 1973, 1975, 1977) :

- Particular grammar: system of language-specific, construction-specific rules
- Universal grammar: grammatical metatheory specifying a broad format for rules and some general principles on rule application (A over A, Island constraints, etc.)
- Acquisition: rule induction

The theory was really centered on the notion of particular grammars as systems of rules specific to a particular language and construction-specific: there were phrase structure rules for the NP and the VP, and construction-related transformational rules like passive, question formation etc. which constituted the grammar of English, and similar rule systems were postulated for Italian, Chinese, etc.. Universal Grammar was thought of as a kind of a 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 a few others. 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 his particular grammar were on the basis of the format provided by Universal Grammar and of the empirical evidence presented to him.

There were some obvious problems with this way of looking at things. One critical problem was that a system based on language specific rules wasn't 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 one couldn't really put the finger on the primitive properties that remained uniform or that varied, a rather frustrating state of affairs. Comparative syntax wasn't really feasible on that basis because the fundamental invariant and variable elements could not be isolated in a sufficiently transparent manner.

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

Another serious problem was that this system couldn't really address the problem of acquisition because there weren't precise enough ideas about how rule induction could work. So that technically the analyses at that time did not attain the level of adequacy that Chomsky (1964) had called "explanatory adequacy", the level which is reached when an analysis somehow comes with an account of how the relevant properties could be acquired by the language learner. It was clear at that time that one could hope to successfully address this problem only by radically restricting the options offered by Universal Grammar, i.e. by making the rule systems among which the child was assumed to choose more and more impoverished.

Things changed around the second half of the 1970s. Recently, I came across the passage in (3) in Chomsky's *Conditions on rules of grammar*; as far as I can tell, this is 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.

Chomsky somehow considered the abstract 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 with the case of extraction from indirect questions, the selective violation of wh-islands. It turned out that in some languages it is possible to extract an element from an indirect question as in (5) in Italian, while in other languages this option doesn't exist.

(3) Ecco un incarico [$_{CP}$ che [$_{IP}$ non so proprio [**$_{CP}$**a chi [$_{IP}$ potremmo affidare ___]]]]
'Here is a task that I really don't know to whom we could entrust'

In my original analysis (Rizzi 1978) I compared Italian and English but in English things turned out to be quite complex (see, e.g., Grimshaw 1986), so for the purposes of this illustration I have used a German example here in (4), a language which manifests a robust impossibility of extracting something from an indirect question. If you take the word by word equivalent of (3) in German, modulo word order, etc. you get an ungrammatical sentence:

(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'

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

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) was 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 between the two languages: i.e., by taking CP as the clausal bounding node for Italian, and IP for German (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 CP's would be crossed in the double wh island (5), thus accounting for the deviance of the structure in Italian[69].

This turned out to be a rather peripheral parameter in retrospect (in fact, one that is not easily amenable to the general format to be discussed later on), but the important point is that it was soon realized that one could entertain the ambitious program of dealing with the whole cross-linguistic variation in terms of parametric choices; the postulation of a set of language specific rules could 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 wouldn't 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 analyzed in a comparative perspective in terms of the parametric model. Moreover, the theory of principles and parameters provided a promising model of the

[69] Certain varieties of German are very restrictive on wh extraction, banning extraction even from embedded declaratives and permitting the expression of question 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 parametrisation of bounding nodes, e.g. in Freidin (1988), Other varieties, spoken e.g. in Southern Germany and Austria, permit extraction from declaratives and still manifest a robust wh island banning examples like (4).

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

acquisition of syntax *qua* parameter setting, a much more appealing conception than one based on an obscure notion of rule induction. Reaching the level of explanatory adequacy thus became a feasible enterprise, even if by no means an obvious one.

2. From "parameters expressed in principles" to "parameters in the functional lexicon"

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 clearly there were ideas on the locus. The assumption was made that, as the first parameter looked like an option specified on a principle, perhaps that was the locus of parameters in general. so that UG principles would somehow express parameters.

(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 idea about the possible size of the set of parameters: as there were few principles, one would expect few parameters. It may be that we have a dozen principles so we may have a dozen or maybe two dozen parameters, something like that. It also gave rise to the so-called switchboard model. I think, the image is originally due to James Higginbotham and essentially is that the child is confronted with a little switchboard with principles specifying parameters, and then 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.

As I said, 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 parameterization. In (7), I give a little list of parameters that were identified around the late 1970's 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 1983)
- the head precedes/follows the complement,
- V moves to I (Emonds 1978, Pollock 1989),
- V moves to C (V-2 Germanic: Den Besten 1977/1982)
- N incorporates into V (Baker 1988)
- Nominative is assigned under agreement (SVO) or under government (VSO) (Koopman & Sportiche 1991)
- there are long-distance anaphors (English vs Icelandic, etc.: Manzini & Wexler 1986)

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

- wh-movement is overt or covert… (English vs Chinese, etc.: Huang 1982)
- the language is non-configurational (K. Hale)

So, we have bounding nodes, we have the licensing of empty elements, we have certain selectional properties of special verb classes like *believe*-type epistemic verbs, movement properties of various sorts 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 there was something wrong with this view. There were a number of problems. One was the unnatural character of the list in (7). And then, there were the other problems indicated in (8).

(8) Some problems with the model of "parameters as specifications on principles", in addition to the arbitrary-looking character of the list of the first parameters:

a. Some principles didn't appear to be parametrized at all.

b. Some parameters appeared to be directly keyed to the presence of particular lexical items.

c. Other global parameters like non-configurationality turned out to be advantageously reanalysable as conglomerates of more elementary parameters:

One problem was that some principles didn't seem to require or allow a parametrisation. Take for instance hierarchical properties of X-bar theory - always the same across languages, presumably: structures are built by heads projecting and taking complements and specifiers - ; the Theta Criterion (e.g., no known language seems to admit structure 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 (Principle C doesn't seem to be parameterized at all: 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).

The second problem, perhaps more important, was 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

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

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 the language but it's a property of a particular lexical item.

Moreover, it turned out that certain global parameters like non-configurationality could be advantageously reanalyzed as conglomerates of smaller parameters. It is clear that null subject languages are more non-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 also are 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). Languages where it is possible to split the DP have more ordering options than languages which do not permit DP split (Boskovic 2009), etc.. So, one really gets 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 much 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.

A significant shift, directly suggested by problem (8)b, occurred at this point: the view that parameters are expressed on principles was abandoned in favour of the hypothesis that the locus of parameters is the functional lexicon.

(9) Parameters are not specified directly in UG principles, but rather are to be conceived of as featural specifications in the (functional) lexicon.

This shift is clearly expressed in the following quote taken from Hagit Borer.

(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 (1983: 29)

I will basically adhere to this conception in the rest of this talk, 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 so 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

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

complements of epistemic verbs like *believe* in English and Romance as in (11):

(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)
c. *Credo [(di) [Gianni sapere la risposta]]
d. Credo [di [PRO sapere la risposta]]
e. * 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, as in (11). In Romance, one gets 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, 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[70].

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 chamber* vs *enter the room*), etc., and all the item-particular cases in which categorial selection seems to depart from the Canonical Structural Realisation of semantic selection (Grimshaw 1979, Pesetsky 1983) in language-specific, and item-specific ways. A possible solution here may be provided if "selected" prepositions are reanalyzed as being part of the functional structure associated to the verb, as in Kayne (2000).

In the reminder 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 (see also Kayne, 2005, Cinque & Rizzi 2010 for discussion).

[70] 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 parametrisation is limited to the functional lexicon, as Frédérique Berthelot suggests. Thinking of the decomposition of verbs into v and V components, the class could be characterized by the presence of a specially "flavoured" v, say v_{epist}, which could be responsible for the c-selectional properties of the complex v_{epist}+root. 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+root.

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

3. The theory of parameters in current models.

What does a parameter look like in current syntactic theorizing? Building on some suggestions in Rizzi (2009), I would like to propose the following informal characterisation:

(13) A parameter is an **instruction for 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 by acting as a head to be merged with other syntactic entities, 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:

(14) H has F {yes, no}

Where H is an item of the functional lexicon entering syntax as a head, and F is a relevant feature. In order to make the system properly restrictive, we must now specify the range of F more precisely. Features are the expression of properties of various kinds: of sounds, of meanings, 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 and small subset of morphosyntactic features which have the property of triggering the basic syntactic actions. If we assume a highly restrictive theory of possible syntactic actions such as minimalism, parametric features will be restricted to the features triggering the basic 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:
   - H c-selects XP (where XP departs from the canonical structural realisation of the s-selected entity).

2. Move parameters:
   - H attracts X[+F]
   - H attracts XP[+F]

3. Spell-out parameters:
   - H is null
   - H licenses a null Spec.

Merge parameters may primarily express cases in which the head's categorial selection does not immediately reflect principles of canonical structural

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

realisation: e.g., the cases of "truncated" clausal selection of English epistemic verbs referred to in the previous section. Other cases may be the cross-linguistically variable orders 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 (Cinque 1999, Belletti 2001); single or recursive Topic in the left periphery, presence or absence of Focus projections in the CP and/or in the vP periphery specialized to new information or contrast (Rizzi 1997, 2004, Belletti 2004, 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 (Pollock 1989, Cinque 1999, Holmberg & Platzack 1995, Roberts & Holmberg 2005), 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 (which could lead to distinct possible parametrisations on the search operation, and on internal merge). The head-complement ordering parameter may be seen as a Merge parameter in more traditional approaches, or as a Move parameters in antisymmetric approaches (Kayne 1994); or else as a spell-out parameter if ordering is a property confined to externalisation (Chomsky & Berwick 2009). Whatever the exact nature of this property, the crucial feature should be specified on the functional categories assigning the categorical status to lexical roots (i.e., v, n, a, p, etc.), with the greenbergian tendency to uniformity (Greenberg 1963) expressed grammatically (Biberauer, Holmberg, Roberts 2008) or explained 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 head has the capacity to license a null pronominal subject and/or a null pronominal object in some languages (Rizzi 1982, 1986), etc.

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 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 realisation of its s-selectional properties), attract another head or a specifier, be spelled out or not, and govern the spell-out properties of its Spec.

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

4. On the numerosity of parameters.

The view that the functional lexicon is the locus of parameters affects the expectations on the number of parameters:

(16) The size of the set of parameters is not determined by the number of principles, but by the size of the (functional) lexicon.

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 set of principles. Moreover, if cartographic studies are on the right track (Cinque ed. 2002, Belletti ed. 2004, Rizzi ed. 2004, Cinque & 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 suggests 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?

It seems to me that this argument overlooks the important distinction between the locus and the format of parameters. Under the current conception, the loci of parameters are quite numerous and diverse, a direct function of the size of the (functional) lexicon, as we have seen; but the format is extremely restrictive, as determined by the restrictiveness of minimalist syntax. The syntactic actions that a featural specification can triggered 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.

So, assimilating the two kinds of models overlooks what seems to me to be genuine and substantial progress in the identification of the basic ingredients of linguistic computations over more than thirty years of syntactic research.

Of course, the choice of a particular terminology is largely an arbitrary decision, inasmuch as it does not affect the nature of the devices referred to. So, one may decide not to use the term "parameters" for the devices referred to by (13), (14), (15) and call them "language-particular rules", "item-specific rules", or the like without changing in any way the structure of the approach. Nevertheless,

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

different terminological choices may not be completely innocent and neutral, as they may have very different connotations, somehow linked to the intellectual history of the field. In our particular case, using the term "rule" in connection to such specification as (13), (14), (15) seems to me to be highly misleading. First, because the term "rule" evokes the very complex phrase structure and transformational rules of pre-parametric models (the precise counterparts of the millenary tradition of language description through construction-based rules), which have nothing to do with the highly restrictive conception expressed in (13)-(14)-(15): an instruction to trigger one of very few syntactic actions made available by a very austere theory of syntax. Second, because the shift of the locus for parameters from UG principles to the functional lexicon took place already around the mid-1980's, 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 2000, and many contributions in Cinque & Kayne 2005). In the absence of a clear conceptual or formal shift, I think it would be very misleading to introduce a new terminology, or go back to a highly connotated old terminology.

The assumption we are now making on the numerosity of parameters has other consequences. If parameters are so numerous, it's unlikely that a single parameter will fully control complex sets of properties, simply because there will be too many interactions. Many parameters imply many intricate interactions. It is a little bit like the fact that it is unlikely that a single gene will control very complex aspects of the structure of the body, say the whole shape and internal structure of an organ, simply because there are too many genes and there would be too many interactions.

If complex arrays of properties cannot be made to follow in any simple manner from single parametric values, this does not mean that parametric values only have "local" consequences, and a parameter-based system will have no deductive depth: quite the contrary is true. Parametric choices will typically have consequences well beyond the simple property they express because the system has a very tight deductive structure, and a small difference in one point will typically have systemic repercussions.

Pursuing our analogy with DNA, that is essentially like the repercussions that a single gene would have. A small mutation on a single regulatory gene could have radical and diversified consequences for the structure of the body, affecting different organs and cognitive capacities, precisely because of the tight interconnections in the system. The action of a single gene is local – it may be limited to turning on or off another gene, but this may have cascading effects with pervasive consequences. The same seems to be basically true with parameters. Their action is very local, for instance the licensing of a null argument by a functional head. But then some of these 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. For instance the licensing

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

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 (two properties that may well be closely related: Rizzi 2006b, Rizzi & Shlonsky 2007), properties of the Case-agreement system, etc. So we observe that null subject languages typically have null expletives (for the formal satisfaction of the EPP property), and typically don't manifest that-trace effects, as the availability of the null expletive offers a device to escape that-trace[71]. Should we then expect a perfect correlation between null subjects and the violability of that-trace? In this particular case, the connection may hold with remarkable systematicity (Nicolis 2005), but we can't 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, Sobin 2002, in Norwegian, Taraldsen 1986, etc.. Along similar lines, Romance null subject languages permit focalisation of the subject in a low, clause final position (Belletti 2004), a property which 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 determine a complex cluster of properties. More precisely, that would be possible "all other things being equal", i.e. in the abstract 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; reasonable approximations may be found through the microcomparison of historically very close grammatical systems, i.e. in the cases provided by dialectological studies (the closest approximation to a controlled experiment in comparative syntax, as Richard Kayne pointed out: see Kayne 2000 for discussion).

[71] In the terms of Rizzi and Shlonsky, op. cit., 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).

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

Part II: On the acquisition of variable properties.

5. On the early acquisition of certain abstract grammatical properties.

In this second part I would like to address the question of how the child acquires cross-linguistically variable properties. I would like to try to phrase this question in the context of a broader issue: is linguistic variability constrained by a dedicated "language faculty"? or is it just a particular case of cultural, historically determined variation, with no domain-specific constraints? The former position has been a central assumption throughout the history of generative grammar; the theory of parameters is a particularly precise and empirically successful version of this position. The latter position is assumed by a large spectrum of viewpoints, recently taken up by Evans and Levinson (2009): there are certain contents to express, and there are indefinitely many imaginable ways to do so, hence an indefinite cross-linguistic variability is to be expected; linguistic communities make particular choices and language learners figure out what these choices are, as in any other case of cultural acquisition, through their general intelligence and domain-general problem-solving capacities.

The study of acquisition may provide relevant evidence on this broad divide. The timing of the acquisition process matters. So, we should really pay attention to how fast or how slow acquisition is. The first approach, let's call it the "language faculty" approach in a synthetic form, naturally leads to the expectation of a fast acquisition of the cross-linguistically variable properties. The problem that the language learner is confronted with is very well defined and narrowly circumscribed and the learner is guided by task-specific cognitive resources which allow her to quickly converge to the correct parametric values. The second approach, which I will refer to as the "constructivist" 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. For instance, one would expect a certain correspondence between the acquisition of variable properties of language and the acquisition of culturally-driven technical abilities of various sorts: reading, writing, drawing, and so on.

Let us address the question of the time course of acquisition in connection with the acquisition of a fundamental cross linguistically variable property: word order, and in particular the VO or 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 older learning English will typically say "eat cake", and the two year older 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

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

of an ordering parameter (a Merge parameter in traditional approaches; a Move parameter in antisymmetric approaches; an externalisation parameter in the analysis of Chomsky & Berwick 2009: the choice is not critical for the point relevant here).

On the other hand, the 'constructivist' hypothesis, proposed in this particular context by Michael Tomasello and his associates in a number of papers (Tomasello 2000, 2003, Achtar and Tomasello 1997, etc.), assumes that the child initially memorizes fragments she hears, and stores in memory individual items with the associated syntactic environments. So, there is no generalization initially, there is only memorization of fragments, individual items with the syntactic structures in which they are found. Initially the child retrieves such item-based knowledge and reproduces it in her early productions; only much later on does the child generalize such item-based knowledge to abstract and general grammatical statements like "my language is OV (or VO)".

The two approaches lead to clearly different expectations on 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 new structures because the relevant knowledge is abstract from early on; on the contrary, the 'item-based' 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. To be fair, neither approach makes 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, while constructivist approaches seem to assume that abstract knowledge will arise through analogical generalisation only well after the third birthday (consider, e.g., the fact that children in the younger group tested by Matthews et al., on which see below, are around age 2.9). So, even though the two approaches do not generate very sharp predictions about the exact time course of the acquisition of abstract knowledge, they clearly lead to quite distinct expectation about the earlier or later character of such acquisition.

6. An experiment.

Let me now present an experiment which bears directly on this issue: Franck, Millotte, Posada &. Rizzi (2011). In order to test the abstract grammatical knowledge of 19 months 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, Fisher and Eisengart

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

2006, and Hirtsh-Pasek & Golinkoff 1997 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 in Abbot-Smith, Lieven & Tomasello 2001; Akhtar & Tomasello 1997; Matthews, Lieven, Theakston & Tomasello 2005; 2007, and adapted to comprehension).

3. Jabberwocky verbs are used, morphophonologically possible items which are not listed in the French lexicon.

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 ass 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 ass 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 "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.

The item based approach, on the other hand, does not predict any preference in either case. As in this approach the infant does not have any abstract grammatical scheme to build on, but only item-based 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

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

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 is predicted in either case.

The experimental evidence clearly is in line with the expectations of the "abstract grammar" approach: it is reported in Franck et al (op.cit.) that infants look at the transitive video significantly more time 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". 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, Nespor, Guasti, & Van Ooyen (2003), Gervain, Nespor , Mazuka, Horie, Mehler (2008).

There is an apparent contradiction between these results and the conclusion reached by Matthews et al. (2005, 2007) on the basis of production experiments. They elicited sentences with jabberwocky 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 schema. 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, Millotte & Lassotta (2010) have redone the Matthews et al. experiments by introducing certain modifications in the methodology, in particular by improving the communicative situation; they found that younger children at 2.11 were not distinguishable from older children at 3.11 in the weird word order task, 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 jabberwocky 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, producing sentences like *La vache, elle a dasé le chien* 'the cow, it has dased the dog' with pronominalisation, dislocation, the introduction of compound tenses etc.. In contrast, children in both groups failed to produce any compound tense, special inflectional properties on the verb, 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 they

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

just repeated the input string in the (rare) occasions in which they reproduced the ungrammatical NP NP V order. These authors therefore conclude that also the younger group 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 about to enter the two-word stage). Franck et al. (2010) 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, as the language faculty approach would lead us to expect.

7. 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 microcomparative analysis of very close varieties, and approximating the ideal of two systems differing for a single parametric value). 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 brief second part, I have broadened the perspective to the general issue of the nature of cross-linguistic variation, and the plausibility of assuming dedicated cognitive resources constraining linguistic variability. Crucial evidence on this issue can come from the study of the timing and characteristics of language acquisition. I have focus 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 seems hard to reconcile with a view 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 I discussed is more readily consistent with a view such as the parametric approach, 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.

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

References.

Aboh, E. (2004) *The Morphosyntax of Complement-Head Sequences*, Oxford University Press, New York.

Akhtar, N. and Tomasello, M. (1997). Young children's productivity with word order and verb morphology. *Developmental Psychology, 33*, 952–965.

Baker, M. (1988). *Incorporation: A Theory of Grammatical Function Changing*. Chicago, Chicago University Press.

Belletti, A. (2001) "Agreement projections", in M. Baltin, C. Collins, eds., *The Handbook of Contemporary Syntactic Theory*, Bleckwell, Oxford – Malden, pp. 483-510.

Belletti, A. (2004) "Aspects of the Low IP Area", in Rizzi, ed. (2004).

Belletti, A. (Ed.), (2004) *Structures and Beyond: The Cartography of Syntactic Structures, Vol. 3*, Oxford University Press, New York.

Belletti, A. (2009) *Structures and Strategies*, Routledge, London.

Berwick, R. & N. Chomsky (2008) "The biolinguistics program: The current state of its development", ms., MIT.

Biberauer, T., A. Holmberg & I. Roberts (2008) "Disharmonic Word Orders and the Final-over-Final Constraint (FOFC)", in A Bisetto & F Barbieri (eds) Proceedings of the XXXIII Incontro di Grammatica Generativa (available online)

Borer, H. (1983) *Parametric Syntax*, Dordrecht: Foris.

Boskovic, Z. (2009) "On NP's and Clauses", talk presented at the International Conference on Sentence Types: Ten Years After, University of Frankfurt, Germany.

Chomsky, N. (1964) "Current Issues in Linguistic Theory". In Fodor, J. & J. Katz, eds. *The Structure of Language*, Prentice Hall.

Chomsky, N. (1973) "Conditions on Transformations". In S. Anderson & P. Kiparsky, eds., *A Festschrift for Morris Halle*, Holt, Rinehart and Winston, New York, 232-286

Chomsky, N. (1976) 'Conditions on Rules of Grammar', *Linguistic Analysis* 2: 303-351.

Chomsky, N. (1977) *Essays on Form and Interpretation*. New York, Amsterdam, London, North Holland.

Chomsky, N. (1981) *Lectures on Government and Binding*. Dordrecht, Foris Publications.

Christophe, A., Nespor, M., Guasti, M.T., & Van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: the case of the head-direction parameter. *Developmental Science, 6*(2), 211-220.

Cinque, G. (1999) *Adverbs and Inflectional Heads*, Oxford University Press, Oxford-New York.

Cinque, G., ed. (2002) The Structure of CP and DP: The Cartography of Syntactic Structures, Vol. 1., Oxford University Press, Oxford – New York.

Cinque, G. & R. Kayne, eds. (2005) *The Oxford Handbook of ComparativeSyntax*, Oxford University Press, Oxford – New York.

Cinque, G. & L. Rizzi (2010) "The Cartography of Syntactic Structures", in *Handbook of Syntactic Theories*, Oxford University Press.

Cruschina, S. (2006) "Informational focus in Sicilian and the left periphery". In M.Frascarelli, ed., *Phases of Interpretation*. 363-385. Berlin: Mouton de Gruyter.

den Besten, H. (1977/1983) 'On the Interaction of Root Transformations and Lexical Deletive Rules', in W. Abraham, (ed.) *On the Formal Syntax of Westgermania*, Amsterdam: Benjamins, 47-131.

Emonds, J. (1978) "The verbal complex V'-V in French". *Linguistic Inquiry*, 9:151-175.

Evans, N. and S.C. Levinson (2009) "The myth of language universals: Language diversity and its importance for cognitive science", in *Behavioral and Brain Sciences*, 32, 429–492.

Felser, C. (2004) Wh-copying, phases and successive cyclicity. *Lingua* 114 (2004) 543-574.

Franck, J., S. Millotte, A. Posada, L. Rizzi (2011) "Abstract knowledge of word order by 19 months: An eye-tracking study", in publication in *Applied Psycholinguistics*.

Franck, J., Millotte, S., & Lassotta, R. (2010) "Early word order representations: Novel arguments against old contradictions", in publication in *Language Acquisition*.

Freidin, R. (1988) "Comments on Lightfoot (1988)", *Brain and Behavioral Sciences*, 12.

Gertner, Y., Fisher, C., & Eisengart, J. (2003). Learning words and rules. *Psychological Science, 17*(8), 684-691.

Gervain, J., M. Nespor , R. Mazuka, R. Horie, J.Mehler (2008) Bootstrapping word order in prelexical infants: A Japanese–Italian cross-linguistic study, Cognitive Psychology, 2008.

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

Greenberg, J. (1963) “Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements”, in J. Greenberg, ed., *Universals of Language*, MIT Press, Cambridge, Mass, 73-113.
Grimshaw, J. (1979) ‘Complement Selection and the Lexicon’, *Linguistic Inquiry* 10: 270-326.
Grimshaw, J. (1986) “Subjacency and the S/S’ Parameter”. *Linguistic Inquiry*, 17:364-369.

Huang, J. (1982) *Logical Relations in Chinese and the Theory of Grammar*. PhD dissertation, MIT.
Hirsh-Pasek, K.R., & Golinkoff, R.M. (1996). *The origins of grammar.* Cambridge, MIT Press.
Holmberg, A. & C. Platzack (1995) *The Role of Inflection in Scandinavian Syntax*, Oxford University Press, Oxford – New York.
Kayne, R. (1983) *Connectedness and Binary Branching*, Foris Publications, Dordrecht.
Kayne, R. (1994) *The Antisymmetry of Syntax*, MIT Press, Cambridge, Mass.
Kayne, R. (2000) *Parameters and Universals.* Oxford and New York, Oxford University Press.
Kayne, R. (2005) *Movement and Silence*. New York: Oxford University Press.
Koopman, H. & D. Sportiche (1991) “The Position of Subjects”. *Lingua*, 85, 211-258
Manzini, M. R. & K. Wexler (1987) “Parameters, binding, and learning theory”, in *Linguistic Inquiry*, 18, 413-444.
Matthews, D., Lieven, E., Theakston, A. & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development, 20*, 121-136.
Matthews, D., Lieven, E., Theakston, A. & Tomasello, M. (2007) French children's use and correction of weird word orders: A constructivist account. *Journal of Child Language, 32(2)*, 381-409.
Moscati, E. (2007) *The Scope of Negation*, Doctoral Dissertation, University of Siena.
Newmeyer, F. (2004) “Against a parameter-setting approach to language variation”, in *Language Variation Yearbook*, vol. 4, Benjamins, Amsterdam, 181-234.
Newmeyer, F. (2005) *Possible and Probable Languages: a Linguistic Perspective on Linguistic Typology*, Oxford University Press, Oxford – New York.
Pesetsky, D. (1982) *Paths and Categories*, PhD dissertation, MIT, Cambridge, Massachusetts.
Pollock, J.-Y. (1989) “Verb Movement, Universal Grammar, and the Structure of IP”. *Linguistic Inquiry*, 20, 365-424.
Postal, P. (1974) *On Raising*, MIT Press, Cambridge, Mass.
Rizzi, L. (1978) “Violations of the Wh Island Constraint in Italian and the Subjacency Condition”. *Montreal Working Papers in Linguistics*, 11.
Rizzi, L. (1982) *Issues in Italian Syntax*. Dordrecht, Foris Publications.
Rizzi, L. (1986) “Null Objects in Italian and the Theory of pro", *Linguistic Inquiry*, 17.3, 1986, pp. 501-557.
Rizzi, L. (1997) “The Fine Structure of the Left Periphery”, in L. Haegeman (Ed.), *Elements of Grammar*, Kluwer, Dordrecht.
Rizzi, L. (2000) *Comparative Syntax and Language Acquisition*, Routledge, London.
Rizzi, L. (Ed.), (2004b) *The Structure of CP and IP – The Cartography of Syntactic Structures vol 3*, OUP.
Rizzi, L. (2004b) Locality and Left Periphery. In A.Belletti, ed., *Structures and Beyond. The Cartography of Syntactic Structures, vol.3*, New York: Oxford University Press
Rizzi, L. (2006a) Grammatically-Based Target-Inconsistencies in Child Language. In Deen, K.U., J. Nomura, B. Schulz and B.D. Schwartz (eds), *The Proceedings of the Inaugural Conference on Generative Approaches to Language Acquisition–North America*, Honolulu, HI. University of Connecticut Occasional Papers in Linguistics, 4, 2006, pp 19-49.
Rizzi, L. (2006b) . “On the Form of Chains: Criterial Positions and ECP Effects”, in L. Cheng, N. Corver, eds, *On Wh Movement*, MIT Press, Cambridge, Mass.
Rizzi, L. (2009) ”Language Invariance and Variation”, in M. Piattelli-Palmarini, J. Uriagereka, P. Salaburu *Of Minds and Language- The Basque Country Encounter with Noam Chomsky*, Oxford University Press, Oxford-New York, 2009, pp. 211-220.
Rizzi, L. and U. Shlonsky. (2007) “Strategies of Subject Extraction”, in H.-M.Gärtner and U. Sauerland (eds). *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics*. 115-16. Berlin: Mouton de Gruyter, 538-553.
Roberts, I. & H. Holmberg (2005) “On the role of parameters in Universal Grammar: A reply to Newmeyer”, in Broekhuis et al. (2005) *Organizing Grammar*, Mouton de Gruyter, Berlin – New York

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)

Sobin, N. (2002) "The Comp-trace effect, the adverb effect and minimal CP", in *Journal of Linguistics* 38, 527-560.
Sportiche, D. (1981) "Bounding Nodes in French". *The Linguistic Review*, 1:219-246.
Taraldsen, Knut T. (1978) "On the NIC, Vacuous Application and the That-Trace Filter" Bloomington, Indiana University Linguistics Club.
Taraldsen, Knut T. (1986) "Som and the Binding Theory" in L. Hellan and K. K. Christensen, eds., *Topics in Scandinavian Syntax*, Reidel, Dordrecht.
Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*, 209-253.
Tomasello, M. (2003). Constructing a language: A usage-based theory of language acquisition. Cambridge, M.A.: Harvard University Press.
Tomasello, M., Akhtar, N., Dodson, K., & Rekau, L. (1997). Differential productivity in young children's use of nouns and verbs. *Journal of Child Language, 24*, 373-387.
Zanuttini, R. (1997) *Negation and Clausal Structure: A Comparative Study of Romance Languages*, New York and Oxford: Oxford University Press.

BDD-A22709 © 2010 Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Provided by Diacronia.ro for IP 216.73.216.103 (2026-01-20 20:36:01 UTC)