

## ON THE STUDY OF LANGUAGE AS A COGNITIVE CAPACITY

**Luigi Rizzi**

*Università di Siena*

**Abstract:** Linguistics in the second half of the century is marked by the cognitive revolution, profoundly influenced and partly triggered by Chomsky's ideas. Universal Grammar is the key concept for the investigation of predetermined linguistic knowledge; it can be studied by integrating evidence from poverty of the stimulus, language comparison and language development. In the last twenty years, "Principles and Parameters" models of Universal Grammar have renewed comparative syntax and the study of language acquisition.

**Keywords:** linguistics, cognitive sciences, language acquisition, syntax, comparative syntax, morphology, universal grammar, parameters.

The topic of this section is much too broad for me to aim at even minimal systematicity. So I will limit these notes to a rather narrow range of research topics, sufficiently interconnected to form a coherent body of research programs, and selected primarily on the basis of personal competence and taste. My focus will be on some research trends which have studied language as a human cognitive capacity through the use of precise formal models in the second half of the century. Even within such a limited scope, no attempt will be made to aim at exhaustiveness.

### 1. A NATURALISTIC APPROACH TO THE STUDY OF LANGUAGE AND MIND.

Linguistics in the second half of the century is marked by the cognitive revolution, which was in turn profoundly influenced and partly triggered by Chomsky's ideas on language and mind. These ideas renewed the study of language by redefining objects and methods of linguistic inquiry. At the beginning of the century, the attempt to define the object of the discipline had

led Ferdinand de Saussure to draw a distinction between "langue" and "faculté du langage". Saussure's "langue" is "...un ensemble de conventions nécessaires adoptées par le corps social pour permettre l'usage de la faculté du langage chez les individus...", "un produit social de la faculté du langage", an entity which is only partially represented in an individual mind/brain ("...elle n'existe parfaitement que dans la masse..."); on the other hand, the language faculty is "une faculté que nous tenons de la nature, ... un instinct naturel..." (Saussure 1985, *passim*), in modern terminology, a species-specific and task-specific cognitive capacity. Saussure then went on to build his approach to linguistics based on the study of "langue". About half a century later, Noam Chomsky proposed to go the other way, and centered his reflection on language on the study of the language faculty.

The idea of focusing on the language faculty was not new. The broader perspective of studying language as a "mirror of the mind", as a domain offering a privileged access to the study of the human cognitive capacities, had been consistently pursued in the history of the rational inquiry on language. On the other hand, Chomsky was the first to show that the scientific study of the language faculty was a feasible enterprise: it could be conducted through the use of precise formal models of language knowledge, language acquisition and language use, models possessing a non-trivial deductive structure, rich in heuristic value and accessible to a variety of modes of empirical validation (Chomsky 1957, 1959, 1965). This is the essence of the "Chomskian revolution": the language faculty is accessible to study in the "Galilean style", the style of inquiry which established itself in the natural sciences since the XVII century; the ramifications of this approach have extended well beyond the boundaries of technical linguistics and have profoundly influenced the cognitive studies in the second half of this century.

It is often said that the modern study of language as a "mirror of the mind" revolves around a host of basic questions such as the following:

- What is knowledge of language?
- How is it acquired?
- How is it accessed in language use?
- How is it concretely represented in the human brain?

The question of language knowledge turned out to be of critical importance for the program to get started. The first fragments of generative grammars in the fifties and sixties (e.g., in Chomsky (1957), Fodor & Katz (1964)) showed, on the one hand, that linguistic knowledge was accessible to study through certain precise techniques which had their roots in the theory of formal systems; on the other hand, these studies quickly brought to light the richness and complexity of the intuitive knowledge that speakers share: every speaker implicitly knows about the structure of his language much more than he is consciously aware of; in fact, every speaker implicitly masters a rich system of precise instructions to form and interpret linguistic expressions; this system, somehow concretely represented in our brains and largely beyond the reach of introspection, is constantly used by the speaker to produce and understand novel sentences, a normal characteristic of ordinary language use.

The discovery of the structural richness of the implicit knowledge of language immediately underscored the complexity of the cognitive task that the language learner is confronted with. Humans acquire a natural language early in life, without specific instruction, apparently in a non-intentional manner, with limited individual variation in spite of the fragmentary and individually variable courses of experience which ground individual knowledge of language.

More importantly, the precise understanding of fragments of the adult knowledge of language quickly underscored the massive presence of "poverty of stimulus" situations: our adult knowledge of language is largely underdetermined by the data available in childhood, which would be consistent with innumerable generalisations over and above the ones that speakers seem to unerringly converge to. This empirical observation is of great importance, as it grounds the necessity of postulating a structured system of predetermined linguistic principles which guide language acquisition; it also leads to the expectation of a fundamental cross-linguistic uniformity of human languages, hence it lays the ground of the recent, cognitively constrained, comparative perspective; it also opens new directions of research in the study of linguistic change and language development.

## 2. ON POVERTY OF STIMULUS

Given the importance of this point, it is worthwhile to illustrate it in some detail through the analysis of a concrete case. Every speaker of English has intuitive knowledge of the fact that a pronoun and a noun phrase can corefer (refer to the same individual(s)) in some structural environments but not in others: for instance, in (1) and (3), but not in (2) and (4)

- (1) John thinks that he will win the race
- (2)\* He thinks that John will win the race
- (3) John's opinion on his father is surprising
- (4)\* His opinion on John's father is surprising

(2) and (4) are of course possible if the pronominal forms he and his refer to some other individual, Peter for instance, but coreference with John is barred. Clearly, speakers of English tacitly possess some procedure for the interpretation of pronouns that they can quickly and efficiently use to compute coreference possibilities in new sentences. This procedure is non-trivial: it is not simply a question of linear precedence: in the following examples the pronominal element precedes the noun phrase, and still coreference is fine:

- (5) When he wins, John is very happy
- (6) All the people who know him well say that John can win the race
- (7) His father thinks that John can win the race

The converse case also exists: in some surface configurations, a pronoun cannot be coreferential to a preceding NP:

- (8)\* To John's father, he does not speak anymore

About twenty years ago, Tanya Reinhart and Howard Lasnik identified the relevant structural configuration and the operative principle, respectively. The configuration is defined by one of the fundamental structural relations, c-command (Reinhart 1976); the operative principle is that coreference is barred when the pronoun c-commands the NP (Lasnik 1976). Simplifying things somewhat, we can say that the domain of c-command (or c-domain) of an element is the

category that immediately contains that element in the phrase structure representation. In the following examples I have indicated by a pair of brackets the c-domain of the pronoun:

- (1') John thinks that [he will win the race]
- (2')\* [He thinks that John will win the race]
- (3') John's opinion on [his father] is surprising
- (4')\* [His opinion on John's father] is surprising
- (5') When [he wins], John is very happy
- (6') All the people who [know him well] say that John can win the race
- (7') [His father] thinks that John can win the race

What singles out (2) and (4) is that only in these structures does the NP fall within the c-domain of the pronoun, regardless of linear ordering: Lasnik's non-coreference principle states that coreference is barred in this configuration, while it is free elsewhere (a few years later Chomsky (1981) showed that Lasnik's statement could be subsumed under a more general binding principle, but this reduction is not of immediate relevance here).

This approach seems insufficient to deal with cases like (8): here, the c-domain of the pronoun is the clause he does not speak anymore; since John falls outside, coreference should be allowed, contrary to fact:

- (8')\* To John's father, [he does not speak anymore]

To make things even more complicated, one can observe that, with somewhat different lexical choices, the same structural configuration seems to allow coreference, as Lasnik's approach would predict:

- (9) In John's (most recent) picture, [he doesn't look in good shape]

The solution to the problem comes from the theory of reconstruction. In (8), the clause initial PP is a selected argument of the verb speak, it has been preposed from a clause-internal complement position, where a trace bound by the preposed element was left. We can think of the trace as a full copy of the preposed element, except that it is not pronounced (under the recent copy theory of traces of the Minimalist Program: Chomsky (1995)), so that the complete representation is the following (the unpronounced copy within angled brackets):

- (8'') To John's father [he does not talk anymore <to John's father>]

If the mind assigns a representation like (8'') to sentence (8), then the non-coreference principle applying on it will correctly compute non-coreference, as the (unpronounced occurrence of) the NP John now is within the domain of the pronoun. On the other hand, the initial PP in (9) is not a verbal argument, but a kind of "scene-setting" adjunct; the mind won't see any unpronounced clause-internal copy here, hence the possibility of coreference will be computed. A complex array of behavioral data, the interpretive judgments that speakers give, then follows from the assumption that certain precise structural principles guide our mental computation of

linguistic representations. The principles that turn out to be empirically adequate are at the same time quite elegant and general (in that they abstract away from the specificity of individual constructions); on the other hand, they are by no means obvious, not amenable to logical necessity, nor to some general condition of communicative efficiency; in fact, they are highly specific to the linguistic computation, as far as we can see.

We should now ask the question of acquisition. How do all speakers come to know that the non-coreference principle holds, and that its effects must be computed over abstract representations containing unpronounced positions like (8''), rather than on concrete representations more directly reflecting the physical structure of the sentence like (8'), etc.? Obviously, such abstract properties are not explicitly taught, adult speakers are totally unaware of possessing such elements of tacit knowledge.

The common sense solution that immediately comes to mind is the one in line with the traditional empiricist approach to the acquisition of knowledge: perhaps, children can somehow figure out such properties via some kind of unconscious induction from their experience through simple mechanisms of analogy and generalization. No matter how a priori appealing such approaches may seem, they do not resist more careful scrutiny stemming from elementary considerations on the kind of empirical evidence that the child has access to. In the domain that concerns us here, what children hear is a series of utterances containing pronouns, manifesting the sentence types of (1) through (8) and indefinitely many others; some such pronouns will be intended by the speaker as coreferential to another utterance-internal NP, others would be intended as non-coreferential utterance-internally (I am assuming that enough contextual information is available to the child to decide if coreference is intended or not in a number of actual utterances available to him). As no negative information is ever directly provided to the learner ("coreference is barred in this particular environment"), the evidence available to him is consistent with the simplest assumption that coreference is free across the board. Why should the learner then assume a ban on coreference in a particular, and very specific, structurally defined environment, as the adult behavioral data unerringly manifest?

One could still object that the assumption on the non availability of negative evidence is too strong. Clearly, the negative information "no coreference in this structural environment" is not directly provided to the learner; still, couldn't it be the case that the learner can figure out the existence of a ban indirectly, from the failure of expected feedback? Suppose for instance the following hypothetical scenario: the language-learner tries to utter structures like (2) and (4) intending coreference; he is then made to realize, through the hearer's feedback, that communication of the intended interpretation failed; could he not then conclude that a non-coreference ban must be postulated in the relevant environments on the basis of such indirect negative evidence? A moment's reflection suffices to make such a variant of the empiricist approach highly implausible. If the acquisition of the non-coreference principle was input-driven, we would expect a wide variety of generalizations across speakers, somehow reflecting the variability of available experience, of the types of indirect negative evidence occurring in the individually available input. So, some speakers would end up barring coreference in (2) but not in (4), or in (4) but not in (2) (or in any other arbitrary subset of the indefinitely many NP-pronoun configurations), others would assume a generalization based on precedence, uniformly barring coreference in (2)-(4)-(5)-(6)-(7), as opposed to (1) and (3); yet other speakers would analogize (8) on (9), allowing coreference in both cases, or (9) on (8), excluding coreference in both; other more bizarre generalizations could be assumed by other speakers; yet other

speakers could have had a particularly uninformative input, as for indirect negative evidence, hence postulate no ban at all, just free coreference everywhere, etc.

No such variation is found at all across speakers of the same linguistic community and, in fact, as we shall see in a moment, across speakers of different languages; the strong uniformity of the shared knowledge, in this and many other similar cases, is clearly inconsistent with any approach considering the knowledge input-driven (even through indirect negative evidence); such cases then forcefully argue for the predetermined character of the relevant elements of tacit knowledge.

There should be little doubt that we are biologically predisposed to acquire a human language. Our cognitive system makes the leap possible between a limited and fragmentary course of experience and the rich system of linguistic knowledge that every speaker possesses. Moreover, as far as the nature of the predetermined system of computational principles is concerned, we have no indication that such principles would cut across different cognitive domains. There seem to be special principles dedicated to linguistic computation, with properties which are quite distinct from the principles that rule other aspects of the human cognition: principles for the interpretation of visual stimuli into coherent scenes, principles determining the planning of action, principles of logical and probabilistic reasoning which ground our capacity to solve non-linguistic problems, to take rational decisions, etc. The cognitive segregation of the system dedicated to linguistic computation is also strongly supported by the study of pathology, with sharp cases of selective impairment (or selective preservation) of linguistic capacities with respect of other cognitive domains, a trend highlighted in particular in much recent work on Specific Language Impairment. To the extent to which these conclusions are correct, a powerful argument comes from modern linguistics for the modular theory of mind, a theory which sees the mind as a system of mental organs, each highly specialized to efficiently perform a certain kind of mental computation (Chomsky 1975, Fodor 1983).

Of course, we should relativize such conclusions to our current level of understanding, as is normal in science, and we should keep an eye open on the conceivable formal developments and empirical discoveries which may lead us to question such conclusions in the future. Advances in connectionist modeling could tell us how far induction and analogy can lead in extracting complex patterns from experience, and on the basis of how much predetermined structure. It is to be hoped by any researcher working within symbolic approaches to cognition that connectionist modeling will start addressing higher level aspects of the cognitive capacities, such as the problem discussed here, rather than confining itself to the acquisition of low level semi-regularities which must involve elements of data-driven learning under any reasonable approach. Such advances could make the boundary between innate and acquired more precise than it is possible today, in the absence of empiricist models structured enough to reach even minimal empirical adequacy in the higher domains of the linguistic capacities. On the other hand, many variants of hybrid models combining the two paradigms are conceivable: for instance, it may be worthwhile to explore connectionist models of parameter setting, with learning functions, thresholds of activation, etc., mechanisms which may be able to give a precise content to the notion of indirect negative evidence, sometimes hinted at in the literature in a richer sense than the one mentioned before (e.g., discarding a grammatical option on the basis of the non-occurrence of its manifestations in environments in which they would be likely to occur). Of course, it could also be that newly based research programs, as the gap with the neurosciences is progressively reduced, may lead to the discovery of overarching principles

cutting across different cognitive domains. What can be safely claimed is that to the best of our current knowledge, the evidence overwhelmingly favors the modular view, a conclusion that may well be definitive at the level of description of cognition which our current knowledge is based on.

### 3. UNIVERSAL GRAMMAR: A THEORY OF PREDETERMINED LINGUISTIC KNOWLEDGE.

Various considerations stemming from sources as diverse as learnability theory, developmental psycholinguistics, comparative linguistics support the view that there is a strong predetermined basis for the acquisition of linguistic knowledge. Poverty of stimulus considerations put a lower bound on the richness of the predetermined basis: if a property of adult knowledge cannot be figured out from experience, it must be predetermined. On the other hand, languages differ, knowledge of Italian and knowledge of Mandarin Chinese are quite different systems of knowledge which, at least as far as practical purposes of communication are concerned, seem to have little if any intersection, and even varieties which are closely related historically may quickly become mutually incomprehensible. Language variation sets an obvious upper bound to the postulation of a predetermined basis. In other words, the innate basis to be postulated must be rich enough to be consistent with the fact that language acquisition is possible, and impoverished or flexible enough to be consistent with the observed cross-linguistic variation. The tension between these two poles has been a major triggering force for the progress of the field in the second half of the century. It triggered, among other things, a renewed interest for comparative studies and the growth of a theory-guided comparative approach.

The comparative studies of the last twenty years or so have shown that the paradox between biological predisposition and language diversity is only apparent. If the description is expressed at the appropriate level of abstraction, language diversity tends to fade away and the massive underlying uniformity of natural languages sticks out. Of course, the difficulty of the exercise consists of identifying the appropriate level of description. The need was to identify a level which could effectively capture the cross linguistic uniformity, and at the same time manage not to lose the expression of the structural complexity and richness of the individual natural languages. The history of the modern concept of Universal Grammar is, in a sense, the history of successive approximations in the search of such an optimal level of description, and of a technical language able to express it in a precise way.

Universal Grammar (UG) can be thought of as a theory of the biologically necessary linguistic invariance. The qualification "biologically necessary" expresses the cognitive basis of the concept. It underscores the difference between UG and a simple list of empirically observed language universals (which could include accidental linguistic universals), and also the difference between UG and the purely conceptual exploration of the notion of possible semiotic or communicative system. UG is the theory of an empirical object, a component of the mind of the speaker which expresses predetermined linguistic knowledge. As the ability to acquire a language is uniformly distributed across the species, UG expresses linguistic universals, but only inasmuch as they are determined by the inner nature of our cognitive system for language.

How can one argue for ascribing a given element of the adult knowledge of language to UG? At least three kinds of evidence can be provided stemming, respectively, from considerations of learnability, comparative linguistics and developmental psycholinguistics. Learnability

considerations allow us to identify situations of poverty of the stimulus of the kind informally discussed above, sometimes with great formal precision (in work such as Wexler & Culicover (1980), Pinker (1984), etc.). The postulation of an element of innate knowledge on the basis of learnability considerations immediately leads to the prediction of its universality: under the plausible and empirically supported assumption of the fundamental uniformity of the human genome for language, the innate element should be universal, at least in the sense that, if a language has the structural characteristics which would make it possible, the element will occur.

Comparative considerations become relevant at this point to check the prediction. For instance, to the best of our current knowledge, the non-coreference effect illustrated above is universal: no known natural language allows coreference between a pronoun and a NP in its domain. The hypothesis of the predetermined character of the non-coreference effect is thus empirically supported.

Some obvious precautions must be taken in the evaluation of the universality of a given property, and more importantly on the legitimacy of ascribing it to the human cognitive system, on the basis of comparative evidence. On the one hand, a language may not have the structural characteristics which would make the manifestation of the property transparent, or possible. If a language does not have a morphologically identified class of pronouns, the testing of the non coreference effect in such a language will have to be different than in English, if possible at all. If a language does not have overt movement of interrogative elements to the front, it will not manifest familiar locality effects which are transparently observable with overt movement, etc.

On the other hand, universality *per se* is no guarantee of the predetermined status, as there is the logical possibility of a (cognitively) accidental universal: it is logically possible that human languages may have converged to all have the same property at some point of their history even if the property is not enforced by the human cognitive constraints on language knowledge. It is also logically possible that all human languages may have preserved some (non-cognitively driven) original common property. The possibility of a cognitively accidental convergence, driven by historical and cultural factors, is rather obvious for trivial lexical "universals" (most languages may well share common terms to designate high technology objects, proper names of personalities of international relevance, etc.) but it is extremely implausible for non-trivial structural universals.

Neither the possibility of an accidental convergence nor the possibility that linguistic universals may be reduced to some unique ancestor of all modern human languages can provide an adequate explanation whenever linguistic universals exist in situations of poverty of stimulus. In such cases, every new generation of speakers of each human language would have to unerringly converge to postulate the same abstract property (e.g., the non-coreference principle) on the basis of an impoverished and individually highly variable evidence, a pure miracle in the absence of specific cognitive constraints enforcing the convergence on an endogenous basis. So, in our search of the predetermined cognitive mechanisms underlying our linguistic ability, we want the comparative evidence to be supported by poverty of stimulus considerations, which provide the primary impulse for the whole research program.

A third kind of evidence comes from developmental psycholinguistics and more generally from the experimental study of the linguistic capacities of children at different ages. If a given element of linguistic knowledge is predetermined, we expect it to manifest itself as soon as the child is able to process structures of sufficient complexity to make detection of the element

possible. For instance, Stephen Crain and others experimentalists have provided evidence that children show sensitivity to the non-coreference effect as early in life as the effect is experimentally testable, thus supporting the hypothesis that the effect is predetermined (Crain 1991).

An important trend of developmental psycholinguistics is now working at techniques and paradigms that make it possible to test sophisticated elements of linguistic knowledge with younger and younger children. Well-established techniques exist to test certain aspects of the linguistic capacities of infants in the domain of the sound systems already a few days (or even hours) after birth. In classical work by Peter Eimas, Jacques Mehler and others, these techniques have made it possible to tap directly on the initial cognitive state of the child, and have already uncovered a good deal in the domain of the innate predispositions of the child for the acquisition of the sound systems of natural languages, as well as on the time course of the early fixation of certain phonological parameters (see Mehler & Dupoux (1992) for a review). Much progress is to be expected in the development of analogous techniques to test predispositions and early mastery of knowledge on the higher levels of morphology, lexicon, syntax and semantics.

Here too, a word of caution is necessary. It is entirely possible (even plausible in some cases) that maturational processes may be involved in language development. If so, an element of adult knowledge could be fully predetermined, but be absent from the initial cognitive state (narrowly construed) and come to existence at some point of the cognitive development following an endogenous maturational schedule. So, UG can be thought of as a theory of the initial cognitive state for language, as is often said, but only if we understand "initial cognitive state" in a way comprehensive enough to include whatever condition of the organism will later determine linguistic maturation (much as the initial state of the visual system can be legitimately construed as including whatever determines the maturation of binocular vision). By combining learnability and comparative considerations (uncovering the elements of predetermined linguistic knowledge) with the experimental study of language development (uncovering what element of knowledge emerges at what developmental stage in language acquisition) we should be able to isolate maturational factors, understand their role in language development, and ultimately relate them to biological indices of the growth of the child.

In conclusion, poverty of stimulus considerations provide the primary source of evidence for UG models; hypotheses constructed on the basis of such evidence immediately invite further empirical testing in the domains of comparative and development studies. Hence the research program quickly branches and diversifies into lines of inquiry involving different disciplinary techniques, and organizes around the question of linguistic knowledge much interdisciplinary research on the human cognitive capacity for language.

#### 4. HISTORY OF UG: FROM GRAMMATICAL METATHEORY TO SYSTEM OF PRINCIPLES AND PARAMETERS

Early models of Universal Grammar were based on (a more formal version of) a rather traditional view of the nature of individual grammars. The system of linguistic knowledge that the adult speaker of a particular language possesses was thought of as a system of language-specific rules, explicit instructions to form and interpret linguistic expressions of a particular human language. Correspondingly, Universal Grammar was thought of as a kind of grammatical metatheory, a theory expressing conditions on the form and functioning of

particular grammars. So, Universal Grammar was supposed to express the format of particular grammars, defining rule types, the technical vocabulary from which particular rules could be constructed, etc. In this sense, UG was supposed to limit the search space available to the language learner. The task of the language learner was seen as the identification of the language particular rules forming an individual grammar on the basis of his linguistic experience and within the search space defined by UG. On the other hand, the constraints on rule format were much too rudimentary till the early or mid seventies (as they admitted fairly unconstrained rewriting and transformational rules) to narrow down the search space in a satisfactory manner: just too many particular analyses remained consistent with the data accessible to the learner. In order to account for the rapidity and effectiveness of language acquisition, it was then assumed that UG provides an evaluation measure which ranks grammatical analyses consistent with a given course of experience along a gradient of simplicity. The assumption was that the language learner selects the most highly valued (or simplest) grammar consistent with his experience.

In the course of the seventies, the two putative components of UG had a very different fate. On the one hand, the notion of evaluation measure remained a purely programmatic idea in syntax (in spite of the considerable success achieved by the same notion in phonology: Chomsky & Halle 1968): it was not amenable to refinements and formal implementation which could give empirical content and explanatory capacity in specific cases to the notion of a formally induced grammatical ranking. As no progress was made, the notion was progressively abandoned in syntactic theory. On the other hand, very fast and substantial progress was made on the effort to constrain the format of particular grammars, to an extent that the whole problem quickly appeared under a very different light. First of all, conceptual and formal work pursued within the so-called Extended Standard Theory showed that the rule format for particular grammars could be radically impoverished without loss of empirical adequacy by assuming a much richer internal structure for UG. Concomitantly, the first systematic attempts at a theory-guided comparative syntax gave strong support to the view of a more fundamental cross-linguistic uniformity than previous models could have led one to expect. For instance, a wealth of cross-linguistic research in Relational Grammar around the mid-seventies managed to show (in work by Perlmutter, Postal and others: Perlmutter 1983) that A(rgumental) structures undergo syntactic processes obeying a fixed set of laws which are by and large constant across languages, a trend which gave rise, among other things, to much important research on the Unaccusative Hypothesis. And the first systematic generative studies on different Romance and Germanic languages, much under the influence of Richard Kayne's work (1975, 1983) and primarily conducted in Europe, strongly supported the attempts to constrain the grammatical format that Chomsky and his students were pursuing in these years. This trend then quickly extended to Semitic, American Indian, East-Asian and other language groups.

Much formal and empirical progress in this period led to a radical transfer of structure from particular grammars to UG. The study of general constraints on rules (Ross 1967, Emonds 1976, Chomsky 1973) and representations (Chomsky 1976) made possible an extreme simplification and impoverishment of the formal tools to be postulated for particular grammars without a loss of empirical adequacy. Many properties arbitrarily expressed by particular rule systems in previous approaches could be advantageously restated as principles of a more richly articulated UG. UG then ceased to be an abstract grammatical metatheory and became an integral component of particular grammars, directly expressing the fundamental cross-linguistic uniformity. Language diversity could be expressed by assuming in the system of UG principles

a finite number of parameters, mostly binary choice points expressing the irreducible differences between grammatical systems. So, a particular grammar could be seen as a particular instantiation of UG under a specific set of parametric values.

The "principles and parameters" (P&P) approach, first developed in a systematic form in Chomsky's seminars at the Scuola Normale Superiore of Pisa in the spring 1979, introduced some radical innovations in the study of language (Chomsky 1981). On the one hand, it broke with a consolidated tradition, fully inherited by early generative grammar from more traditional approaches, which looked at individual languages as characterized by specific rule systems (obeying certain general guidelines): within P&P, an individual language is defined by the general principles that characterize the other human languages plus a finite, possibly quite small, set of specific parametric values. On the other hand, the model of language acquisition was correspondingly simplified in a radical manner: within P&P, the task of the language learner, as far as the computational component of his linguistic ability is concerned, is the task of fixing the parameters of UG on the basis of experience. No complex induction of specific rule system is to be postulated, as there is no specific rule system to be figured out in the first place. The task of the language learner simply is to select, among a predetermined set of parametric values, the values that are supported by his linguistic experience. There is no complex transfer of structure (or instruction) from experience to the internalized system; the basic contribution of experience (apart from the very fact of somehow "turning on" the language faculty) is to motivate selection of certain options over other options, within the range of possibilities generated by the mind.

## 5. SOME TOPICS IN THE THEORY OF PARAMETERS.

The introduction of parametric theory in the late seventies had the immediate effect of triggering a renewed interest for comparative work. The new formal tool quickly turned out to be well adapted to concisely express similarities and differences between closely related systems, by putting a strong emphasis on linguistic uniformity, but at the same time allowing a precise characterization of language variation, its patterns and limits.

After some initial work focused on the parametrization of locality principles (Subjacency), the empirical research turned to the study of major domains of cross-linguistic variation, such as the language-particular licensing of null pronouns and word order phenomena. The hope was raised that certain patterns of variation uncovered by typological studies could be amenable to a deeper explanation in terms of parametric theory. If parameters express the primitive differences between grammatical systems, it is reasonable to conjecture that the irreducible bifurcation induced by a binary parameter will interact with other elements of the tightly deductive system of UG; as a consequence of that interaction, a single parametric difference will be causally related, through a possibly complex deductive chain, to other observable differences between two grammatical systems.

Much attention was initially devoted to the application of this mode of explanation to the null subject pattern. I will then focus on this class of phenomena for an illustration essentially based on Rizzi (1982), but many others could be chosen. Some languages, such as Italian, allow phonetically null pronominal subjects in tensed clauses with both referential and non-referential interpretation; other languages like French do not have this option and require the expression of an overt (referential or expletive) pronoun:

(10)a \_\_\_\_ ha detto la verità  
 b \_\_\_\_ pioveva da tre giorni  
 c \_\_\_\_ è probabile che...

(11)a \* \_\_\_\_ a dit la vérité  
 b \* \_\_\_\_ pleuvait depuis trois jours  
 c \* \_\_\_\_ est probable que...

Italian allows free (untriggered) inversion of the subject, while French does not:

(12) \_\_\_\_ te lo dirà Gianni  
 (13) \* \_\_\_\_ te le dira Jean

Italian allows free subject extraction from an embedded clause (across an overt complementizer), French does not:

(14) Chi pensi che te lo dirà?  
 (15) \* Qui penses-tu que te le dira?

This is not an isolated pattern: the other Romance languages seem, by and large, to pattern with Italian, as does Icelandic (modulo its Verb Second properties, and except for the non-availability of referential null subjects) within Germanic, as well as many other null subject languages; English and Continental Scandinavian (Platzack 1987) pattern with French, as do other non-null subject languages outside the Indo-European family, such as the Kwa languages.

The explanation of this pattern presupposes certain invariable UG principles: the Extended Projection Principle, requiring that all clauses have a structurally represented subject position, and the Empty Category Principle, requiring that traces be limited to occur in certain special, properly governed environments (proper government can be assimilated, for our purposes, to lexical government); it also presupposes a parameter, the statement of a minimal difference between the two language types, which can be expressed by assuming that the Italian language type disposes of a phonetically null pronominal subject pro (whose licensing is in turn determined by certain morphological properties of the verbal inflection) and the French type does not (at least, it does not have a null pronoun licensed in the same environments and under the same formal conditions). (10)-(11) is explained at once: by the Extended Projection Principle, the mind postulates a structural subject position in each clause, even if no referential semantic role is lexically assigned by the predicate to the subject (as in (b-c)): (10) is fine with a mental representation in which pro fills the subject position; as French does not have a well-formed null pronoun for this environment, (11) is ill-formed. The same analysis straightforwardly extends to (12)-(13): a French-type language does not have a null filler to plug in the structural subject position when the overt lexical subject is left VP-internally. As for (15), traces, phonetically null copies of displaced elements, are disallowed in non-properly-governed position by the Empty Category Principle, which rules out representation (15'):

(15') \* Qui penses-tu que <qui> te le dira

On the other hand, Italian allows free subject inversion, so that (14) admits a representation in which the subject trace is VP-internal (postverbal), i.e., in a legitimate position for a trace, the preverbal subject position being plugged in by *pro* as in any other subject inversion sentence:

(14') Chi pensi che pro te lo dirà <chi>

The whole pattern is then fundamentally reduced to a unique primitive difference, the different setting of the null subject parameter.

This mode of explanation was widely explored starting from the late seventies. A host of other properties plausibly related to the null subject parameter was taken into account, ranging from the existence of quirky subjects, agreement patterns in inverse copular constructions and other kinds of inversion, the existence and range of definiteness effects, etc. A lively debate was focused on the attempt to give precise content to the intuitive observation that the licensing of null subjects is restricted to languages with a rich morphological specification of agreement, and the possibility of integrating within a morphologically based generalization the null subject properties of East Asian languages (Huang 1984). The licensing conditions on null pronominal subjects were also expressed within wider attempts to address licensing conditions on other types of null pronominal elements, and in terms of more global approaches to the formal licensing and identification of the different kinds of null elements.

A prominent aspect of cross-linguistic variation involves word order differences, a domain which has been extensively studied in terms of the principles and parameters approach. Consider for instance the (at least) four positions that a verb can occupy in French, depending on its morphological form and certain global properties of the structure (the three positions not occupied by the verb in a specific example being designated by X):

- (16)a X ne X pas X complètement comprendre la théorie (c'est décevant)
- b X ne X pas comprendre complètement X la théorie (c'est décevant)
- c X il ne comprend pas X complètement X la théorie
- d Ne comprend-il X pas X complètement X la théorie?

Under the influential research trend established by Jean-Yves Pollock's theory of verb movement (Pollock 1989), all these cases are reducible to a unique underlying structure, with the lexical verb VP- internal and adjacent to the direct object it selects, as in (16)a, plus a general process of head to head movement, which raises the verb to a higher functional head depending on its morphological shape and other properties of the structure:

(17) C° il ne AGR° pas T° complètement comprend- la théorie

So, a non-finite verb may remain in the position of head of the VP, as in (16)a, or optionally move to a functional head expressing tense (T°) higher than certain adverbs but lower than negation, as in (16)b; a finite verb must raise to an agreement head (AGR°) higher than negation to pick up agreement morphology, as in (16)c (I follow here the ordering argued for in Belletti (1990)); in questions, the verb continues its trip to the next higher functional head, the complementizer (C°), to fulfill certain construction-specific well-formedness requirements,

as in (16)d. Different languages exploit the head movement mechanism in different ways: some never raise the lexical verb out of the VP (English), others raise finite and non-finite verbs on a par to higher functional heads (Italian), others systematically exploit the verb movement possibility to C° in a wider range of cases (Verb Second languages, some VSO languages), etc. The patterns are many, varying across constructions and languages, but they are all reducible to extremely elementary computational mechanisms and parameters: a phrase structure consisting of lexical and functional heads and their phrasal projections, head to head movement (also covering different types of incorporation in Mark Baker's (1988) sense), certain parametrized principles determining the (partly language-specific) morphosyntactic conditions triggering head movement.

A wealth of empirical results made possible by the parametric approach over fifteen years or so gives us a rather precise idea of the domains of UG allowing parametrization: the most plausible and best studied parameters seem to always involve properties of heads, the most crucial choice points for the computational system possibly being restricted to heads of the functional lexicon. So, the language learner acquires the elements of the linguistic computation, the set of items of the substantive and functional lexicon (nouns, verbs, adjectives, etc.; morphemes expressing tense, agreement, case, definiteness, etc.), in part under the guidance of universal principles, in part by fixing some parameters associated to specific items on the basis of experience. Such items then enter the computational component as heads, each projecting a phrase which combines with other phrases on the basis of certain elementary principles of phrasal construction. The computational component is unique across languages, except for the parameters associated with individual heads.

In recent years, an important attempt has been made to unify the expression of different word order parameters under the format of Chomsky's (1995) Minimalist Program. Under this approach, syntactic movement is a last resort option, severely constrained by economy principles and triggered by the satisfaction of (concrete or abstract) morphological requirements. Then, the parametrization determining movement is morphological in nature, and expressible through a system of checking of morphological features. For instance, movement of the French inflected verb to the AGR head in (16)c is motivated by the need of checking the morphological features of person and number agreement on the verb, movement of the subject NP to the specifier of the AGR head is motivated by the need of checking its Case features, in Vergnaud's (1982) sense, and so on. Putting this approach together with Richard Kayne's principled restatement of the universal base hypothesis (under his Linear Correspondence Axiom: Kayne 1994), the tempting conclusion suggests itself that all the basic word order variation may follow from a small set of morphological parameters expressed within the straightforward format of feature checking.

These recent proposals have raised empirical and conceptual questions. Is the satisfaction of morphological properties the only causal factor of movement, or should this mode of explanation be restricted to A(rgument) movement? Are movements in the A' (or operator) system triggered by other kinds of interface conditions such as the creation of operator-variable structure, the satisfaction of A'-criteria, and other properties of Logical Form (May 1985).? If so, how is the relevant parametrization to be expressed? Can the feature checking technology be adapted to cover cases of cross-linguistic variation not reducible to word order, such as the licensing of different kinds of null elements? These and many other related questions have triggered an important debate, and still are very much on focus in the current theoretical discussions. Whatever firm conclusions theoreticians may ultimately agree on, it is

clear that in the course of the nineties, partly under the impulse of the Minimalist Program, definite progress has been made on the identification of what parts of UG are parametrized, as well as on the optimal format for parameters.

## 6. BROADENING THE EMPIRICAL BASIS.

In these final remarks, I will touch upon other domains of inquiry which have contributed to the research program presented above, and/or which are likely to provide a substantive enrichment of its empirical basis in the near future.

The study of closely related grammatical systems offers a very firm testing ground for hypotheses on parametrisation and on the properties governed by each parametric value through deductive steps: by studying coherent microsystems consisting of many closely related varieties, researchers can hope to achieve a better control of possible factors interfering with the basic patterns under investigation. So, it is not surprising that the parametric approach has raised much new theoretical interest for dialectological studies, with particular emphasis on Romance and Germanic dialectology, domains in which literally hundreds of fairly well-defined varieties can be brought to bear on the comparative questions under investigation. Among the positive fallouts of this trend is the fact that dialectological studies, traditionally focused on morphophonology and the lexicon, are progressively filling the gap in the domain of syntax, both synchronic and diachronic.

More generally, language change offers a crucial testing ground for parametric analyses, providing critical natural experiments for the clustering of variable properties. Much work has been devoted to the points of "catastrophic" change in the history of English, French, Scandinavian and other European languages and dialectal varieties (e.g., Lightfoot 1988, Roberts 1993), sometimes with illuminating results, e.g., on the concomitance between the impoverishment of certain morphological paradigms and the loss of syntactic options (verb movement to a functional head, the licensing of null subjects, etc.) in ways predicted by parametric analyses. The comparative study of Creole languages and creolisation processes has also been brought to bear on the theory of parameters, in an effort to understand certain striking similarities between historically unrelated Creole languages through markedness assumptions in parameter setting.

Another area of research which is quickly growing has to do with the very nature of structural representations. The last ten years of research have shown that syntactic representations are much more intricate than we had thought. Each lexical head and phrase is completed by a very rich functional structure, a set of functional heads projecting their own phrases. Through the discovery of such functional structure it has been possible to integrate and partially unify very different kinds of structural and interpretive information, much as in the French example discussed before: syntactic information (e.g., by providing enough positions to integrate different sorts of adverbial modifiers, expressing ordering constraints in principled manners, etc.: Cinque 1997), morphological information (for instance, by relating the position of the verb to its morphological shape), semantic information (by providing positions responsible for interpretive properties such as tense, aspect, definiteness, specificity, etc.). Configurations are then rich and intricate, and yet extremely simple, as the building brick is always the same, the minimal phrase structure projected by each lexical or functional head, and combining with other such syntactic atoms through a very elementary recursive procedure. After a period of very fast progress in the discovery of the fine configurational properties of the IP, CP, DP, etc,

we now start to see the limits of the size of fine-grained structural representations, and we have a clearer and broader picture of the patterns of organization which determine the shape, interpretation and morphological make-up of syntactic structures. The time seems to be ripe to attempt to draw very refined maps of syntactic configurations aiming at exhaustiveness. Large projects of structural cartography are likely to flourish in the years to come.

Although language acquisition has been among the central empirical problems in the theoretical agenda over the last 40 years or so, till fairly recently theoretical linguists did not devote much attention to the study of language development; the study of language acquisition was primarily addressed in the abstract terms of learnability theory abstracting away, among other things, from the actual time course of the acquisition process. From the mid eighties on, things have significantly changed. Under the initial impulse provided in work by Hyams, Roeper, Wexler and others, a number of theoretical linguists have started to seriously take into account developmental data and discover or rediscover the theoretical relevance of a rich descriptive literature in developmental psycholinguistics. *Prima facie*, development offers cases of both continuity and discontinuity in the acquisition of linguistic knowledge. On the one hand, as was mentioned earlier, certain UG principles appear to be operative as soon as we are able to test them, strongly supporting the view that early grammars are UG constrained systems, cast in the same mold as adult grammars. On the other hand, children around the age of two appear to possess grammatical systems which diverge in a number of respects from the target languages, in ways that are amenable to precise grammatical analysis. After over ten years of intensive research in this area, the interplay of continuity and discontinuity has been brought to light in much descriptive and theoretical work, with clear benefits for both theoretical linguists and developmental psycholinguists. From the viewpoint of the linguist, early grammars seem to offer variations on the theme of UG constrained systems which are radical enough to allow us to see things that are not immediately accessible in adult systems (as much recent literature on the so called root infinitive construction shows, for instance), thus providing new kinds of natural experiments and significantly expanding the empirical basis of our models. Conversely, a theory guided, UG based approach has proved to be crucial for drawing a rational chronological map of language development, a central goal for developmental psychology, and an essential tool for topics as diverse as the comparison between first and second language acquisition, and the study of developmental pathologies.

Needless to say, many questions remain open. We want to know how and with what time course parameter fixation takes place in development, if parameters are preset on some initial values determined by general principles of markedness, if parameter refixation is possible and gives rise to observable developmental effects. We want to know what role maturation has in language development, if certain UG principles become operative following an endogenous maturational schedule, if and how linguistic maturation can be related to other measurable dimensions of biological growth, if and how developmental pathologies specific to language, which appear to affect highly specialized elements of the system computing linguistic representations, can be interpreted as involving a perturbation of the maturational schedule. These questions are very much the focus of collaborative projects involving theoretical linguists and psychologists. There are concrete hopes that the study of these and other related questions in structured cross-disciplinary research projects will determine a substantial enrichment of the empirical and conceptual basis for the study of the human cognitive capacity for language.

## REFERENCES

Baker, M. (1988) *Incorporation*, Chicago University Press, Chicago.

Belletti, A. (1990) *Generalized Verb Movement*, Rosenberg & Sellier, Turin.

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

Chomsky, N. (1959) "A Review of B.F. Skinner's *Verbal Behavior*", *Language*, 35, 1, 26-58.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*, MIT Press, Cambridge, Mass.

Chomsky, N. (1973) "Conditions on transformations", in S. Anderson & P. Kiparski, A *Festschrift for Morris Halle*, Holt, Rinehart & Winston, New York.

Chomsky, N. (1975) *Reflections on Language*, Pantheon Books, New York.

Chomsky, N. (1976) "Conditions on Rules of Grammar", *Linguistic Analysis*, 2.4.

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

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

Chomsky, N. & M. Halle (1968) *The Sound Pattern of English*, Harper & Row, New York.

Cinque, G. (1997) *Adverbs and Functional Heads: A Cross-linguistic Perspective*, to be published by Oxford University Press, Oxford.

Crain, S. (1991) "Language acquisition in the absence of experience", *Brain and Behavioral Sciences*.

Emonds, J. (1976) *A Transformational Approach to English Syntax*, Academic Press, New York.

Fodor, J. & J. Katz (1964) *The Structure of Language*, Prentice-Hall, Englewood Cliffs.

Fodor, J. (1983) *The Modularity of Mind*, MIT Press, Cambridge, Mass.

Huang, J. (1984) "On the distribution and reference of empty pronouns", *Linguistic Inquiry*, 15, 531-574.

Kayne, R. (1975) *French Syntax*, MIT Press, Cambridge, Mass.

Kayne, R. (1983) *Connectedness and Binary Branching*, Foris Publications, Dordrecht.

Kayne, R. (1994) *The Antisymmetry of Syntax*, MIT Press, Cambridge, Mass.

Lasnik, H. (1976) "Remarks on Coreference", *Linguistic Analysis*, 2.1.

Lightfoot, D. (1988) "The child's trigger experience: Degree 0 learnability", *Behavioral and Brain Sciences*, 16.2.

May, R. (1985) *Logical Form*, MIT Press, Cambridge, Mass.

Mehler, J. & E. Dupoux (1992) *Naître humain*, Editions Odile Jacob, Paris.

Perlmutter, D., ed. (1983) *Studies in Relational Grammar 1*, The University of Chicago Press, Chicago.

Pinker, S. (1984) *Language Learnability and Language Development*, Harvard University Press, Cambridge, Mass.

Platzack, Ch. (1987) "The Scandinavian languages and the null subject parameter", *Natural Language and Linguistic Theory*, 5.3, 377-402.

Pollock, J.-Y. (1989) "Verb Movement, UG and the structure of IP", *Linguistic Inquiry*, 20, 365-424.

Reinhart, T. (1976) *The Syntactic Domain of Anaphora*, PhD Dissertation, MIT.

Rizzi, L. (1982) *Issues in Italian Syntax*, Foris Publications, Dordrecht.

Roberts, I. (1993) *Verbs and Diachronic Syntax*, Kluwer, Dordrecht.

Ross, J. R. (1967) *Constraints on Variables in Syntax*, PhD Dissertation, MIT.

Saussure, Ferdinand de (1985) *Cours de linguistique générale*, Payot, Paris.

Vergnaud (1982) *Dépendances et niveaux de représentation en syntaxe*, Thèse de doctorat d'état, Université de Paris VII.

**Wexler, K. & P. Culicover (1980) Formal Principles of Language Acquisition, MIT Press,  
Cambridge, Mass.12**