

On Nature and Language

NOAM CHOMSKY

with an essay on
“The Secular Priesthood and the
Perils of Democracy”

Edited by
ADRIANA BELLETTI AND
LUIGI RIZZI



CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS

The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa
<http://www.cambridge.org>

© Noam Chomsky, Adriana Belletti, Luigi Rizzi 2002

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2002
Reprinted 2003

Printed in the United Kingdom at the University Press, Cambridge

Typeface *Quadraat* 9.5/14 pt System \LaTeX 2_ε [TB]

A catalogue record for this book is available from the British Library

ISBN 0 521 81548 7 hardback
ISBN 0 521 01624 X paperback

Contents

Preface vii

- 1 Editors' introduction: some concepts and issues in linguistic theory 1
- 2 Perspectives on language and mind 45
- 3 Language and the brain 61
- 4 An interview on minimalism 92
- 5 The secular priesthood and the perils of democracy 162

Notes 187

References to chapters 1–4 191

Index 201

Chapter 1

Editors' introduction: some concepts and issues in linguistic theory

1 The study of language in a biological setting

Dominant linguistics paradigms in the first half of the twentieth century had centered their attention on Saussurean “*Lange*,” a social object of which individual speakers have only a partial mastery. Ever since the 1950s, generative grammar shifted the focus of linguistic research onto the systems of linguistic knowledge possessed by individual speakers, and onto the “*Language Faculty*,” the species-specific capacity to master and use a natural language (Chomsky 1959). In this perspective, language is a natural object, a component of the human mind, physically represented in the brain and part of the biological endowment of the species. Within such guidelines, linguistics is part of individual psychology and of the cognitive sciences; its ultimate aim is to characterize a central component of human nature, defined in a biological setting.

The idea of focusing on the *Language Faculty* was not new; it had its roots in the classical rationalist perspective of studying language as a “*mirror of the mind*,” as a domain offering a privileged access to the study of human cognition. In order to stress such roots, Chomsky

On nature and language

refers to the change of perspective in the 1950s as “the second cognitive revolution,” thus paying a tribute to the innovative ideas on language and mind in the philosophy of the seventeenth to early nineteenth centuries, with particular reference to the Cartesian tradition. What is new in the “second cognitive revolution” is that language is studied for the first time, in the second half of the twentieth century, with precise formal models capable of capturing certain fundamental facts about human language.

A very basic fact of language is that speakers are constantly confronted with expressions that they have never encountered in their previous linguistic experience, and that they can nevertheless produce and understand with no effort. In fact, normal linguistic capacities range over unbounded domains: every speaker can produce and understand an unbounded number of linguistic expressions in normal language use. This remarkable capacity, sometimes referred to as a crucial component of the “creativity” of ordinary language use, had been noticed at least ever since the first cognitive revolution and had been regarded as a crucial component of human nature. Nevertheless, it had remained fundamentally unexplained in the classical reflection on language. For instance, we find revealing oscillations in Ferdinand de Saussure’s *Cours* on this topic. On the one hand, the *Cours* bluntly states that “la phrase, le type par excellence de syntagme . . . appartient à la parole, non à la langue” (p. 172) [the sentence, the type of phrase par excellence, belongs to *parole*, not to *langue*], and immediately after this passage, the text refers back to the definition of *parole* as “un acte individuel de volonté et d’intelligence . . . [which includes] les combinaisons par lesquelles le sujet parlant utilise le code de la langue en vue d’exprimer sa pensée personnelle . . .” (p. 31) [an individual act of will and intelligence . . . which includes the combinations by which the speaking subject utilizes the code of *langue* in view of expressing

his personal thought]. The freedom of the combinations of elements which characterizes a sentence is “le propre de la parole.” On the other hand, “il faut attribuer à la langue, non à la parole, tous les types de syntagmes construits sur des formes régulières . . . , des groupes de mots construits sur des patrons réguliers, des combinaisons [which] répondent à des types généraux” [it is necessary to attribute to *langue*, not to *parole*, all the types of phrases built on regular forms . . . , groups of words built on regular patterns, combinations which correspond to general types](p. 173). The *Cours*'s conclusion then seems to be that syntax is half way in between *langue* and *parole*: “Mais il faut reconnaître que dans le domaine du syntagme il n'y a pas de limite tranchée entre le fait de langue, marqué de l'usage collectif, et le fait de parole, qui dépend de la liberté individuelle” (p. 173) [but it is necessary to recognize that in the domain of the phrase there is no sharp limit between the facts of *langue*, marked by collective usage, and the facts of *parole*, which depend on individual freedom]. The source of the oscillation is clear: on the one hand, the regular character of syntax is evident; on the other hand, the theoretical linguist at the beginning of the twentieth century does not have at his disposal a precise device to express the astonishing variety of “regular patterns” that natural language syntax allows. See also Graffi (1991: 212–213) for a discussion of this point.

The critical formal contribution of early generative grammar was to show that the regularity and unboundedness of natural language syntax were expressible by precise grammatical models endowed with recursive procedures. Knowing a language amounts to tacitly possessing a recursive generative procedure. When we speak we freely select a structure generated by our recursive procedure and which accords with our communicative intentions; a particular selection in a specific discourse situation is a free act of *parole* in Saussure's sense, but the underlying procedure which specifies the possible “regular patterns”

is strictly rule-governed. Over the last fifty years, the technical characterization of the recursive property of natural language syntax has considerably evolved, from the assumption of “generalized transformations” forming complex constructions step by step beginning with those underlying the simplest sentences (Chomsky 1957), to recursive phrase structure systems (Katz and Postal 1964, Chomsky 1965) capable of producing deep structures of unbounded length, to a recursive X-bar theory (Chomsky 1970, Jackendoff 1977), to the minimalist idea that the basic syntactic operation, “merge,” recursively strings together two elements forming a third element which is the projection of one of its two constituents (Chomsky 1995a, 2000a). Nevertheless, the fundamental intuition has remained constant: natural languages involve recursive generative functions.

The new models built on the basis of this insight quickly permitted analyses with non-trivial deductive depth and which, thanks to their degree of formal explicitness, could make precise predictions and hence could be submitted to various kinds of empirical testing. Deductive depth of the models and experimental controls of their validity: these are among the basic ingredients of what has been called the “Galilean style,” the style of inquiry that established itself in the natural sciences from the time of Galileo Galilei (see chapters 2 and 4 for further discussion of this notion). Showing that the language faculty is amenable to study within the guidelines of the Galilean style, this is then the essence of the second cognitive revolution in the study of language. Initiated by Chomsky’s contributions in the 1950s, this approach has profoundly influenced the study of language and mind ever since, contributing in a critical manner to the rise of modern cognitive science (see, in addition to the references quoted, and among many other publications, Chomsky’s (1955) doctoral dissertation, published in 1975, Chomsky (1957) and various essays in Fodor and Katz (1964)).

2 Universal Grammar and particular grammars

The modern study of language as a mirror of the mind revolves around a number of basic research questions, two of which have been particularly prominent:

- What is knowledge of language?
- How is it acquired?

The first question turned out to be of critical importance for the program to get started. The first fragments of generative grammar in the 1950s and 1960s showed, on the one hand, that the implicit knowledge of language was amenable to a precise study through models which had their roots in the theory of formal systems, primarily in the theory of recursive functions; on the other hand, they immediately underscored the fact that the intuitive linguistic knowledge that every speaker possesses, and which guides his linguistic behavior, is a system of extraordinary complexity and richness. Every speaker implicitly masters a very detailed and precise system of formal procedures to assemble and interpret linguistic expressions. This system is constantly used, in an automatized and unconscious manner, to produce and understand novel sentences, a normal characteristic of ordinary language use.

The discovery of the richness of the implicit knowledge of language immediately raised the question of acquisition. How can it be that every child succeeds in acquiring such a rich system so early in life, in an apparently unintentional manner, without the need of an explicit teaching? More importantly, the precise study of fragments of adult knowledge of language quickly underscored the existence of “poverty of stimulus” situations: the adult knowledge of language is largely underdetermined by the linguistic data normally available to the child,

which would be consistent with innumerable generalizations over and above the ones that speakers unerringly converge to. Let us consider a simple example to illustrate this point. Speakers of English intuitively know that the pronoun “he” can be understood as referring to John in (1), but not in (2):

- (1) John said that he was happy
- (2) *He said that John was happy

We say that “coreference” between the name and the pronoun is possible in (1), but not in (2) (the star in (2) signals the impossibility of coreference between the underscored elements; the sentence is obviously possible with “he” referring to some other individual mentioned in the previous discourse). It is not a simple matter of linear precedence: there is an unlimited number of English sentences in which the pronoun precedes the name, and still coreference is possible, a property illustrated in the following sentences with subject, object and possessive pronouns:

- (3) When he plays with his children, John is happy
- (4) The people who saw him playing with his children said that John was happy
- (5) His mother said that John was happy

The actual generalization involves a sophisticated structural computation. Let us say that the “domain” of an element A is the phrase which immediately contains A (we also say that A c-commands the elements in its domain: Reinhart (1976)). Let us now indicate the domain of the pronoun by a pair of brackets in (1)–(5):

Editors' introduction

- (6) John said that [he was happy]
- (7) * [He said that John was happy]
- (8) When [he plays with his children], John is happy
- (9) The people who saw [him playing with his children] said that John was happy
- (10) [His mother] said that John was happy

The formal property which singles out (7) is now clear: only in this structure is the name contained in the domain of the pronoun. So, coreference is excluded when the name is in the domain of the pronoun (this is Lasnik's (1976) Principle of Non-coreference). Speakers of English tacitly possess this principle, and apply it automatically to new sentences to evaluate pronominal interpretation. But how do they come to know that this principle holds? Clearly, the relevant information is not explicitly given by the child's carers, who are totally unaware of it. Why don't language learners make the simplest assumption, i.e. that coreference is optional throughout? Or why don't they assume that coreference is ruled by a simple linear principle, rather than by the hierarchical one referring to the notion of domain? Why do all speakers unerringly converge to postulate a structural principle rather than a simpler linear principle, or even no principle at all?

This is one illustration of a pervasive situation in language acquisition. As the experience is too impoverished to motivate the grammatical knowledge that adult speakers invariably possess, we are led to assume that particular pieces of grammatical knowledge develop because of some pressure internal to the cognitive system of the child. A natural hypothesis is that children are born with a "language faculty" (Saussure), an "instinctive tendency" for language (Darwin); this

cognitive capacity must involve, in the first place, receptive resources to separate linguistic signals from the rest of the background noise, and then to build, on the basis of other inner resources activated by a limited and fragmentary linguistic experience, the rich system of linguistic knowledge that every speaker possesses. In the case discussed, an innate procedure determining the possibilities of coreference is plausibly to be postulated, a procedure possibly to be deduced from a general module determining the possibilities of referential dependencies among expressions, as in Chomsky's (1981) Theory of Binding, or from even more general principles applying at the interface between syntax and pragmatics, as in the approach of Reinhart (1983). In fact, no normative, pedagogic or (non-theory-based) descriptive grammar ever reports such facts, which are automatically and unconsciously assumed to hold not only in one's native language, but also in the adult acquisition of a second language. So, the underlying principle, whatever its ultimate nature, appears to be part of the inner background of every speaker.

We can now phrase the problem in the terminology used by the modern study of language and mind. Language acquisition can be seen as the transition from the state of the mind at birth, the initial cognitive state, to the stable state that corresponds to the native knowledge of a natural language. Poverty of stimulus considerations support the view that the initial cognitive state, far from being the *tabula rasa* of empiricist models, is already a richly structured system. The theory of the initial cognitive state is called Universal Grammar; the theory of a particular stable state is a particular grammar. Acquiring the tacit knowledge of French, Italian, Chinese, etc., is then made possible by the component of the mind-brain that is explicitly modeled by Universal Grammar, in interaction with a specific course of linguistic experience. In the terms of comparative linguistics, Universal

Grammar is a theory of linguistic invariance, as it expresses the universal properties of natural languages; in terms of the adopted cognitive perspective, Universal Grammar expresses the biologically necessary universals, the properties that are universal because they are determined by our in-born language faculty, a component of the biological endowment of the species.

As soon as a grammatical property is ascribed to Universal Grammar on the basis of poverty of stimulus considerations, a hypothesis which can be legitimately formulated on the basis of the study of a single language, a comparative verification is immediately invited: we want to know if the property in question indeed holds universally. In the case at issue, we expect no human language to allow coreference in a configuration like (2) (modulo word order and other language specific properties), a conclusion which, to the best of our current knowledge, is correct (Lasnik (1989), Rizzi (1997a) and references quoted there). So, in-depth research on individual languages immediately leads to comparative research, through the logical problem of language acquisition and the notion of Universal Grammar. This approach assumes that the biological endowment for language is constant across the species: we are not specifically predisposed to acquire the language of our biological parents, but to acquire whatever human language is presented to us in childhood. Of course, this is not an a priori truth, but an empirical hypothesis, one which is confirmed by the explanatory success of modern comparative linguistics.

3 Descriptive adequacy and explanatory adequacy

It has been said that language acquisition constitutes “the fundamental empirical problem” of modern linguistic research. In order to underscore the importance of the problem, Chomsky introduced,

in the 1960s, a technical notion of explanation keyed to acquisition (see Chomsky (1964, 1965) for discussion). An analysis is said to meet “descriptive adequacy” when it correctly describes the linguistic facts that adult speakers tacitly know; it is said to meet the higher requirement of “explanatory adequacy” when it also accounts for how such elements of knowledge are acquired. Descriptive adequacy can be achieved by a fragment of a particular grammar which successfully models a fragment of adult linguistic knowledge; explanatory adequacy is achieved when a descriptively adequate fragment of a particular grammar can be shown to be derivable from two ingredients: Universal Grammar with its internal structure, analytic principles, etc., and a certain course of experience, the linguistic facts which are normally available to the child learning the language during the acquisition period. These are the so-called “primary linguistic data,” a limited and individually variable set of utterances whose properties and structural richness can be estimated via corpus studies. If it can be shown that the correct grammar can be derived from UG and a sample of data which can be reasonably assumed to be available to the child, the acquisition process is explained. To go back to our concrete example on coreference, descriptive adequacy would be achieved by a hypothesis correctly capturing the speaker’s intuitive judgments on (1)–(5), say a hypothesis referring to a hierarchical principle rather than a linear principle; explanatory adequacy would be achieved by a hypothesis deriving the correct description of facts from general inborn laws, say Chomsky’s binding principles, or Reinhart’s principles on the syntax–pragmatics interface.

A certain tension arose between the needs of descriptive and explanatory adequacy in the 1960s and 1970s, as the two goals pushed research in opposite directions. On the one hand, the needs of descriptive adequacy seemed to require a constant enrichment of the

descriptive tools: with the progressive broadening of the empirical basis, the discovery of new phenomena in natural languages naturally led researchers to postulate new analytic tools to provide adequate descriptions. For instance, when the research program was extended for the first time to the Romance languages, the attempts to analyze certain verbal constructions led to the postulation of new formal rules (causative formation transformations and more radically innovative formal devices such as restructuring, reanalysis, clause union, etc.: Kayne 1975, Rizzi 1976, Aissen and Perlmutter 1976), which seemed to require a broadening of the rule inventory allowed by Universal Grammar. Similarly, and more radically, the first attempts to analyze languages with freer word order properties led to the postulation of different principles of phrasal organization, as in much work on so-called "non-configurational" languages by Ken Hale, his collaborators and many other researchers (Hale 1978). On the other hand, the very nature of explanatory adequacy, as it is technically defined, requires a maximum of restrictiveness, and the postulation of a strong cross-linguistic uniformity: only if Universal Grammar offers relatively few analytic options for any given set of data is the task of learning a language a feasible one in the empirical conditions of time and access to the data available to the child. It was clear all along that only a restrictive approach to Universal Grammar would make explanatory adequacy concretely attainable (see chapter 4 and Chomsky (2001b) on the status of explanatory adequacy within the Minimalist Program).

4 Principles and parameters of Universal Grammar

An approach able to resolve this tension emerged in the late 1970s. It was based on the idea that Universal Grammar is a system of principles and parameters. This approach was fully developed for the first time in

informal seminars that Chomsky gave at the Scuola Normale Superiore of Pisa in the Spring semester of 1979, which gave rise to a series of lectures presented immediately after the GLOW Conference in April 1979, the Pisa Lectures. The approach was refined in Chomsky's Fall 1979 course at MIT, and then presented in a comprehensive monograph as Chomsky (1981).

Previous versions of generative grammar had adopted the view, inherited from traditional grammatical descriptions, that particular grammars are systems of language-specific rules. Within this approach, there are phrase structure rules and transformational rules specific to each language (the phrase structure rule for the VP is different in Italian and Japanese, the transformational rule of causative formation is different in English and French, etc.). Universal Grammar was assumed to function as a kind of grammatical metatheory, by defining the general format which specific rule systems are required to adhere to, as well as general constraints on rule application. The role of the language learner was to induce a specific rule system on the basis of experience and within the limits and guidelines defined by UG. How this induction process could actually function remained largely mysterious, though.

The perspective changed radically some twenty years ago. In the second half of the 1970s, some concrete questions of comparative syntax had motivated the proposal that some UG principles could be parametrized, hence function in slightly different ways in different languages. The first concrete case studied in these terms was the fact that certain island constraints appear to be slightly more liberal in certain varieties than in others: for instance, extracting a relative pronoun from an indirect question sounds quite acceptable in Italian (Rizzi 1978), less so in other languages and varieties: it is excluded in German, and marginal at variable degrees in different varieties of

English (see Grimshaw (1986) for discussion of the latter case; on French see Sportiche (1981)):

- (11) Ecco un incarico [_S che [_S non so proprio [_S a chi
[_S potremmo affidare ___]]]]
Here is a task that I really don't know to whom we could
entrust

- (12) * Das ist eine Aufgabe, [_S die [_S ich wirklich nicht weiss
[_S wem [_S wir ___ anvertrauen könnten]]]]
Here is a task that I really don't know to whom we could
entrust

It is not the case that Italian allows extraction in an unconstrained way: for instance, if extraction takes place from an indirect question which is in turn embedded within an indirect question, the acceptability strongly degrades:

- (13) * Ecco un incarico [_S che [_S non so proprio [_S a chi
[_S si domandino [_S se [_S potremmo affidare ___]]]]]]
Here is a task that I really don't know to whom they
wonder if we could entrust

The suggestion was made that individual languages could differ slightly in the choice of the clausal category counting as bounding node, or barrier for movement. Assume that the relevant principle, Subjacency, allows movement to cross one barrier at most; then, if the language selects S' as clausal barrier, movement of this kind will be possible, with only the lowest S' crossed; if the language selects S, movement will cross two barriers, thus giving rise to a violation of subjacency. Even if the language selects S', movement from a double Wh island will be barred, whence the contrast (11)–(13) (if a language were to select both S and S' as bounding node, it was observed, then

even movement out of a declarative would be barred, as seems to be the case in certain varieties of German and in Russian: see the discussion in Freidin (1988)).

In retrospect, this first example was far from an ideal case of parameter: the facts are subtle, complex and variable across varieties and idiolects, etc. Nevertheless, the important thing is that it quickly became apparent that the concept of parameter could be extended to other more prominent cases of syntactic variation, and that in fact the whole cross-linguistic variation in syntax could be addressed in these terms, thus doing away entirely with the notion of a language-specific rule system. Particular grammars could be conceived of as direct instantiations of Universal Grammar, under particular sets of parametric values (see Chomsky (1981) and, among many other publications, different papers collected in Kayne (1984, 2001), Rizzi (1982, 2000)).

Within the new approach, Universal Grammar is not just a grammatical metatheory, and becomes an integral component of particular grammars. In particular, UG is a system of universal principles, some of which contain parameters, choice points which can be fixed in one of a limited number of ways. A particular grammar then is immediately derived from UG by fixing the parameters in a certain way: Italian, French, Chinese, etc. are direct expressions of UG under particular, and distinct, sets of parametric values. No language-specific rule system is postulated: structures are directly computed by UG principles, under particular parametric choices. At the same time, the notion of a construction-specific rule dissolves. Take for instance the passive, in a sense the prototypical case of a construction-specific rule. The passive construction is decomposed into more elementary operations, each of which is also found elsewhere. On the one hand, the passive morphology intercepts the assignment of the external Thematic Role (Agent, in the example given below) to the subject

position and optionally diverts it to the *by* phrase, as in the underlying representation (14a); by dethematizing the subject, this process also prevents Case assignment to the object (via so-called Burzio's generalization, see Burzio (1986)); then, the object left without a Case moves to subject position, as in (14b) (on Case Theory and the relevance of Case to trigger movement, see below):

- (14) a. ___ was washed the car (by Bill)
b. The car was washed ___ (by Bill)

None of these processes is specific to the passive: the interception of the external thematic role and optional diversion to a *by* phrase is also found, for instance, in one of the causative constructions in Romance (with Case assigned to the object by the complex predicate *faire+V* in (15)), movement of the object to a non-thematic subject position is also found with unaccusative verbs, verbs which do not assign a thematic role to the subject as a lexical property and are morphologically marked in some Romance and Germanic languages by the selection of auxiliary *be*, as in (16) in French (Perlmutter 1978, Burzio 1986):

- (15) Jean a fait laver la voiture (par Pierre)
Jean made wash the car (by Pierre)
- (16) Jean est parti ___
Jean has left

So, the “passive construction” dissolves into more elementary constituents: a piece of morphology, an operation on thematic grids, movement. The elementary constituents have a certain degree of modular autonomy, and can recombine to give rise to different constructions under language-specific parametric values.

A crucial contribution of parametric models is that they provided an entirely new way of looking at language acquisition. Acquiring a

language amounts, in terms of such models, to fixing the parameters of UG on the basis of experience. The child interprets the incoming linguistic data through the analytic devices provided by Universal Grammar, and fixes the parameters of the system on the basis of the analyzed data, his linguistic experience. Acquiring a language thus means selecting, among the options generated by the mind, those which match experience, and discarding the other options. So, acquiring an element of linguistic knowledge amounts to discarding the other possibilities offered a priori by the mind; learning is then achieved “by forgetting,” a maxim adopted by Mehler and Dupoux (1992) in connection with the acquisition of phonological systems: acquiring the phonetic distinctions used in one’s language amounts to forgetting the others, in the inventory available a priori to the child’s mind, so that at birth every child is sensitive to the distinction between /l/ and /r/, or /t/ and /t./ (dental vs. retroflex), but after a few months the child learning Japanese will have “forgotten” the /l/ vs. /r/ distinction, and the child learning English will have “forgotten” the /t/ vs. /t./ distinction, etc., because they will have kept the distinctions used by the language they are exposed to and discarded the others. Under the parametric view, “learning by forgetting” seems to be appropriate for the acquisition of syntactic knowledge as well.

The Principles and Parameters approach offered a new way of addressing the logical problem of language acquisition, in terms which abstract away from the actual time course of the acquisition process (see Lightfoot (1989) and references discussed there). But it also generated a burst of work on language development: how is parameter fixation actually done by the child in a concrete time course? Can it give rise to observable developmental patterns, e.g. with the resetting of some parameters after exposure to a sizable experience, or under

the effect of maturation? Hyams's (1986) approach to subject drop in child English opened a line of inquiry on the theory-conscious study of language development which has fully flourished in the last decade (see, among many other references, the discussion in Friedemann and Rizzi (2000), Rizzi (2000), Wexler (1994, 1998) and the references quoted there; on the connections between language acquisition, language change and creolization in terms of the parametric approach, see Degraff (1999)).

5 Parametric models and linguistic uniformity

The development of parametric models was made possible by an important empirical discovery: human languages are much more uniform than was previously thought. Let us illustrate this point through some simple examples.

5.1 Overt vs. covert movement

Consider first question formation. Human languages generally take one of two options to form constituent questions. The option taken by English (Italian, Hungarian, etc.) consists of moving the interrogative phrase (*who*, etc.) to the front, to a position in the left periphery of the clause; the option taken by Chinese (Japanese, Turkish, etc.) consists of leaving the interrogative phrase in situ, in the clause-internal argument position in which it is interpreted (e.g. in (18) as the internal argument of *love*):

(17) Who did you meet ____?

(18) Ni xihuan shei?
You love who?

Colloquial French allows both options in main clauses:

- (19) a. Tu as vu qui?
You have seen who?
- b. Qui as-tu vu ___?
Who have you seen?

The very existence of only two major options is already an indication of uniformity. In no known language, for instance, is the question formed by moving the interrogative phrase to a lower structural position in the syntactic tree, say from the main clause to an embedded complementizer position. Moreover, there are good reasons to believe that the uniformity is even deeper. At Logical Form, an abstract level of mental representation at the interface with thought systems (on which see May (1985), Hornstein (1984)), movement seems always to be required, also in Chinese and colloquial French, giving rise to structures in which the interrogative phrase binds a clause-internal variable:

- (20) For what *x*, you met/saw/love *x*?

Important empirical evidence for the idea that movement applies covertly in these systems was provided by Huang's (1982) observation that certain locality constraints hold uniformly across languages. For instance, an interrogative adverb cannot be extracted from an indirect question in English-type interrogatives, a property related to the operation of a fundamental locality principle, giving rise to violations which are much more severe and linguistically invariable than the extraction cases discussed in connection with (11) and (12):

- (21) * How do you wonder [who solved the problem ___]?

For instance, the equivalent of (21) is also strongly excluded in Italian, a language which rather freely allows extraction of argumental material from indirect questions, as we have seen:

(22) * Come ti domandi [chi ha risolto il problema ___]?

How do you wonder who solved the problem?

The constraint violated in (21) and (22) is, according to Huang's original approach, the Empty Category Principle (ECP), a principle giving rise to stronger and cross-linguistically invariant violations than Subadjacency: in a nutshell, the Wh adverb cannot be connected to the embedded clause across another Wh element; see, among many other references, Lasnik and Saito (1992), Rizzi (1990, 2000, 2001a,b), Cinque (1990), Starke (2001) on the different behavior of argument and adjunct extraction in this environment, and the discussion of locality below.

In parallel with (21) and (22), an interrogative adverb within an indirect question cannot be interpreted as a main question element in Chinese-type languages, Huang showed. The parallel is immediately shown by French: starting from a structure like (23a), a main interrogative bearing on the embedded adverb is excluded, whether the adverb is moved or not (NB these judgments hold with normal stress contour; if the interrogative element *in situ* is heavily stressed the acceptability improves: see Starke (2001) for a discussion of the relevance of the stress contour in these cases):

(23) a. Tu te demandes qui a résolu le problème de cette manière
You wonder who solved the problem in this way

b. * Comment te demandes-tu qui a résolu le problème ___?
How do you wonder who solved the problem?

c. * Tu te demandes qui a résolu le problème comment?
You wonder who solved the problem how?

This is immediately explained if speakers of Chinese, colloquial French, etc. assign Logical Forms like (20) to *in situ* interrogatives through covert movement of the interrogative phrase. The same locality principles apply that are operative in cases like (21) and (22),

barring overt and “mental” movement on a par. So, it appears that, in abstract mental representations, questions are represented uniformly, in a format akin to (20); what varies is whether movement to the front has audible consequences, as in English, or is covert as in Chinese, etc., a difference expressible through a straightforward parametrization (e.g. in the feature system of Chomsky (1995a)). A single locality principle applying on uniform Logical Forms accounts for the ill-formedness of overt extraction in the English and Italian structures and for the absence of main clause interpretation in the Chinese structure, with French instantiating both cases. Analogous arguments for covert Wh movement can be based on the uniform behavior of moved and *in situ* interrogative elements with respect to the possibility of binding a pronoun (Weak Crossover Effects), an extension of the classical argument for covert movement in Chomsky (1977: ch. 1). (See also Pollock and Poletto (2001), who reinterpret certain apparent *in situ* cases as involving leftward movement of the Wh element, followed by “remnant movement” of the rest of the clause to an even higher position, in terms of Kayne’s (1994) approach; and Watanabe (1992), Reinhart (1995), Fox and Nissenbaum (1999) for alternative approaches to covert movement.)

The syntax of questions already looks rather uniform on a superficial analysis, but other aspects of syntax seem to vary considerably across languages at first glance. What the work of recent years consistently shows is that, as soon as the domain is studied in detail and with appropriate theoretical tools, much of the variability dissolves and we are left with a residue of few elementary parameters.

5.2 *Adverbs and functional heads*

One aspect with respect to which natural languages seem to vary a lot has to do with the position of adverbials. For instance, certain low