Rizzi, Luigi. 2014. On the Study of Language as a Cognitive Capacity: Results and Perspectives. In On Peripheries, Anna Cardinaletti, Guglielmo Cinque, Yoshio Endo, 61 – 99. Tokyo: Hituzi Syobo

On the Study of Language as a Cognitive Capacity: Results and Perspectives *

Luigi Rizzi

1. Introduction: Some Fundamental Questions

In this paper, I would like to focus on some major directions of research on language within the cognitive sciences. I will illustrate the roots of these directions in work of the last half century or so, I will focus on the current understanding of the issues, and speculate on their possible future projections. The first issue is the conception of language as a cognitive capacity, and the idea of modelling that capacity as the possession of a computational system, a system that tacitly computes linguistic expressions. This is the approach which gave linguistics the status of one of the main founding disciplines of the cognitive sciences in the fifties and sixties, and which determined the profound influence of the study of language on the cognitive sciences. The second direction concerns the fundamental issue of the universality and variability of language, and the idea of addressing this problem through a system of principles and parameters, which separates the general properties of human language from the specificities of individual languages; this approach paved the way to modern comparative linguistics, and at the same time offered a simple and attractive model of language acquisition, which provided solid foundations for developmental psycholinguistics. The third direction is the reflection on language design and the attempt to identify the ultimate, irreducible ingredients of the language faculty, and separate them from more general properties of the human cognitive capacities, an attempt which has recently received special emphasis within the Minimalist Program.

These three issues have always been present in the development of linguistic theory over the last half century, but the emphasis has somewhat shifted from one to the other in different periods. Very roughly these periods correspond to the phases, respectively, of classical generative grammar (standard and extended standard theory), of the principles and parameters approach, and of the minimalist program. It is common, in the presentations of this development, to underscore the hiatuses between these phases, but I personally find the continuity more significant than the discontinuity. The same fundamental questions have been present throughout; even if the emphasis shifted and, of course, important new ideas were introduced and other ideas were abandoned, the development has largely been cumulative, incremental, in that a new phase has taken into account the results of the previous phase and taken them as a point of departure.

Let us start from the "cognitive revolution" and the modern study of language as a "mirror of the mind", according to a classical definition. This approach, also sometimes called the biolinguistic approach or generative grammar, started developing in the 1950's under the impulse of Noam Chomsky's revolutionary ideas. It revolves around some fundamental questions:

(1) What is knowledge of language?
(2) How is it acquired in childhood?
(3) How is it put to use in production and comprehension?
(4) How is it physically expressed in the human brain?
(5) How did the capacity for language develop in the species?

The first three questions about knowledge, acquisition and use defined the program of generative grammar from the outset. The question of knowledge of language had a certain historical priority in the development of this approach, for good reasons: in order to ask meaningful questions about acquisition and use it is necessary to have a precise idea of the internal structure of the system which is acquired and used. For instance, one cannot seriously study language acquisition without a detailed characterisation of the end point of the acquisition process, the stable cognitive state of adult speakers for language: this is essential to fully appreciate the nature and dimension of the cognitive task that every child is confronted with, acquiring an extremely rich and complex system in the first years of life by simple exposure, without explicit teaching. The same considerations hold for the study of language use, the classical disciplinary domain of

experimental psycholinguistics. The fourth questions is the traditional domain of neuropsychology and neurolinguistics: it is observed that patients with lesions in a specific brain area manifest specific kinds of language deficits, and the inference is made that the neural circuitry in that area is crucially involved in that particular kind of linguistic computation. In the last twenty years or so, much progress has been made on this question thanks to the new techniques of brain imaging; the rapidity of technological development, as well as the accumulation of results, suggests that crucial advances are to be expected on this question. The progressive integration of the study of the brain hardware with the functional modelling of the cognitive capacities was also underscored by the new name of "cognitive neurosciences" that the federating framework of these studies has assumed in the last two decades. As for the fifth question, it is hard at the moment to translate it into a series of precise, testable hypotheses, so it still belongs to the domain of speculation more than to proper scientific inquiry. And yet, the interest of investigating the biological roots of a capacity which makes us humans very different from any other species is so great that we can't resist reflecting and speculating on it, even if the arguments remain at the borders of scientific inquiry.

2. Knowledge of Language as the Possession of a Computational Capacity

The first issue can be rephrased as follows: what does it mean "to know a language"? What kind of system of knowledge does the native speaker of a natural language tacitly possess? One salient property of the human knowledge of language is the so-called creativity manifested in normal language use: we can constantly produce and understand new linguistic objects, sentences that we had never encountered in our previous linguistic experience, and our linguistic capacities give us the possibility of expressing an indefinitely large number of messages. No other species possesses a communication system with such characteristics. The importance of this creative aspect is not a new observation: it was clear, in essence, as early as in the seventeenth century. René Descartes pointed out that the capacity to organize words into an unlimited number of appropriate sentences distinguishes the dumbest man from the most intelligent ape, and from the most sophisticated machine.

(6) "Car on peut bien concevoir qu'une machine soit tellement faite qu'elle

profère des paroles… mais non pas qu'elle les arrange diversement, pour répondre au sens de tout ce qui se dira en sa présence, ainsi que les hommes les plus hébétés peuvent faire » (p. 86)
« Because we can well conceive that an engine be built in such a way that it may produce words, but not that it may arrange words in different ways, in order to respond to the meaning of all that is said in its presence, as the dumbest men can do"

« …il n'est pas croyable qu'un singe ou un perroquet, qui serait des plus parfaits de son espèce, n'égalât en cela un enfant des plus stupides, … si leur âme n'était d'une nature du tout différente de la notre… » (p. 87)
« … it is not believable that an ape or a parrot, which would be among the most perfect of their species, didn't equal in this the dumbest child, … if their soul was not of an utterly different nature from ours…"
(Descartes 1637)

Incidentally, Descartes' "test" based on language to distinguish man and machine remains essentially valid today, in spite of the extraordinary progress of computer science: no computer can approximate the linguistic capacities of a 5 years old child, a fact which is revealing of how complex language is, and of how good our species is at language use.

What is the secret of this creative aspect? At the roots of it, there must be a combinatorial capacity. More or less at the same time of Descartes' writing, Galileo Galilei praised the invention of alphabetic writing systems as the most magnificent achievement of mankind: a device which made it possible, in Galileo's words, to communicate the most hidden thoughts across time and space by the extremely simple device of combining twenty little characters on a paper:

(7) "…ma sopra tutte le invenzioni stupende, quale eminenza di mente fu quella di colui che s'immaginò di trovar modo di comunicare i suoi più reconditi pensieri a qualsivoglia altra persona, benché distante per lunghissimo intervallo di luogo e di tempo? Parlare con quelli che son nell'Indie, parlare con quelli che non sono ancora nati, né saranno se non di qua a mille o dieci mila anni? E con qual facilità? Con i vari accozzamenti di venti caratteruzzi sopra una carta."
(Galilei 1630)

"… but above all the marvellous inventions, what eminence of mind was the one of the man who imagined to find a way to communicate his most hidden thoughts to any other person, though at the greatest distance in place and time? Speak with those who are in the Indies, speak with those whose who aren't born, and will not be for one thousand or ten thousand years? And with what ease? With the different combinations of twenty little characters on a paper"

Galileo speaks of "a marvellous invention", and writing certainly is an invention, but one that is based on an empirical discovery on the way in which natural language works: language has an inherently combinatorial character, which gives rise to infinite combinations of finite means, and writing is a cultural innovation which exploits and reproduces this natural property.

What was missing till the middle of the twentieth century was a technical device to precisely address this creative aspect. Structural linguistics in the first half of the century conceived of language (saussurean "langue") as a systematic inventory of linguistic signs, each of which consisting of a sound-meaning pairing, basically a theory of the lexicon. But an inventory is limited by definition, so, this approach was intrinsically unable to address the fundamental question of creativity: the infinite possible combinations of linguistic signs was relegated by Saussure to "parole", the actualisation of the system of "langue" in individual linguistic acts. Saussure was probably dissatisfied with this conclusion, as certain oscillations in the *Cour de linguistique générale* — and in the notes taken by Saussure's students — suggest (see Saussure 1916/1985). Natural language syntax is clearly regular, a rule-governed process, but linguistics at the beginning of the twentieth century did not dispose of the formal device to express this regularity.

Chomsky showed that the core notions of the theory of recursive functions, developed in the study of the foundations of mathematics, could be adapted to language. A recursive procedure is one that can indefinitely reapply to its own output, giving rise to a hierarchical structure. The following examples illustrate some simple cases of recursion in natural language, in which a phrase of a given type can be embedded into a phrase of the same type (phrases are delimited by brackets), and this can go on indefinitely:

(8) The preface [of the first book [on the discovery [of the …]]]

(9) I believe [that people wonder [whether Mary thinks [that someone said ...]]]

(10) I met [the boy [who bought [the book [which pleased [the critics [who wrote [the review ....]]]]]]]

Notice that it is not the mere iterability of the procedure which makes it recursive, but its capacity to create a hierarchical structure, as happens at different levels of organisation of linguistic structures. So, for instance the iteration of the motor program activated in walking, one step after the other, can go on indefinitely, but it doesn't give rise to any hierarchical structure, while the stringing together of words in a sentence does, determining bracketed representations like (8)-(10) or, more perspicuously, tree-like representations like the ones we will consider in a moment.

Phrasing things more generally, we can say that to know a language means to possess certain inventories of elements (lists of distinctive features, phonemes, syllable structures, morphemes, words, idiomatic expressions), somehow stored in memory, and certain computational procedures putting together the elements of the inventories to form entities of a higher order; crucially, when this cascade of levels leaves the lexicon and enters syntax, and we start putting words together, the computational procedures become recursive, and give rise to higher order entities, phrases and sentences, which are indefinitely extendable.

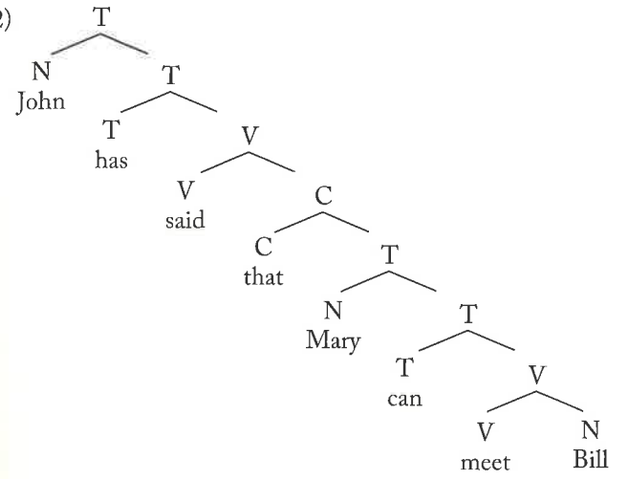
Various recursive techniques have been adopted in the different linguistic models which have been proposed from the 1950's on, ever since *Syntactic Structures*. Jumping across fifty years of syntactic research, the ultimate distilled format of syntactic recursion is the operation Merge, the fundamental structure-building procedure assumed by the Minimalist Program (Chomsky 1995 and much related work), which takes two elements A and B, either directly from the lexicon or as expressions already formed by previous applications of Merge, to form a composed expression C:

(11)

In order for Merge to apply there must be some selectional relation between A and B, so if A is a transitive verb and B is a noun, merge can apply forming a

transitive verb phrase (say, *meet Bill*). In the obtained configuration, the selector has the label which projects, and gives a name to the whole structure; so, in the case of a verb-object construction the obtained C constituent would be a verbal projection. Successive applications of merge can give rise to complex structures like the following, expressed here in terms of a current approach known as Bare Phrase Structure, a component of Minimalism:

(12)



The verb *meet* is merged with the noun *Bill* giving rise the verbal constituent *meet Bill*, which is then merged with the tense bearing element, here the modal *can*. After merger of the subject, the sentence thus created *Mary can meet Bill* is merged with the complementizer *that*, an element which transforms a sentence into a complement, available to be selected by, and merged with, a higher verb, *say* in our case, and so on.

This conception of the structure of linguistic expressions raises issues concerning the whole set of questions (1)-(5), which are actively investigated now and will certainly continue to be pursued in the future. Jumping now from more technical issues of linguistic analysis to the broader issues (4) and (5), let us venture on some more general questions raised by the approach to syntactic structures I just presented. The recursive property is critical in characterizing the two systems of discrete infinities which humans are good at: language and the number system. No other animal species has a communication system which permits the mastery of an infinity of messages, and no other animal species can

count (many species are good at roughly estimating quantities, but that's a very different kind of cognitive ability from precise counting: Dehaene 1997).

(13)a. Is there an identifiable neural substrate which implements the recursive property for language and number in the brain? If so, at what granularity of analysis does it emerge? How does it relate to classical areas dedicated to linguistic (Grodzinsky 2000) and numerical (Dehaene 1997) capacities?

b. Do these mechanisms relate in a non-trivial way to the mechanisms responsible for other kinds of hierarchical structures in other cognitive domains, vision, motor control, the theory of mind and the other cognitive capacities ruling social interactions?

c. How did the mastery of recursion for communication and other human cognitive systems evolve in the natural history of the species? (Hauser, Chomsky, Fitch 2002).

We do not know what the answer may be to these questions, but it is not too far fetched to imagine that partial answers are within reach, through the conjoined efforts of formal modelling of cognitive capacities, the study of pathology, and the brain imaging techniques; these questions will probably keep the scientific community of cognitive neuroscientists busy in the years to come.

Immediately connected to question (4) is the phylogenetic question (5). Hauser, Chomsky and Fitch (2002) speculate that the availability of recursion for the communication system (and perhaps derivatively for the number system) may have been a sudden and recent event in evolutionary history, perhaps the major consequence of a minor reorganization of the brain which had been growing steadily for millions of years, and that this single evolutionary event may be at the root of the acquisition of what paleoanthropologist Ian Tattersal calls "the human capacity", the collection of cognitive capacities which makes our species so different from the others. These speculations currently are at the extreme periphery of scientific understanding; nevertheless, here too it is imaginable that a joint interdisciplinary effort of formal modelling, the brain sciences, and comparative ethology may permit serious advances on evolutionary aspects of language and other human cognitive capacities, which have so far resisted serious scientific inquiry.

3. Language Invariance and Variation

In his later work *The Descent of Man*, Charles Darwin addressed the question of what now is called the naturalistic study of the human cognitive capacities. The following famous passage underscores well the special position of language, in between nature and culture:

(14) [Language] certainly is not a true instinct, as every language has to be learnt. It differs, however, widely from all ordinary arts, for man has an instinctive tendency to speak, as we see in the babble of our young children, while no child has an instinctive tendency to brew, bake, or write… (Darwin 1871)

So, there is a critical instinctive basis in language, more so than in other domains of human knowledge and activity, but the biological endowment for language leaves room for experience-based learning, hence for culture.

A few decades after Darwin, Ferdinand de Saussure gave a name to Darwin's instinctive tendency to acquire a language, and called it *faculté du langage*, the language faculty. Much of modern linguistics is an attempt to understand the nature and properties of the language faculty.

The cognitive study of language must come to terms with the phenomenon of language variation. In this sense, the study of the language faculty is distinct from the study of other basic cognitive faculties, the faculty of vision, for instance, which does not give rise to significant variation in the acquired systems. On the one hand, we want to express the fact that the language faculty is uniform across the species, much as the faculty of vision: any normal child acquires whatever linguistic system he is exposed to in early childhood, a cognitive achievement which is completely unattainable for differently constituted organisms. On the other hand, individual languages vary, in ways that go beyond the simple fact of mutual incomprehensibility, which could simply arise from very superficial distortions in the sound systems leaving the structural constitution entirely identical: clearly, language diversity involves deeper structural properties then just a superficial variation in the sound output. Mark Baker has underscored the point in an interesting way (Baker 2001), pointing out that the military history of the 20th century shows us that natural languages used as codes are much harder to

break by the enemy's intelligence than artificially devised codes, as the famous case of the "Navajo code talkers" indicates (orders expressed in Navajo could not be decoded by the Japanese intelligence during World War II). So, the issue of properly characterizing invariance and variation, an issue common to any science having to do with a rich phenomenology, arises in a specially interesting way in the case of language.

Generative grammar traditionally addressed the issue through the two concepts of Particular Grammars and Universal Grammar: Particular Grammars express the specific properties of individual languages, Universal Grammar (UG) expresses linguistic universals. In cognitive terms, UG expresses the properties of the initial cognitive state for language, the properties given to us by our biological endowment; Particular Grammars express the attained systems at the end of the acquisition process, the speaker's intuitive knowledge of his native language.

The assumed relation between particular grammars and universal grammar has evolved considerably in the last half century, not gradually, but through a sudden conceptual and formal change which took place in the late seventies with the introduction of the Principles and Parameters framework.

Till the early and middle seventies, particular grammars were conceived of as language specific, construction specific rule systems, essentially following a millenary tradition of description of languages as particular systems of grammatical rules. From this perspective, English has a particular phrase structure rule for the noun phrase, another one for the verb phrase, a particular transformational rule for passive, one for the interrogative, one for the relative construction, etc. Hungarian has other language specific rules for the same constructions, Chinese has other such rules, etc. Universal Grammar was conceived of as a grammatical meta theory: it specified the format for rules, and some general principles on rule application, for instance, principles responsible for the Island Constraints which had been discovered in the late sixties (Ross 1967, 1986, etc.). Language acquisition was seen, from this viewpoint, as a process of grammatical induction: the child, equipped with the principles of Universal Grammar, tries to figure out the rule system characterizing the language he is exposed to within the space of possibilities defined by Universal Grammar. We can characterize this approach as the (extended) standard approach, encompassing different kinds of models from the mid sixties (the so-called Standard Theory) to the mid seventies (different varieties of the Extended Standard Theory).

(15) The traditional approach: (Extended) Standard Theory, etc.
- Particular Grammar: a system of language-specific rules.
- Universal Grammar: grammatical metatheory, expressing the format for rules, and certain universal conditions on rule application.
- Language acquisition: rule induction.

The (extended) standard approach made much progress possible through careful analyses of the syntax of English and a few fragments of other languages, but it had serious drawbacks. A major difficulty had to do with language acquisition. There were no precise ideas on how grammatical induction could really work: attempts to elaborate an evaluation metric making the choice possible between alternative analyses remained largely programmatic, and the question of how exactly the child formulates grammatical hypotheses, tests and compares them against the empirical evidence offered to him, remained essentially unanswered.

Things changed around the late seventies in a radical and rather abrupt manner, even though the shift had been prepared by about a decade of work on the simplification of rule systems, along the lines of Chomsky (1973). The first technical step was to introduce the concept of parameter. Certain general UG principles, the so-called Island Constraints of Ross 1967, seemed to give slightly different results in different languages. For instance, extracting a relative pronoun from an indirect question was quite natural in some languages and impossible in others, like German (English was originally discussed in this connection, but the facts appear to be sharper in German):

(16) Ecco un incarico [$_{CP1}$ che [$_{TP1}$ non so proprio [**CP2** a chi [$_{TP2}$ potremmo affidare ___]]
'Here is a task that I really don't know to whom we could entrust'

(17) *Das ist eine Aufgabe, [$_{CP1}$ die [**TP1** ich wirklich nicht weiss [$_{CP2}$ wem [**TP2** wir ___ anvertrauen koennten]]]].
'Here is a task that I really don't know to whom we could entrust'

Clearly, one cannot simply say that Italian is not sensitive to this kind of island, the Wh Island. If we somewhat complicate the structure, and put an indirect question inside another indirect question, the extraction of a relative pronoun from the double wh Island becomes impossible in Italian as well:

(18) *Ecco un incarico [$_{CP1}$ che [$_{TP1}$ non so proprio [**CP2** a chi [$_{TP2}$ si domandino [**CP3** se [$_{TP3}$ potremmo affidare ___]]]]]]
'Here is a task that I really don't know to whom they wonder if we could entrust'

It was then proposed that these variations could be accounted for by assuming that the relevant UG principle contained a parameter, a choice point, and depending on how the parameter was fixed one would get slightly different effects (Rizzi 1978, 1982, ch. 2). There is a uniform locality principle, Subjacency, stating that movement may cross at most one barrier, the boundary of a certain kind of node. But languages can vary in a limited way in the choice of barriers. Italian selects CP as the clausal barrier, while other languages, like German, may select TP; then movement from an indirect question is possible in Italian as only one barrier (CP2, underscored in (16)) is crossed, but banned in German (two barriers, TP1 and TP2, are crossed in (17)). Extraction from a double Wh Island is excluded in Italian as well, because two barriers are crossed (CP2 and CP1 in (18)).

(19)a Subjacency: movement can cross at most one barrier (Chomsky 1973)
b Barriers: Italian {CP, ...} (Rizzi 1978)
German {TP, ... }

In retrospect, this locality effect turned out to be a rather marginal kind of parameter (not even naturally expressible in the format to be discussed later); but the important thing was that it was quickly realized that through a systematic use of this notion it had become possible to account for all the variation in syntactic computations, and do away completely with the notion of language specific rule system (Chomsky 1981).

Universal Grammar could be rethought of as a system of principles and parameters. The former expressed linguistic invariance, the latter were conceived of as binary choice points expressing the range of possible variation. Particular Grammars could then be seen as UG with parameters fixed in certain ways. The core of language acquisition could be construed as the process of setting parameters on the basis of experience.

(20) The Principles and Parameters approach (Chomsky 1981):

- Universal Grammar: a system of principles and parameters, binary choice points.
- Particular Grammar: UG with parameters set in a specific way.
- Language acquisition: parameter setting.

have the verb following the direct object in the fundamental word order :

Consider a simple example. Some languages have the verb preceding the direct objects, and others.

(21)	[V O]		[O V]
French:	Jean [aime Marie]	Latin :	Tullius [Flaviam amat]
	'Jean loves Marie'		'Tullius Flavia loves'
Chicewa:	Njuchi [zi-na-wa-lum-a alenje]	Japanese:	John-ga [Mary-o butta]
	'bees bit hunters'		'John Mary hit'
Edo:	Ozo [mien Adesuwa]	Navajo:	Ashkii [at'ééd yiyiiltsa]
	'Ozo found Adesuwa'		'Boy girl saw'

Suppose that we express this variation in terms of a binary parameter, which generalizes the option to all kinds of head-complement relation, so, not just verb – object, but also preposition – complement of the preposition, noun – complement of the noun., and so on.

(22) Head precedes/follows complement

On the basis of this parameter, when a new word is merged with a phrase, the new word (the head of the construction) is attached to the left or to the right of the phrase (the complement).

This elementary difference has pervasive consequences on the phrase structure of a language. So, the clausal structure of a coherently head-initial language like English looks like the right-branching tree (24), while the structure of a coherently head final language like Japanese looks like the left- branching tree (25).

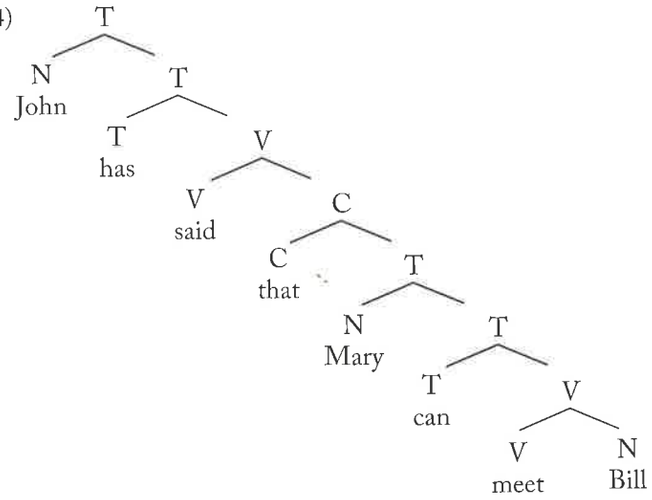
The elementary bifurcation expressed by (22), affecting the pervasive recursive operation merge determines major global differences in the derived structures. The system is not so strict as to forbid "incoherent" languages. In fact, some languages have verbs following objects but prepositions preceding their

complements (e.g., German, as is clear in non V-2 environments). About 90% of the word's languages are coherent, as emerges from much typological work stemming from the pioneering research conducted by Joseph Greenberg at Stanford over 40 years ago (see Greenberg 1963 and, for a later assessment, Dryer 1992). We may assess this state of affairs as follows: parameter (22) is fixed for individual heads, so that the possibility of incoherent or "disharmonic" fixations is contemplated by the system. A "harmonic" fixation (in which each ordering value is analogized, e.g., on the very salient VO or OV order) is preferred, as it involves a system with fewer independent specifications, but it is not compulsory, so that we do not have here a strict universal, but only a statistical tendency (but see Biberauer et al. 2010 for the observation that possible kinds of disharmonic orderings must respect certain constraints).

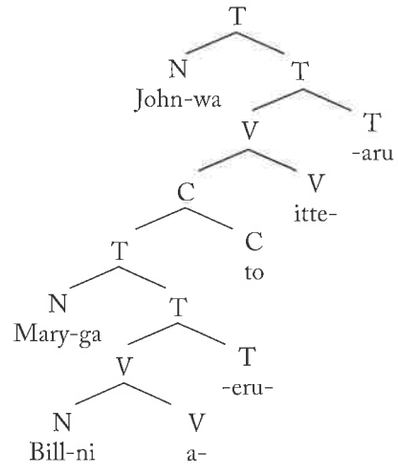
Restricting now our attention to consistently head initial languages and consistently head final languages, we observe that they typically involve tree structures whose shapes differ dramatically: compare sentences (23) in English and Japanese and the corresponding trees:

(23)a John has said that Mary can meet Bill
b John-wa [Mary-ga Bill-ni a-eru-to] itte-aru.
'John-Top [Mary-Nom Bill-Dat meet-can-COMP] said-has

(24)



(25)



A particular grammar could be seen as a specific realization of UG under a particular set of parametric values. The theory of parameters introduced analytical tools well adapted to express language invariance and variation, to factor out linguistic universals and capture patterns of variation. So, it is not surprising that a new comparative syntax flourished, with comparative studies quickly extending from Romance and Germanic to Semitic, East Asian and African languages, Austronesian, Amerindian and Australian Aboriginal languages.

Some major parameters affect the word order of the language; some operate on Merge directly, as the head complement parameter which determines the order of two merged elements. Other word order parameters affect the result of movement. One typical case is the Verb Second parameter, the property which characterizes all modern Germanic languages, except English, which demands the placement of the inflected verb in second position in main clauses, after the first constituent, which can be the subject, as in (26a), an adverbial, as in (26b), etc.; in a non V-2 language, like Modern English, the inflected verb does not have to be in second position, it can be in third position, as in (26c), in fourth position, as in (26d), and so on.

(26)a. Die Kinder **sahen** den Film
'the children saw the film'
b. Gestern **sahen** die Kinder den Film
'yesterday saw the children the film'

c. Yesterday, the children saw the film
d. Yesterday, the student carefully read the paper

In addition to parameters on Merge and Move, other much studied parameters affect the spell-out operation leading to the pronunciation of the structure, and determine the possibility of not pronouncing certain elements or structural positions. Of relevance here is the so called Null Subject Parameter, differentiating languages like Italian and Spanish, for instance, from languages like English and French. The study of the Null Subject Parameter, as of the early Eighties, led to the attempt of going beyond the mere separation of invariant and variable properties, by trying to identify, within variation, certain recurrent patterns.

A useful heuristic strategy consisted in comparing languages relatively close in grammatical structure and sufficiently well-studied to minimise the possible interference of spurious factors. So, if we compare Romance and (non-Verb Second) Germanic, two basic patterns seem to emerge, illustrated, respectively, by Italian and French-English:

(27) a. Io parlo italiano — a'. ___ parlo italiano
b. Je parle l'italien — b'. * ___ parle l'italien
c. I speak Italian — c'. * ___ speak Italian
(28) a. Gianni ha telefonato — a'. ___ ha telefonato Gianni
b. Jean a téléphoné — b'. * ___ a téléphoné Jean
c. John telephoned — c'. * ___ telephoned John
(29) a. Chi credi che ___ verrà?
b. *Qui crois-tu que ___ viendra?
c. *Who do you think that ___ will come ?
(30) It.: parl-o, parl-i, parl-a, parl-iamo, parl-ate, parl-ano
Fr.: /parl/, /parl-ō/, /parl-é/
Eng.: speak, speak-s

The possibility of a null pronominal subject in languages like Italian, as in (27a') may seem a trivial option, but in fact it appears to be connected, directly or indirectly, to other rather different properties of the language. So, it was observed that Null Subject Languages like Italian also typically allow the free positioning of the subject in clause-final position, as in (28a'), and allow free subject extraction, as in (29a'), options which are typically excluded in non Null Subject Languages like English and French.

The key property appeared to be that languages of the Italian kind have a much richer specification of agreement on the verbal morphology. So, it was proposed that a verbal inflection reaching a certain threshold of richness can licence a null pronominal subject (referred to as *pro*), determining, directly or through the interaction with other principles and parameters, the range of formal and interpretive properties illustrated by (27)-(30) (Rizzi 1982, ch 4, 1986).

(31) The verbal inflection licenses *pro* ? {yes, no}

As the comparative work on this topic developed over the last twenty five years or so it became progressively clear that the pattern is not controlled in its entirety by a single parameter, and partially independent parametric properties must be postulated. In particular the possibility of subject inversion in (28), turned out to be a particular device to focalise the subject (Belletti 2001), typically found in Null Subject Languages, but not necessarily associated with the other properties: for instance, many Bantu languages display the fundamental Null Subject properties but do not have subject inversion (of the Italian kind). Other parametric distinctions are necessary to separate languages permitting a full or reduced range of interpretations for the null pronominal (Rizzi 1986).

That complex cross-linguistic patterns are not fully controlled by individual parameters, but by the interaction of several parameters, is made immediately plausible if we reflect on the natural format for parameters, and on reasonable estimates of the number of parameters that must be postulated. Where and how are parameters expressed in UG?, and how many of them should be assumed?

The initial view, largely influenced by the Subjacency parameter around 1980, was that parameters were sorts of switches expressed on UG principles. But this "switches on principles" model was quickly abandoned, in part because some principles didn't look parametrized at all, and, more importantly, because some parametric values seemed to be intimately connected to particular elements of the functional, or grammatical lexicon, as was pointed out by Borer 1983: the Null Subject parameter is a property of the functional head expressing tense and agreement, the V-2 parameter is a property of the complementizer, etc.

So, the conception of the general format for parameters that researchers started assuming, explicitly or implicitly, was something like the following:

(32) H has F
H = a functional head
F = a feature determining a merge, move or spell-out operation.

where H is a functional head, an element of the functional or grammatical lexicon (the expression of T or other inflectional properties, a complementizer, a determiner, and the like) entering syntax, and F is a feature determining one of the basic syntactic actions, operations of Merge, or Move, or Spell-out related to H. In this view, the parametrisation is confined to properties of the functional lexicon, the inventory of elements which define the syntactic configurations in which the elements of the contentive lexicon (nouns, verbs, adjectives, etc.) are inserted, and parameter fixation by the child amounts to the acquisition of the functional lexicon.

As for the headedness parameter, it can be construed as a property of the functional lexicon in the following manner. Marantz (1997) has proposed that the syntactic categories are assigned to lexical roots by merging the latter with functional categories v, n, a, turning the unspecified roots into verbs, nouns, adjectives, etc.. Then the OV or VO order can be seen as a merge parameter specified on v: in some languages v+root is merged with its complement to the right, in others to the left. Languages tend to specify this ordering parameter in a uniform manner for all functional elements, including those assigning categorial status to roots, but the system tolerates the complication of specifying the order separately for distinct functional elements, whence the possibility of "incoherent" languages, as observed by Greenberg. If one follows the alternative approach to "basic" word order advocated by Kayne (1994), according to which the complement – head order is derived via movement from an underlying head – complement order, the fundamental ordering parameter is to be stated as an instruction on move, rather than on (external) merge.

In this view of the locus and format of parameters, the order of magnitude of parameters is determined by the size of the functional lexicon, not of the principles. And as we will see in a moment, much recent work, particularly within the cartographic projects, suggests that the functional structure of clauses and phrases is quite rich, consisting of dozens of functional heads, and possibly more. In short, parametric choices are more numerous than the initial switch-board model would have led us to expect. And if parametric choices are more numerous, their interactions will be more intricate, and it will be unlikely that

complex cross-linguistic patterns may be entirely governed by single parameters.

The observed numerosity of parameters has sometimes been interpreted as a kind of *reductio ad absurdum* of the parametric approach: if variation results from hundreds, perhaps thousands of choice points, how is this different from the view that each language is characterized by language specific rules, generating the language's constructions? If the language learner must make hundreds or thousands of choices, how is this different from saying that s/he induces language-specific rules or constructions? Aren't we essentially back to preparametric models, with individual grammars consisting of rules specific to the particular language?

In fact, considering the sheer numerosity of choices is highly misleading, as it obscures the fundamental restrictiveness of the system. To clarify things, it is useful to distinguish between *locus* and *format* of parameters: the *format* of parameters is extremely impoverished and restrictive, basically as in (32): a parameter is an instruction for a syntactic action encoded in a functional head; as the system permits very few syntactic actions, the format only allows for very few cases: can the element be merged with another element? To the left or to the right? Does it attract another element to its specifier? Is it overt or null? Does it license a null specifier or complement? On the other hand, the *locus* of parameter expression is the functional lexicon: as the size if the functional lexicon is large, it gives rise to numerous choice points, but this should not obscure the high restrictiveness of the format. This is very different from going back to highly permissive phrase structure and transformational rule systems of the Standard and Extended Standard Theory, permitting innumerable patterns not attested in any natural language and raising insurmountable problems for acquisition models.

The observation of the numerosity of parameters also motivated a particular methodology for doing comparative syntax, the so called microparametric approach, strongly advocated by Richard Kayne in particular, consisting in taking very closely related languages or dialects as the starting points of comparison, in order to control the explosion of interactions. The comparison of closely related varieties is the best approximation to a controlled experiment in comparative syntax, as Kayne (2005) notes. So it is not surprising that dialectological studies have acquired much general importance in this microparametric perspective.

If the importance of starting small is clear, one should also not lose sight of the grand picture. General typological patterns do exist, and if it is unlikely that they may be fully controlled by individual parameters, we want to understand what parameters are involved, and how they interact. In the domain of null

subjects, the heuristic validity of the method of comparing closer systems, and then gradually extending the comparison, was shown by the fact that similar clusterings have been observed well beyond Romance and Germanic; for instance, Kenstowicz (1989) showed that different Arabic dialects neatly subdivide along the lines of the observed clustering, and different families of African languages turned out to be similarly subdivided: the Gbe languages are English like (Aboh 2004), many Bantu languages (Swahili, Lingala, etc.) are Italian like, with the systematic proviso that they disallow the VS order. Once the causal factors are properly teased apart, and the multiplicity of parametric choices affecting patterns is properly acknowledged, statements of universal validity are still derivable from the system which withstand cross-linguistic verification through the survey method (Nicolis 2005).

The parametric model offered an effective device to compare languages, by factoring out the common properties and isolating the irreducible points of variation. The introduction of these ideas had a dramatic impact on the field: all of a sudden, an appropriate technical language to do comparative generative grammar had become available, and comparative studies flourished, starting from Romance and Germanic, but quickly extending to other language families.

4. Parameter Setting, Language Acquisition, "Learning by Forgetting"

The principles and parameters approach also had a profound influence on the study of acquisition. The task of the language learner could be reconstructed as an operation of parameter setting: there is no induction of grammatical rules, because there are no language specific grammatical rules; what the language learner must do, from this perspective, is figure out, on the basis of his experience, what option his language has taken at every choice point. Setting parameters in the empirical conditions which characterise normal language acquisition is not a trivial task, as the data available to the child may remain ambiguous between different settings. Nevertheless, parameter setting is a much more operative concept than an obscure notion of grammatical induction, it is a clear concept on which formal and empirical studies of language acquisition can be solidly based. The rapid growth of comparative acquisition studies from the mid eighties on (ever since Hyams 1986: see, for instance, much work published in *Language Acquisition*, the journal which more directly reflects this trend) strongly confirmed the theoretical and operative value of the approach.

From the viewpoint of the interdisciplinary interactions, it is worth noticing that the principles and parameters approach, and the model of acquisition of syntax which it embodies, have found important analogues in the study of other language-related domains, and also, more broadly, in the study of learning mechanisms in biological systems. An important discovery of the experimental study of the initial cognitive state is that newborn babies are able to discriminate all the phonetic distinctions that adult languages use, but a few months later babies remain sensitive only to the phonetic distinctions used in the language they have come in contact with. This mechanism has also been called "learning by forgetting": learning the inventory of phonetic distinctions of one language amounts to forgetting the distinctions used by the other languages, which are all present at birth:

(33) "L'apprentissage d'une langue se traduit en fait par une perte partielle, par la sélection de certains contrastes et l'oubli d'autres, qui ne sont pas pertinents… Les tout petits bébés discriminent tous les contrastes auxquels ils sont confrontés, tandis que vers huit ou dix mois , ils montrent la même aptitude, mais seulement pour les contrastes de leur propre langue. A douze mois, ils se conduisent comme des adultes et ne sont plus sensibles aux contrastes étrangers" (Mehler and Dupoux 1990: 236).
« The learning of a language results in fact in a partial loss, in the selection of some contrasts and the forgetting of other constrasts, which are not relevant… Small babies discriminate all the contrasts they are confronted with, whereas around eight months they show the same attitude, but only for the contrasts of their own language. At twelve months, they behave like adults, and are not sensitive anymore to the foreign contrasts ».

So, at birth, babies exposed to English are able to discriminate between dental and retroflex /t/, but at 8–10 months they lose this distinction, while babies exposed to Hindi keep it, in fact for the rest of their lives, as it is used in the language they are acquiring (Werker et al., 1981). In this domain, acquisition and learning seem to involve a reduction, rather than a growth, of the initial repertoire of options, and the role of experience is not to import knowledge from the external world, but rather to select certain options which are already internal to the learning system, and inhibit and eventually discard other options.

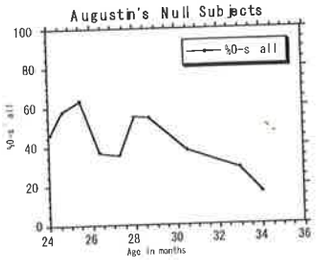
Under the parametric approach, also the acquisition of higher levels of linguistic knowledge, such as syntax, is, in essence, a selective mechanism: the child figures out, on the basis of the linguistic data he is exposed to, which direction to take at each parametric bifurcation, and discards the other possible direction, so, in a sense, he "forgets" options used by other languages. The analogy of parameter setting with learning by forgetting is further reinforced by the observation that in their early productions children systematically use certain grammatical options which are not in their target grammar, but which are found in other adult systems: the child is apparently driven by some internal necessity to explore, for some time, certain parametric options which are not supported by experience; the pressure of experience eventually prevails, and such non-target consistent options are later abandoned, or "forgotten".

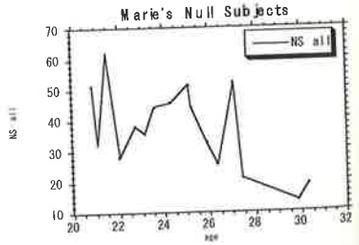
One well-studied case in point is the selective subject omission which is persistently observed, throughout the third year of life, in the productions of children acquiring non-Null Subject languages. For instance, in corpora of child English and French productions we typically find examples like the following:

(34) Child English (Brown 1973)
a. ___ was a green one (Eve, 1;10)
___ falled in the briefcase (Eve 1;10)

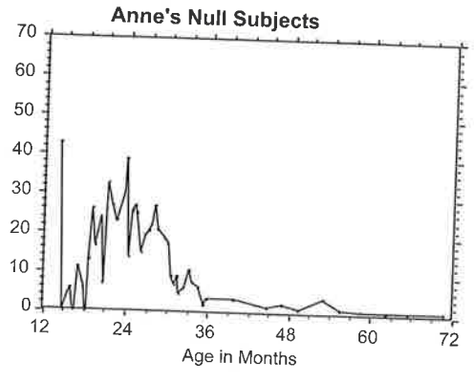
(35) Child French (Hamann, Rizzi, Frauenfelder 1996)
___ a. tout tout tout mangé (Augustin 2,0)
'___ has all all all eaten'

(36) Subject Drop in Child French (from Hamann 1997)





(37) Subject drop in Child Danish (from Hamann and Plunkett 1997)



(from Hamann and Plunkett 1997)

This kind of subject drop differs from the one found in adult Null Subject languages in that it is limited to the initial position. In non-initial position, for instance in the position following a preposed interrogative pronoun (as (38b)), or in embedded subject position (as (38d)), subjects are typically pronounced in child productions:

(38) Child English (and French, German, Dutch, Swedish,....)
a. Where dis goes?
b. *Where ___ goes?
c. ___ know what I maked
d. *___ know what ___ maked

This dropping of the initial subject, root subject drop, is systematically observed in natural production corpora throughout the third year of life. Is this phenomenon restricted to child language? No, various adult languages, as diverse as some Arabic dialects, several Romance dialectal varieties, Brazilian Portuguese, etc. have this peculiar subject drop option, while others do not, e.g. French, Standard English, etc., so the option appears to be a genuine parameter of Universal Grammar.

(39) Gruyère Franco-Provençal (De Crousaz and Shlonsky 2003)

a. (i) travayè pra
'(s/he) works a lot'

b. Portyè *(i) travayè?
'Why s/he works'

c. (i) travayè kan?
'S/he works when?'

(40) Levantine Arabic (Kenstowicz 1989)

a. ___ istarat l-fustaan
'___ bought the dress'

b. *Fariid kaal innu ___ istarat l-fustaan
'Fariid said that ___ bought the dress'

c. Fariid kaal inn-ha istarat l-fustaan

(41) Brazilian Portuguese (Figueiredo-Silva 1994)

a. ___ comprei um carro ontem
'(I) bought a car yesterday'

b. *O que que ___ comprei ontem?
'What that (I) bought yesterday?'

It is also found in specific registers of other languages, such as the "diary registers" of English and French (Haegeman 2000).

In conclusion, children systematically entertain a parametric value which is not target consistent in the course of the third year of life. Why do they do so? A plausible speculation is that young children follow a simplification strategy like the following:

(42) Adopt parametric values which reduce the computational burden of an immature production system and are not directly contradicted by the data of experience. (Rizzi 2000)

In particular, grammatical options are recruited which license different kinds of ellipsis, thus permitting the saving of the articulatory effort involved in the pronunciation of the omitted positions. As the production system matures, the option is abandoned, or "forgotten" by the child, if not supported by experience, and kept otherwise. If this interpretation is on the right track we have here an example of "learning" as "reduction of the options" also on a higher level

of linguistic organisation such as syntax. This simplifying strategy may remain active also in early second language acquisition, given certain similarities in the developmental course (White 2003), and speakers may go back to it when using registers specifically connotated as "abbreviated" (Haegeman 2000)).

"Learning by forgetting" in grammar may just be a particular instance of a selective mode of acquisition of knowledge generally operative in biological systems. In his Nobel lecture entitled The Generative Grammar of the Immune System, immunologist Niels Jerne (1985) has underscored the fact that other biological systems, such as the immune system, typically "learn" through selective mechanisms of this sort:

(43) "Looking back into the history of biology, it appears that wherever a phenomenon resembles learning, an instructive theory was first proposed to account for the underlying mechanisms. In every case this was later replaced by a selective theory" ...[in immunology, an animal] cannot be stimulated to make specific antibodies, unless it has already made antibodies of this specificity before the antigen arrives"
... [selective mechanisms are mechanisms] through which products which are already present in the system prior to the arrival of the signal are selected and amplified". (Jerne 1985)

"Experience", or the contact with the external world, has the role of selecting and amplifying elements which are already present in the system, and no transfer of structure ("instruction") takes place from the world to the learning system (see also the discussion of these points in Piattelli-Palmarini (1989) and, on the neurophysiological bases of learning, the comprehensive discussion in Changeux 2002).

5. Minimalism and Language Design

If the introduction of the principles and parameters framework represented a real break with respect not only to previous generative approaches, but also to an ancient tradition of language description, the Minimalist program represents, I believe, a development in full continuity with the principles and parameters guidelines. Even though various technical assumptions are changed, the fundamental empirical results and formal achievements of the parametric approach are

assumed, and the concept of Universal Grammar as a system of principles and parameters is fully preserved and taken as a basis to ask further questions.

The fundamental question which motivated the minimalist program in the early nineties was the following:

(44) Is there a deeper level of explanation of the language faculty, going beyond the precise identification of principles and parameters?

If a successful theory of UG is said to reach "explanatory adequacy", the goal of the Minimalist Program is the quest for a level of explanation "going beyond explanatory adequacy" (Chomsky 2004). Why does UG specify the empirically observed principles and parameters, rather than some other logically conceivable alternatives?

One line of inquiry which proved fruitful to ask such further questions was to try to understand how the language faculty connects and interacts with other systems and faculties that the individual possesses. Language is sound with meaning, and the language faculty allows us to associate representations of sounds and representations of meanings over an unlimited range. We possess systems capable of producing and analysing sounds, the articulatory/auditory systems, whose components are involved in functions quite independent from language, nutrition, respiration, perception of background noise, etc. And systems of thought, difficult to study in isolation from language in our species (as language offers such an efficient way to precisely structure and express thought), but which plausibly do not completely overlap with language: we can think non-linguistically, e.g. through images, and non human primates, or humans affected by severe language pathologies can engage in mental activities which we would want to characterize as "thought". So, we can think of the language faculty as a device to connect representations of sounds and representations of meanings, both "legible", and usable in various ways, by language-external systems, articulatory-auditory systems, and systems of thought. According to Phase Theory, developed by Chomsky (2001) and Nissenbaum (2000), the outcome of the syntactic computation is periodically sent to the sound and meaning interfaces, and made available to the external system for use.

If one reasons along these lines, a natural dimension to pursue question (44) is to try to see if certain properties of the language faculty are tailored on the needs of the external systems at the interface. So, some properties of the language faculty, say some version of Kayne's (1994) Linear Correspondence Axiom, could be required in order to encode the inherently hierarchical linguistic representations into linear sequences, to make them compatible with the human articulatory system (Saussure's principle of « Linearité du signifiant »). On the meaning side, considerations of this sort are more speculative, as so little is known on the properties and demands of language-independent thought; nevertheless it is not implausible that certain properties like the theory of arguments structures, Theta Theory, the theory of binding and referential dependencies, perhaps the very existence of movement, etc. may be required by the needs and demands of human thought systems, to encode structures properly interpretable and able to express an appropriate range and variety of meanings.

So, one partial positive answer to question (44) may involve the following line of argument:

(45) Some properties of the language faculty are determined by the requirements of the systems of sound and thought at the interfaces.

This mode of "further explanation" is very plausible in some cases, highly speculative in others. Even granting its potential, it clearly could not provide the basis for a complete "further explanation" of the properties of the language faculty: the computing machine has very subtle and complex modes of functioning, highlighted by many years of studies of the fine properties of structures and movement, which seem to respond to some inner logic of the system. Can one spell out this inner logic? One inherent component in the computing machine is the fundamental recursive mechanism, which we have already described. What else?

That principles of economy may shape language design and language functioning is not a recent idea: economy plays a basic role in the classical structuralist idea of the systematic organisation of linguistic inventories, and is central in the Gricean approach and other theories of language use for communication (e.g. Relevance Theory, see Sperber and Wilson 1986). Within generative grammar, economy, in the form of a simplicity measure, was the assumed basis for the evaluation metric, in the attempts to model language acquisition in pre-parametric approaches. In more recent years, a number of independent trends of research highlighted the role of economy in syntax (Collins 1997, Chomsky 2000, 2001). One important economy concept is *locality*, the fact that syntactic

processes take place involving limited portions of structure, so that only minimal memory resources are needed to perform an operation (Rizzi 1990). Another economy concept is a kind of "minimal effort" or "last resort" principle which justifies the application of a given procedure only if it determines some interface effect, so that there is no true optionality in natural language syntax. Other economy principles ensure the optimal simplicity of the derivational procedure and of the representations which are computed. If these lines of inquiry are on the right track, the "inner logic" ruling the functioning of the language faculty is the logic of simple, parsimonious, efficient computation, minimizing entities, steps and resources to achieve the structural results demanded by interface systems.

(46) The language faculty computes sound/meaning pairings in accordance with general principles of economy and efficient computation

Chomsky explicitly connects this factor to the general underlying assumption of the "perfection of nature", an assumption which has profoundly influenced the birth of modern science. Another quote from Galileo, which I borrow from Chomsky (2003), is appropriate here:

(47) "...we have been guided...by our insight into the character and properties of nature's other works, in which nature employs the least elaborate, the simplest and easiest of means. For I do not believe that anybody could imagine that swimming or flying could be accomplished in a simpler or easier way than that which fish and birds actually use by natural instinct."
(Galileo Galilei 1963)

Chomsky's idea is that this fundamental underlying intuition can be vindicated for the naturalistic study of cognitive systems, and of the language faculty, in particular.

Guidelines (45) and (46) thus offer some promise that question (44) may be successfully addressed. If these ideas are pursued, the picture of the language faculty which emerges is highly simplified. The internal structure of the language faculty and its way of functioning is either determined by the needs of the external systems it connects, systems dealing with sounds and meanings, or by general principles of economy and efficient computation which constrain the application of the fundamental recursive operation, principles which may be seen

to be operative in completely different domains of the activity of the mind/brain, and in other complex systems of the biological and physical world.

6. Minimalism and Language Design

Minimalism has achieved promising results in simplifying the inventory of computational mechanisms and principles to a bare minimum. But if the basic ingredients of computations are simple, this does not imply that the representations which are computed are simple too: complexity may quickly arise from the interplay of very elementary mechanisms.

I would like to conclude this chapter by illustrating a particular trend of research which, while adopting basic minimalist guidelines, gave itself the task of studying the complexity of syntactic representations in great detail. Representations like (24) are only very rough approximations which do not do justice to the complexity of syntactic structures. In fact if one starts studying the fine properties of the different configurations, one quickly arrives at the postulation of much richer representations, in which the nodes and branches in (24) correspond to fully articulated structural zones. The study of the fine details of syntactic representations has been actively pursued over the last ten years in the so called "cartographic projects", projects which try to draw maps as detailed and realistic as possible of syntactic structures (for some results, see Belletti, ed. 2004, Cinque, ed. 2002, Rizzi, ed. 2004, and for a general assessment Cinque & Rizzi 2010, Shlonsky 2010). For instance, as soon as we try to investigate in detail the left periphery of the clause, the simple C (complementizer) system expressed in (24) reveals a complex internal structure, able to host different positions (relative and interrogative operators, topic, focus, etc.) in a fixed order; at least something like the following articulation seems to be needed:

(48) The cartography of the C(omplementizer) system, or left periphery of the clause (Rizzi 1997, 2004):
Force > Topic > Interrogative yes/no > Topic > Focus > Modifier > Q > Finiteness > Clause

The following example in Italian illustrates the simultaneous utilization of several left peripheral positions in Italian:

(49)	Credo	che	a Gianni,	QUESTO,	domani,	gli
		Force	Top	Foc	Mod	Fin
	dovreste dire					
	Clause					
	'I believe	that	to Gianni	THIS	tomorrow	you should
	tell him'					

And, for instance, the necessity of differentiating the position of yes/no question markers from the target of wh movement (Q in (48)) is shown, in languages like Italian, by the fact that the yes/no interrogative marker necessarily precedes the focus position, while the target of wh movement necessarily follows it in indirect questions:

(50)a. Mi domando se A GIANNI potremmo dire questo, non a Piero
'I wonder if TO GIANNI we could say this, not to Piero'
b. Mi domando A GIANNI che cosa potremmo dire, non a Piero
'I wonder TO GIANNI what we could say, not to Piero'

There simply isn't enough room in representations like (24) to accommodate all these elements, and the idea of having them adjoined to TP cannot naturally express the fact that, when more elements can cooccur, they typically occur in a fixed order, partly universal and partly subjected to parametric variation. Much work has been devoted to the study of the clause-peripheral positions across languages along these lines: Romance (Rizzi 1997, Benincà & Poletto 2004 and many other contributions in Rizzi ed., 2004), Germanic (Grewendorf 2002, Haegeman 2006), West African languages (Aboh 2004, 2007), Creole languages (Durrleman 2007), East Asian languages (Endo 2007, Tsai 2007, Saito 2010).

The cartographic projects focus on this complexity as an independent topic of inquiry, worth pursuing for it's own sake, but clearly interacting with virtually all aspects of syntactic research.

Along similar lines, focusing on the fine structure of clauses, Guglielmo Cinque has identified a very detailed functional structure, which may be expressed by different devices in different languages (adverbs, particles, verbal morphology), but whose hierarchical order remains remarkably constant across languages. Cinque (1999, 2004, 2006) expresses this universal backbone of the clausal structure in terms of a universal hierarchy of functional heads, expressing

mood and modality, tense, aspect, and voice:

(51) $MoodP_{speech\ act}$ > $MoodP_{evaluative}$ > $MoodP_{evidential}$ > $ModP_{epistemic}$ > TP(Past) > TP(Future) > $MoodP_{irrealis}$ > $ModP_{alethic}$ > $AspP_{habitual}$ > $AspP_{repetitive(I)}$ > $AspP_{frequentative(I)}$ > $ModP_{volitional}$ > $AspP_{celerative(I)}$ > TP(Anterior) > $AspP_{terminative}$ > $AspP_{continuative}$ > $AspP_{retrospective}$ > $AspP_{proximative}$ > $AspP_{durative}$ > $AspP_{generic/progressive}$ > $AspP_{prospective}$ > $ModP_{obligation}$ > $ModP_{permission/ability}$ > $AspP_{Completive}$ > VoiceP > $AspP_{celerative(II)}$ > $AspP_{repetitive(II)}$ > $AspP_{frequentative(II)}$

The need of complex representations is again directly illustrated by sentences in which different adverbials cooccur in a fixed order:

(52) Da allora non hanno di solito mica rimesso più sempre completamente tutto bene in ordine
'Since then, they haven't usually not any longer always put everything well in order'

Other zones of the syntactic tree have been explored along similar lines, giving rise to other detailed maps of nominal expressions and other phrasal categories, and these cartographic projects are now actively pursued to get global structural maps of what is possible in linguistic structures (see Cinque 2002, 2005, Belletti 2004, Rizzi 2004 and the subsequent volumes of the subseries "The cartography of syntactic structures" of the Oxford Studies in Comparative Syntax, among many other publications).

Cartographic projects have been developed in parallel with the development of minimalism, following partially independent trails. There are clear points of connection, such as the central role of economy considerations and the emphasis on the interfaces. There are also points of theoretical tension, at least at first sight. The cartographic projects underscore the richness and complexity of syntactic structures and try to provide realistic descriptions of this complexity. Minimalism tries to capture the fundamental empirical results of syntactic theory through a set of descriptive tools which is substantially impoverished with respect to previous versions of the Principles and Parameters framework. The apparent tension manifests itself very directly in the fact that syntactic representations in much minimalist literature (starting from Chomsky 1995, ch. 4

and much subsequent work) look somewhat simpler than the representations normally assumed some ten years ago, while cartographic representations (e.g. in Cinque 1999, Cinque, 2002, Belletti, ed. 2004, and in this volume) look substantially enriched.

My view is that minimalism and cartography are fully consistent with one another, and the apparent tension is in fact a matter of complementarity and division of labour. Minimalism tries to identifies the basic computational processes, and cartography is concerned with the inventory and combinatorial properties of the syntactic atoms which enter the computations.

One driving factor of the cartographic endeavour is a fundamental intuition of simplicity, clearly akin to core ideas of minimalism. Complex structures arise from the proliferation of structural units which are extremely simple. Ideally, one structural unit (a head and the phrase it projects) is defined by a single syntactically relevant property; so, in the ideal case we have a correspondence between one property, one feature and one head: one relevant morphosyntactic property is formally expressed by a single feature, which enters the syntactic computation in the form of a single head. Complex heads obviously exist in syntax, much as complex words exist in morphology (e.g., a verb inflected for voice, aspect, tense, mood and agreement), but they are not syntactic primitives: they can be assembled by the operation of head to head movement, the only device available to create conglomerates of syntactically relevant properties. Local simplicity is preserved by natural languages at the price of accepting a higher global complexity, through the proliferation of structural units and the recursive application of Merge and Move.

The same intuition, preservation of local simplicity, seems to underlie the functional motivation of phrasal movement in the Minimalist Program, at least of certain kinds of phrasal (A') movement to the left periphery. These kinds of movement exist, it is assumed, to allow elements to carry two types of interpretive properties: argumental and scope-discourse. Through movement, an element can occur in distinct positions specialized for the two kinds of interpretive properties (see, in particular, Chomsky 2001, 2004). So for instance the nominal expression *[D books]* must be interpreted as an argument of the embedded verb *read*, and as a question operator, with main clause scope in (53a). This result is achieved by having the element occur twice in representations like (53b), in positions assigning to it the two kinds of properties (only the higher copy is pronounced, the lower copy, the "trace" of movement, remaining silent):

(53)a. Which book do you think I should read?
b. Which book do you think I should read <which book>?

But why couldn't languages attribute both kinds of interpretive properties to the same position, thus avoiding multiple occurrences of elements? Again, preservation of local simplicity seems to be the key factor: natural language design favours local attribution of single properties, and is prepared to pay the price of multiplying the occurrences: recursion, and proliferation of copies, is cheap, local computation is expensive and to be reduced to the bare minimum.

What particular kind of simplicity natural language design chooses to favour is an empirical question. To quote Chomsky:

(54) " "Good design" conditions are in part a matter of empirical discovery, though within general guidelines of an aprioristic character, a familiar feature of rational inquiry ... Even the most extreme proponents of deductive reasoning from first principles, Descartes for example, held that experiment is critically necessary to discover which of the reasonable options was instantiated in the actual world". (Chomsky 2001a:1–2)

In different domains, the empirical evidence seems to suggest that natural languages favour local simplicity, and accept to pay the price of ending up with global representations involving such complex properties as multiple occurrences of elements (movement), and a very rich articulation of functional structures.

In conclusion, the cartographic results are fully consistent with the basic computational ingredients postulated by bare phrase structure and minimalism (merge and the connected principles of tree construction), and with the conceptual and methodological guidelines of minimalism, but they point to the necessity of substantially enriching the vocabulary of operative heads. So, there is a fruitful tension between the desideratum of simplification of minimalism and the discovery and description of complexity stemming from cartography. We expect from the future of these studies the identification of a point of equilibrium between these factors.

7. Questions Revisited

In this final section I would like to propose a rephrasing of our initial

questions on the basis of recent findings, some of which we reviewed, and with an eye to what, on the basis of our current understanding, looks like a plausible future line of development.

(1') Can language reduce to recursion plus interface requirements, plus economy? How can this reductionist approach be reconciled with the complexity emerging from the cartographic studies?

This is the fundamental question of minimalism, as we have seen; the years to come should tell us if the bold reductionist working hypothesis of minimalism can withstand detailed empirical scrutiny, and, among other things, can be made fully consistent with the results of the cartographic studies and the many sources of empirical evidence which bear on these questions.

(2') What is the format of parameters? How does parameter setting take place? Where do first and second language acquisition converge? Where do they diverge and why? What is the status of bilingualism?

This set of questions is likely to acquire an increasingly central role in future acquisition studies, both for its inherent intellectual interest, and for its potentially wide social impact, in an increasingly multiethnic society, in which bilingualism and the acquisition of a second language in different periods of life will become more and more common and necessary. The identification of a restrictive format for parameters and the drawing of a detailed temporal chart of parameter fixation will be essential to fully understand the acquisition process, and also to have a sound basis for the study of developmental pathologies affecting language. Conducting a very detailed comparison between first language acquisition and the acquisition of a second language at different ages will be crucial to fully appreciate what aspects of the language faculty remain active throughout the human life, and what aspects weaken, and within what temporal windows, with implications for methodologies in language teaching and policies in the school systems.

(3') How does the grammar relate to the parser and the production system?

This question is critical for the integration of linguistic models into the

performance models of experimental psycholinguistics. Well beyond that, it has substantial implications for the language related computational technology which pervades our society, from speech recognition to automatic translators, to all sorts of man-machine interaction systems. These technologies are often developed without taking into account the results of the fundamental research on language, but that should be corrected, sooner or later, as a principle-based approach to applications offers the promise of a higher generality and scope of the solutions to be adopted.

(4') What is the neural basis of the UG ingredients: recursion, principles, parameters...? What kind of neural circuitry is involved in first language acquisition, bilingualism, early and adult second language acquisition?

The fourth question has become the focus of an enormous attention, due to the rapid development of brain imaging technology. The attractive horizon for these studies is an integrated approach to the properties and functioning of the mind/brain, an integrated approach to the different levels of analysis of mental computation, from the cellular level to the abstract functional modelling. A detailed integration probably is a distant goal, but one which can guide interdisciplinary research along converging paths. Here too, there is a definite social dimension: in an aging society, it is of great importance to identify in detail the consequences on the language faculty of degenerative diseases, in order to favour early diagnosis and effective rehabilitation or preservation techniques. And a vast array of acquired aphasic syndromes still offers ample room for progress in interdisciplinary neurolinguistic approaches for the refinement of typology, diagnostic and rehabilitation techniques.

(5') How much of the language faculty in a broad sense is unique to our species? How much is shared with higher mammals and other species? How did recursion become available to the human communication system? Why did the language faculty evolve so as to permit more than one language?

We have already commented on these questions, it's a set of issues which is, and will probably remain, at the margins of scientific inquiry: prominent biologists like Richard Lewontin have expressed scepticism on the possibility that

these problems may ever be addressed in scientifically satisfactory terms. Still, we cannot avoid addressing the evolutionary issues in an attempt to understand what makes our species unique. And it is conceivable that the study of language may offer a model of rational inquiry, if not of hard science, for the broader study of the natural bases of human cultural systems and their evolution.

Note

* This paper was presented as a general lecture at McGill University, Montreal, March 27, 2006.

References

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

Aboh, Enoch (2007) Information Structuring Begins with Numeration. MS. ACLC, University of Amsterdam.

Baker, Mark (2001) *The Atoms of Language*. New York: Basic Books.

Belletti, Adriana (2001) "Inversion as Focalization." In Hulke, Aafke and Jean-Yves Pollock (eds.), *Inversion in Romance*. Oxford: Oxford University Press.

Belletti, Adriana ed. (2004) *Structures and Beyond – The Cartography of Syntactic Structures, Vol. 3*. Oxford: Oxford University Press.

Benincà, Paola and Cecilia Poletto. (2004) "Topic, Focus, and V2. Defining the CP Sublayers." In Luigi Rizzi (ed.), *The Structure of CP and IP. The Cartography of Syntactic Structures, vol.2* 52–75. New York: Oxford University Press.

Biberauer, Theresa, Ander Holmberg and Ian Roberts (2010) "A Syntactic Universal and its Consequences." To appear in *Linguistic Inquiry*.

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

Brown, Roger (1973) *A First Language: The Early Stages*. Cambridge, Mass.: Harvard University Press.

Changeux, Jean-Pierre (2002) *L'homme de Vérité*. Paris: Odile Jacob.

Chomsky, Noam (1957) *Syntactic Structures*. Mouton: The Hague.

Chomsky, Noam (1959) "A Review of B.F. Skinner's Verbal Behavior." *Language* 35: 26–58.

Chomsky, Noam (1973) "Conditions on Transformations." In Stephan. Anderson and Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart and Winston.

Chomsky, Noam (1981) *Lectures on Government and Binding*. Dordrecht : Foris.

Chomsky, Noam (1995) *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chomsky, Noam (2000) "Minimalist Inquiries: The Framework." In Roger Martin, David Michaels and Juan Uriagereka (eds.), *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, Mass.: MIT Press.

Chomsky, Noam (2001) "Derivation by Phase." In M. Kenstowicz, (ed.), *Ken Hale: A Life in Language*. Cambridge, Mass.: MIT Press.

Chomsky, Noam (2002) *On Nature and Language*. Cambridge: Cambridge University Press.

Chomsky, Noam (2004) "Beyond Explanatory Adequacy." In Belletti (ed.) (2004).

Cinque, Guglielmo (1999) *Adverbs and Functional Heads*. New York: Oxford University Press.

Cinque, Guglielmo (2002) *Functional Structure in DP and IP*. Oxford: Oxford University Press.

Cinque, Guglielmo and Luigi Rizzi (2010) "The Cartography of Syntactic Structures." *Handbook of Syntactic Theories*. Oxford: Oxford University Press.

Collins, Christopher (1997) *Local Economy*. Cambridge, Mass.: MIT Press.

Crain, Steven (1992) "Language Acquisition in the Absence of Experience." *Behavioral and Brain Sciences* 14: 597–611.

D'Arcy Thompson, Wentworth (1917) *On Growth and Form*. Cambridge: Cambridge University Press.

Darwin, Charles (1871/1981) *The Descent of Man*. Princeton: Princeton University Press.

De Crousaz, Isabella and Ur Shlonsky (2003) "The Distribution of a Subject Clitic Pronoun in a Franco-Provençal dialect." *Linguistic Inquiry* 34: 413–442.

Dehaene, Stanislas (1997) *La bosse des Maths*. Paris: Odile Jakob.

Descartes, René (1637/1951) *Discours de la Méthode*. Paris: Union générale d'éditions.

Dryer, Matthew S. (1992) "The Greenbergian Word Order Correlations." *Language* 68, 81–138.

Durrleman, Stephanie (2007) *The Syntax of Jamaican Creole – A Cartographic Perspective*. Amsterdam: John Benjamins.

Endo, Yoshio (2007) *Locality and Information Structure – A Cartographic Approach to Japanese*. Amsterdam/Philadelphia: John Benjamins.

Figueiredo, Silva Maria (1994) *La Position Sujet en Portugais Brésilien*. Ph.D. dissertation, University of Geneva; published in 1996. Campinas: Editora da Unicamp.

Friedemann, Marc Ariel and Luigi Rizzi (eds.) (2000) *The Acquisition of Syntax*. London: Longman.

Friederici, Angela (2000) "The Neuronal Dynamics of Auditory Language Comprehension." In Marantz et al. (ed.).

Galilei, Galileo (1630/1970) Dialogo Sopra I due Massimi Sistemi del Mondo, (éd.), Torino: Einaudi.

Gibson, Edward and Kenneth Wexler (1994) "Triggers." *Linguistic Inquiry* 25, 407–454.

Greenberg, Joseph (1963) "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements." In Josef Greenberg (ed.), *Universals of Language*, 73–113. Cambridge, Mass.: MIT Press.

Grewendorf, Günther (2002) "Left Dislocation as Movement." In Simon Manck and Jennifer Mittelstaedt, (eds.), *Georgetown University Working Papers in Theoretical Linguistics* 31–81.

Grodzinsky, Yosef (2000) "The Neurology of Syntax: Language Use without Broca's Area." *Behavioral and Brain Sciences*, 23, 1–71.

Grodzinsky, Yosef (2005) "Syntactic Dependencies as Memorized Sequences in the Brain." Ms., McGill, Montreal.

Haegeman, Liliane (2000) "Adult Null Subjects in non pro Drop Languages." In Mark

Friedemann and Luigi Rizzi (eds.) *The Acquisition of Syntax*. Harlow: Longman.
Haegeman, Liliane (2006) "Argument Fronting in English, Romance CILD and Left Periphery." In Raffaela Zanuttini, Hector Campos, Elena Herburger and Paul Portner (eds.), *Crosslinguistic Research in Syntax and Semantics: Negation, Tense and Clausal Architecture*. 27–52. Washington, D.C.: Georgetown University Press.
Haegeman, Liliane (2012) *Adverbial Clauses, Main Clause Phenomena, and Composition of the Left Periphery: The Cartography of Syntactic Structures, Volume 8*. New York: Oxford University Press.
Hamann, Cornelia (1997) *From Syntax to Discourse. Children's Use of Pronominal Clitics, Null Subjects, Infinitives, and Operators*. Habilitation Thesis, University of Tübingen.
Hamann, Cornelia and Kim Plunkett (1997) "Subject Omission in Child Danish." *BU Conference on Language Development* 21, 220–231. Somerville: Cascadilla Press.
Hamann, Cornelia, Rizzi, Luigi and Uli Frauenfelder (1996) "On the Acquisition of Subject and Object Clitics in French." In Herald Clahsen (ed.), *Generative Perspectives in Language Acquisition*, 309–334. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Hauser, Marc (1996) *The Evolution of Communication*. Cambridge, Mass.: MIT Press.
Hauser, Marc, Noam Chomsky and Tecumseh Fitch (2002) "The Faculty of Language: What is it, Who has it, and How did it Evolve?" *Science*, 298.
Hyams, Nina (1986) *Language Acquisition and the Theory of Parameters*. Dordrecht: Reidel.
Kayne, Richard (1994) *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.
Kayne, Richard (2005) *Movement and Silence*. Oxford: Oxford University Press.
Kenstowicz, Michael (1989) "The Null Subject Parameter in Modern Arabic Dialects." In Jaeggli, Osovaldo and Ken Safir (1989) (eds.) *The Null Subject Parameter*. Dordrecht: Kluwer.
Jenkins, Lyle (2000) *Biolinguistics: Exploring the Biology of Language*. Cambridge: Cambridge University Press.
Jerne, Niels (1985) "The Generative Grammar of the Immune System." *Science* 229: 1057–1059.
Lenneberg, Eric H. (1967) *Biological Foundations of Language*. New York : John Wiley & sons.
Marantz, Alec (1997) "No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of your own Lexicon." In Alexis Dimitriadis and Laura Siegel (eds.), *University of Pennsylvania Working Papers in Linguistics*, 4, 201–225.
Marantz, Alec, Yasushi Miyashita, Wayne O'Neil (eds.), (2000) *Image, Language, Brain*. Cambridge, Mass.: MIT Press.
Mehler, Jacques and Emmanuel Dupoux (1990) *Naitre Humain*. Paris: Odile Jacob.
Neville, Helen, Janet Nicol, Andrew Barss et al. (1991) "Syntactically based Sentence Processing Clauses." *Journal of Cognitive Neurosciences* 3, 151–165,
Nicolis, Marco (2005) *On pro-drop*. Ph.D. dissertation, University of Siena.
Nissenbaum, Jon (2000) *Investigations of Covert Phrasal Movement*. Ph.D. dissertation, Cambridge, MA : MIT.
Piattelli-Palmarini, Massimo (1989) "Evolution, Selection and Cognition: from "Learning" to Parameter Setting in Biology and the Study of Language." *Cognition*, 31, 1–44.
Pinker, Steven (1994) *The Language Instinct*. New York: Morrow.
Rice, Mabel L., Kenneth Wexler and Patricia L. Cleave (1995) "Specific Language Impairment as a Period of Optional Infinitive." *Journal of Speech and Hearing Research* 38: 850–863.
Rizzi, Luigi (1978) "Violations of the Wh Island Constraint in Italian and the Subjacency Condition." *Montreal Working Papers in Linguistics* 11.

Rizzi, Luigi (1982) *Issues in Italian Syntax*. Dordrecht: Foris.
Rizzi, Luigi (1986) "Null Objects in Italian and the Theory of *pro*." *Linguistic Inquiry* 17, 501–557. Republished in Rizzi (2000).
Rizzi, Luigi (1990) *Relativized Minimality*. Cambridge, Mass.: MIT Press.
Rizzi, Luigi (1997) "The Fine Structure of the Left Periphery." In Liliane Haegeman (ed.), *Elements of Grammar*, 281–338. Dordrecht: Kluwer.
Rizzi, Luigi (2000) *Comparative Syntax and Language Acquisition*. London: Routledge.
Rizzi, Luigi ed. (2004a) *The Structure of CP and IP. The Cartography of Syntactic Structures, vol.2*. New York: Oxford University Press
Ross, John Robert (1967) *Constraints on Variables in Syntax*. Ph.D. dissertation, MIT.
Ross, John Robert (1986) *Infinite Syntax!* N.J.: Norwood: Ablex.
Saito, Mamoru (2010) "Sentence Types and the Japanese Right Periphery." Ms., Nanzan University, Nagoya.
Saussure, Ferdind de (1916/1985) *Cours de Linguistique Générale*. Paris: Payot.
Shlonsky, Ur (2010) "The Cartographic Enterprise in Syntax." *Language and Linguistics Compass* 4(6).
Sperber, Dan and Deldren Wilson (1986) *Relevance*. Oxford: Blackwell.
Tattersal, Ian (2008) *The World from Beginnings to 4000BCE*. Oxford: Oxford University Press.
Tsai, Dylan (2007) Left Periphery and Why-How Alternations. Ms., Taiwan: National Tsing-Hua University.
Werker, Janet F., Gilbert, John. H. V., Humphreys, Keith and Tees, Richard C. (1981) "Developmental Aspects of Cross-Language Speech Perception." *Child Development* 52, 349–355.
White, Lydia (2003) *Second Language Acquisition and Universal Grammar*. Cambridge: Cambridge University Press.