Metaphysics texts: Rashi
Revised
09 Nov 2026
14:10
KSB

Knowledge of Language:
Its Nature, Origin, and Use

Noam Chomsky

CONVERGENCE

A Series Founded, Planned, and Edited by Ruth Nanda Anshen

PRAEGER

1 Knowledge of Language as a Focus of Inquiry

The study of language has a long and rich history, extending over thousands of years. This study has frequently been understood as an inquiry into the nature of mind and thought on the assumption that "languages are the best mirror of the human mind" (Leibniz). A common conception was that "with respect to its *substance* grammar is one and the same in all languages, though it does vary *accidentally*" (Roger Bacon). The invariant "substance" was often taken to be the mind and its acts; particular languages use various mechanisms—some rooted in human reason, others arbitrary and adventitious—for the expression of thought, which is a constant across languages. One leading eighteenth century rational grammarian defined "general grammar" as a deductive science concerned with "the immutable and general principles of spoken or written language" and their consequences; it is "prior to all languages," because its principles "are the same as those that direct human reason in its intellectual operations" (Beauzée). Thus, "the science of language does not differ at all from the science of thought." "Particular grammar" is not a true "science" in the sense of this rationalist tradition because it is not based solely on universal necessary laws; it is an "art" or technique that shows how given languages realize the general principles of human reason. As John Stuart Mill later expressed the same leading idea, "The principles and rules of grammar are the means by which the forms of language are made to correspond with the universal forms of thought.... The structure of every sentence is a lesson in logic." Others, particularly during the

Romantic period, argued that the nature and content of thought are determined in part by the devices made available for its expression in particular languages. These devices may include contributions of individual genius that affect the "character" of a language, enriching its means of expression and the thoughts expressed without affecting its "form," its sound system and rules of word and sentence formation (Humboldt).

The thesis: to know a language is to be in a state of a distinct faculty of the mind, grown from data too poor to teach it. An old picture, revived as generative grammar (p. 3) and argued from p. 7.

With regard to the acquisition of knowledge, it was widely held that the mind is not "so much to be filled therewith from without, like a vessel, as to be kindled and awaked" (Ralph Cudworth); "The growth of knowledge . . . [rather resembles] . . . the growth of Fruit; however external causes may in some degree cooperate, it is the internal vigour, and virtue of the tree, that must ripen the juices to their just maturity" (James Harris).[1] Applied to language, this essentially Platonistic conception would suggest that knowledge of a particular language grows and matures along a course that is in part intrinsically determined, with modifications reflecting observed usage, rather in the manner of the visual system or other bodily "organs" that develop along a course determined by genetic instructions under the triggering and shaping effects of environmental factors.

this essentially Platonistic conception: *knowledge drawn out of the mind rather than poured in, as Cudworth and Harris say. Language then grows like an organ, triggered by experience, the picture Plato's problem on p. 7 argues for.*

With the exception of the relativism of the Romantics, such ideas were generally regarded with much disapproval in the mainstream of linguistic research by the late nineteenth century and on through the 1950s. In part, this attitude developed under the impact of a rather narrowly construed empiricism and later behaviorist and operationalist doctrine. In part, it resulted from the quite real and impressive successes of historical and descriptive studies conducted within a narrower compass, specifically, the discovery of "sound laws" that provided much understanding of the history of languages and their relationships. In part, it was a natural consequence of the investigation of a much richer variety of languages than were known to earlier scholars, languages that appeared to violate many of the allegedly *a priori* conceptions of the earlier rationalist tradition.[2] After a century of general neglect or obloquy, ideas resembling those of the earlier tradition re-emerged (initially, with virtually no awareness of historical antecedents) in the mid-1950s, with the development of what came to be called "generative grammar"—again, reviving a long-lapsed and largely forgotten tradition.[3]

"generative" means nothing more than "explicit": *a generative grammar states its rules fully enough to fix, for every string, its structure and meaning, with no appeal to the reader's good sense. Nothing is said about how speech is produced.*

The generative grammar of a particular language (where "generative" means nothing more than "explicit") is a theory that is concerned with the form and meaning of expressions of this language. One can imagine many different kinds of approach to such questions, many points of view that might be adopted in dealing with them. Generative grammar limits itself to certain elements of this larger picture. Its standpoint is that of individual psychology. It is concerned with those aspects of form and meaning that are determined by the "language faculty," which is understood to be a particular component of the human mind. The nature of this faculty is the subject matter of a general theory of linguistic structure that aims to discover the framework of principles and elements common to attainable human languages; this theory is now often called "universal grammar" (UG), adapting a traditional term to a new context of inquiry. UG may be regarded as a characterization of the genetically determined language faculty. One may think of this faculty as a "language acquisition device," an innate component of the human mind that yields a particular language through interaction with presented experience, a device that converts experience into a system of knowledge attained: knowledge of one or another language.

"universal grammar" (UG): *the old term meant the grammar common to all languages. Here it names the theory of the innate language faculty, the state the child starts from. On p. 29 he owns it also named that state, S0.*

The thesis answers (i) and (ii) from this standpoint. Its rivals, knowledge as an ability (p. 9) and language as a set of sentences (p. 20), keep the old focus on behaviour and its products.

The study of generative grammar represented a significant shift of focus in the approach to problems of language. Put in the simplest terms, to be elaborated below, the shift of focus was from behavior or the products of behavior to states of the mind/brain that enter into behavior. If one chooses to focus attention on this latter topic, the central concern becomes knowledge of language: its nature, origins, and use.

three basic questions: *the three together are numbered (1), so '(1i)' and '(1iii)' later refer to the first and third of them. The case that (i) and (iii) are distinct questions comes on pp. 9 to 12.*

The three basic questions that arise, then, are these:

(1)
(i) What constitutes knowledge of language?
(ii) How is knowledge of language acquired?
(iii) How is knowledge of language put to use?

a particular generative grammar: *so the answer to (i) is a theory of a state of one person's mind, not a list of sentences and not a description of a community's usage. Chapter 2 draws the contrast as I-language against E-language.*

The answer to the first question is given by a particular generative grammar, a theory concerned with the state of the mind/brain of the person who knows a particular language. The answer to the second is given by a specification of UG along with an account of the ways in which its principles interact with experience to yield a particular language; UG is a theory of the "initial state" of the language faculty, prior to any

linguistic experience. The answer to the third question would be a theory of how the knowledge of language attained enters into the expression of thought and the understanding of presented specimens of language, and derivatively, into communication and other special uses of language.

Nothing is contested yet. What is, a distinct language faculty, becomes an empirical question with a test: are the initial state's properties language-specific? Answered yes on pp. 12 to 13.

So far, this is nothing more than the outline of a research program that takes up classical questions that had been put aside for many years. As just described, it should not be particularly controversial, since it merely expresses an interest in certain problems and offers a preliminary analysis of how they might be confronted, although as is often the case, the initial formulation of a problem may prove to be far-reaching in its implications, and ultimately controversial as it is developed.

Some elements of this picture may appear to be more controversial than they really are. Consider, for example, the idea that there is a language faculty, a component of the mind/brain that yields knowledge of language given presented experience. It is not at issue that humans attain knowledge of English, Japanese, and so forth, while rocks, birds, or apes do not under the same (or indeed any) conditions. There is, then, some property of the mind/brain that differentiates humans from rocks, birds, or apes. Is this a distinct "language faculty" with specific structure and properties, or, as some believe, is it the case that humans acquire language merely by applying generalized learning mechanisms of some sort, perhaps with greater efficiency or scope than other organisms? These are not topics for speculation or *a priori* reasoning but for empirical inquiry, and it is clear enough how to proceed: namely, by facing the questions of (1). We try to determine what is the system of knowledge that has been attained and what properties must be attributed to the initial state of the mind/brain to account for its attainment. Insofar as these properties are language-specific, either individually or in the way they are organized and composed, there is a distinct language faculty.

Is this a distinct "language faculty": *the alternative is the empiricist's: one general learner for language, chess and everything else. The case against it is Plato's problem, pp. 7 to 12, and the Reflections example, pp. 30 to 33.*

Generative grammar is sometimes referred to as a theory, advocated by this or that person. In fact, it is not a theory any more than chemistry is a theory. Generative grammar is a topic, which one may or may not choose to study. Of course, one can adopt a point of view from which chemistry disappears as a discipline (perhaps it is all done by angels with mirrors). In this sense, a decision to study chemistry does stake out a position on

not a theory any more than chemistry: *a topic is a domain of facts, a theory a claim about them. Denying there is anything to study here is a position too, as odd as denying chemistry, so the burden isn't all on Chomsky's side.*

matters of fact. Similarly, one may argue that the topic of generative grammar does not exist, although it is hard to see how to make this position minimally plausible. Within the study of generative grammar there have been many changes and differences of opinion, often reversion to ideas that had been abandoned and were later reconstructed in a different light. Evidently, this is a healthy phenomenon indicating that the discipline is alive, although it is sometimes, oddly, regarded as a serious deficiency, a sign that something is wrong with the basic approach. I will review some of these changes as we proceed.

In the mid-1950s, certain proposals were advanced as to the form that answers to the questions of (1) might take, and a research program was inaugurated to investigate the adequacy of these proposals and to sharpen and apply them. This program was one of the strands that led to the development of the cognitive sciences in the contemporary sense, sharing with other approaches the belief that certain aspects of the mind/brain can be usefully construed on the model of computational systems of rules that form and modify representations, and that are put to use in interpretation and action. From its origins (or with a longer perspective, one might say "its reincarnation") about 30 years ago, the study of generative grammar was undertaken with an eye to gaining some insight into the nature and origins of systems of knowledge, belief, and understanding more broadly, in the hope that these general questions could be illuminated by a detailed investigation of the special case of human language.

computational systems of rules: *cognitive science's working assumption: a mind described as a system that builds and transforms representations by rules, apart from neurons. So a grammar can be a theory of a brain state (p. 23).*

This research program has since been running its course, along a number of different paths. I will be concerned here with only one of these, with the problems it faced and the steps that were taken in an effort to deal with them. During the past 5–6 years, these efforts have converged in a somewhat unexpected way, yielding a rather different conception of the nature of language and its mental representation, one that offers interesting answers to a range of empirical questions and opens a variety of new ones to inquiry while suggesting a rethinking of the character of others. This is what accounts for an unmistakable sense of energy and anticipation—and also uncertainty—which is reminiscent of the period when the study of generative grammar in the modern sense was initiated about 30 years ago.

Some of the work now being done is quite different in character from what had previously been possible as well as considerably broader in empirical scope, and it may be that results of a rather new kind are within reach, or at least within sight. I would like to try to explain why this may be so, beginning with some remarks about goals, achievements, and failures of the past years.

To avoid misunderstanding, I am not speaking here about all of the study of language but rather of generative grammar, and even here I will not attempt anything like a real history of the course of research but rather will give a somewhat idealized picture that is in part clearer in retrospect than it was at the time. Furthermore, what I am describing has represented a minority position throughout, and probably still does, although in my view it is the correct one. A number of different current approaches share properties of the sort discussed here and may be intertranslatable to a considerable extent. I will not consider this important topic here and will also make no effort to survey the range of ideas, often conflicting, that fall within the particular tendency that I will discuss—what is now sometimes called "government-binding (GB) theory."

Two shifts organize the book. The first, from behaviour to states of the mind, is this reading's. The second, from rules to principles with parameters, is only announced, here and on p. 24.

I want to consider, then, two major conceptual shifts, one that inaugurated the contemporary study of generative grammar, and a second, more theory-internal, that is now in process and that offers some new perspectives on traditional problems.[4]

Traditional and structuralist grammar did not deal with the questions of (1), the former because of its implicit reliance on the unanalyzed intelligence of the reader, the latter because of its narrowness of scope. The concerns of traditional and generative grammar are, in a certain sense, complementary: a good traditional or pedagogical grammar provides a full list of exceptions (irregular verbs, etc.), paradigms and examples of regular constructions, and observations at various levels of detail and generality about the form and meaning of expressions. But it does not examine the question of how the reader of the grammar uses such information to attain the knowledge that is used to form and interpret new expressions, or the question of the nature and elements of this knowledge: essentially the questions of (1), above. Without too much exaggeration, one could describe such a grammar as a structured and organized version of the data presented to a child learning a language,

complementary: *the irony: what a good grammar can leave unsaid, since every reader supplies it, is just what generative grammar has to explain. Facts like (2) to (7) are in no grammar book (p. 9).*

with some general commentary and often insightful observations. Generative grammar, in contrast, is concerned primarily with the intelligence of the reader, the principles and procedures brought to bear to attain full knowledge of a language. Structuralist theories, both in the European and American traditions, did concern themselves with analytic procedures for deriving aspects of grammar from data, as in the procedural theories of Nikolay Trubetzkoy, Zellig Harris, Bernard Bloch, and others, but primarily in the areas of phonology and morphology. The procedures suggested were seriously inadequate and in any event could not possibly be understood (and were not intended) to provide an answer to question (1ii), even in the narrower domains where most work was concentrated. Nor was there an effort to determine what was involved in offering a comprehensive account of the knowledge of the speaker/hearer.

Nikolay Trubetzkoy, Zellig Harris, Bernard Bloch: *structuralists of the 1930s and 40s. Trubetzkoy founded Prague-school phonology. Harris, Chomsky's teacher, and Bloch sought mechanical procedures for extracting a grammar from a corpus.*

As soon as these questions were squarely faced, a wide range of new phenomena were discovered, including quite simple ones that had passed unnoticed, and severe problems arose that had previously been ignored or seriously misunderstood. A standard belief 30 years ago was that language acquisition is a case of "overlearning." Language was regarded as a habit system, one that was assumed to be much overdetermined by available evidence. Production and interpretation of new forms was taken to be a straightforward matter of analogy, posing no problems of principle.[5] Attention to the questions of (1) quickly reveals that exactly the opposite is the case: language poses in a sharp and clear form what has sometimes been called "Plato's problem," the problem of "poverty of stimulus," of accounting for the richness, complexity, and specificity of shared knowledge, given the limitations of the data available. This difference of perception concerning where the problem lies—overlearning or poverty of evidence—reflects very clearly the effect of the shift of focus that inaugurated the study of generative grammar.

The key step: Plato's problem. Against overlearning, the examples (2) to (7) show knowledge the data never supplied, so it must come from the learner. *Drawn on p. 12.*

"overlearning": *a habit drilled in by far more repetition than it needs, as in behaviourist learning theory. If that were so the data would overdetermine the grammar, the reverse of what the examples below show.*

"Plato's problem": *in the Meno Socrates draws geometry out of an untaught slave boy and asks how he could know it. For language: how do a child's thin data yield knowledge that is rich, specific and the same for all?*

A great many examples have been given over the years to illustrate what clearly is the fundamental problem: the problem of poverty of evidence. A familiar example is the structure-dependence of rules, the fact that without instruction or direct evidence, children unerringly use computationally complex structure-dependent rules rather than computationally simple rules that involve only the predicate "leftmost" in a linear

structure-dependence of rules: *a rule stated over phrases rather than over the string of words. The example is worked out at the end of this reading,* *Reflections pp. 30 to 33.*

sequence of words.[6] To take some other examples, to which we will return, consider sentences (2)–(7):

I wonder who [the men expected to see them] (2)
[the men expected to see them] (3)
John ate an apple (4)
John ate (5)
John is too stubborn to talk to Bill (6)
John is too stubborn to talk to (7)

Both (2) and (3) include the clause bounded by brackets, but only in (2) may the pronoun *them* be referentially dependent on the antecedent *the men*; in (3) the pronoun is understood as referring in some manner indicated in the situational or discourse context, but not to the men. Numerous facts of this sort, falling under what is now generally called "binding theory," are known without relevant experience to differentiate the cases. Such facts pose a serious problem that was not recognized in earlier work: How does every child know, unerringly, to interpret the clause differently in the two cases? And why does no pedagogic grammar have to draw the learner's attention to such facts (which were, in fact, noticed only quite recently, in the course of the study of explicit rule systems in generative grammar)?

"binding theory": *the principles fixing when a pronoun may take an antecedent in its sentence. The clauses differ in an unheard subject: in (2) the trace of who, in (3) a silent the men, barred to them.* *See p. 267.*

Turning to examples (4)–(7), sentence (5) means that John ate something or other, a fact that one might explain on the basis of a simple inductive procedure: *ate* takes an object, as in (4), and if the object is missing, it is understood as arbitrary. Applying the same inductive procedure to (6) and (7), it should be that (7) means that John is so stubborn that he (John) will not talk to some arbitrary person, on the analogy of (6). But the meaning is, in fact, quite different: namely, that John is so stubborn that some arbitrary person won't talk to him (John). Again, this is known without training or relevant evidence.[7]

a simple inductive procedure: *the induction a child could actually make from what she hears. It gets (5) right and (7) wrong, yet no child goes wrong on (7), so the data can't be what taught her. Plato's problem in one pair.*

The situation is, in fact, more complex. Although plausible, the inductive procedure suggested for the relatively straightforward examples (4)–(5) does not seem correct. As noted by Howard Lasnik, the word *eat* has a somewhat different meaning in its intransitive usage, something like *dine*. One can say "John ate his shoe," but "John ate" cannot be understood to include this case. The observation is general for such cases. The

Howard Lasnik: *Chomsky's student. His point: 'ate' with no object means 'dined', so 'John ate' isn't 'John ate something'. Even the easy induction fails.*

intransitive forms differ from normal intransitives in other respects; for example, we can form "the dancing bear" (corresponding to "the bear that dances"), but not "the eating man" (corresponding to "the man who eats").[8] Such facts pose further problems of poverty of stimulus.

Children do not make errors about the interpretation of such sentences as (6)–(7) past a certain stage of development, and if they did, the errors would largely be uncorrectable. It is doubtful that even the most compendious traditional or teaching grammar notes such simple facts as those illustrated in (2)–(7), and such observations lie far beyond the domain of structural grammars. A wide variety of examples of this sort immediately come to attention when one faces the questions formulated in (1).

largely be uncorrectable: *a correction needs someone who notices the error and evidence showing the right reading, and there's neither: adults don't register these facts and no grammar states them. Feedback isn't the source.*

Knowledge of language is often characterized as a practical ability to speak and understand, so that questions (1i) and (1iii) are closely related, perhaps identified. Ordinary usage makes a much sharper distinction between the two questions, and is right to do so. Two people may share exactly the same knowledge of language but differ markedly in their ability to put this knowledge to use. Ability to use language may improve or decline without any change in knowledge. This ability may also be impaired, selectively or in general, with no loss of knowledge, a fact that would become clear if injury leading to impairment recedes and lost ability is recovered. Many such considerations support the commonsense assumption that knowledge cannot be properly described as a practical ability. Furthermore, even if this view could somehow be maintained, it would leave open all of the serious questions. Thus, what is the nature of the "practical ability" manifested in our interpretation of the sentences (2)–(7), how is it properly described, and how is it acquired?

Is knowing a language a practical ability, (i) and (iii) one question? (8) to (10) show knowledge outrunning our access to it, Jones (p. 10) that the reduction only renames it. Landed on p. 12.

could somehow be maintained: *the second prong grants the reduction for argument's sake. An ability that gets (2) to (7) right would need the same description and origin a grammar and UG give, so nothing is saved.*

Often it is not immediately obvious what our knowledge of language entails in particular cases, a fact illustrated even with short and simple sentences such as (8)–(10):

his wife loves her husband	(8)
John is too clever to expect us to catch Bill	(9)
John is too clever to expect us to catch	(10)

his wife loves her husband: *if her means his wife, can his mean her husband? No: each phrase's reference would wait on the other's, a circle. Seeing it takes thought, yet the fact was known all along.*

In the case of (8), it takes some thought to determine whether *his* can be referentially dependent on *her husband* if *her* is

dependent on *his wife*—that is, if the reference of either *he* or *she* is not somehow contextually indicated.[9] Examples (9) and (10) are, in fact, analogous to (6) and (7), respectively, but again, it takes some thought to discover that (10) means that John is so clever that an arbitrary person cannot expect us to catch him (John), although it is clear at once that it does not mean that John is so clever that he (John) cannot catch some arbitrary person, on the analogy of (9) (and (4), (5)). Our abilities seem limited somehow in such cases (and there are far more complex ones), but it would make little sense to speak of our knowledge of language as "limited" in any comparable way.

Our abilities seem limited: *the asymmetry that parts (i) from (iii): what we know fixes one reading of (10), though it takes work to find it. What's limited is our access to the knowledge, not the knowledge itself.*

Suppose we insist on speaking of knowledge of language as a practical ability to speak and understand. Then normal usage must be revised in numerous cases such as those just discussed. Suppose that Jones takes a public speaking course and improves his ability to speak and understand without any change in his knowledge of English, as we would describe the situation in normal usage. We must now revise this common-sense usage and say, rather, that Jones has improved his ability$_1$ to use his ability$_2$ to speak and understand; similar translations are required in the other cases. But the two occurrences of "ability" in this description are hardly more than homonyms. Ability$_1$ is ability in the normal sense of the word: it can improve or decline, can be inadequate to determine consequences of knowledge, and so on. Ability$_2$, however, remains stable while our ability to use it changes, and we have this kind of "ability" even when we are unable to detect what it entails in concrete cases. In short, the neologism "ability$_2$" is invested with all the properties of knowledge. Note that there are cases when we do speak of abilities that we cannot put to use: for example, the case of swimmers who cannot swim because their hands are tied, although they retain the ability to swim. The cases in question are not of this sort, however.

hardly more than homonyms: *words spelt alike with unrelated meanings, like 'bank'. So the reduction works only by equivocation: ability2 shares nothing with ability but the name. Knowledge is renamed, not reduced.*

swimmers who cannot swim: *an ability kept but not exercisable, and why it doesn't help: the swimmer is stopped by something outside him. Nothing outside stops us reading (10). We just don't yet know what we know.*

The purpose of the attempt to reduce knowledge to ability is, presumably, to avoid problematic features that seem to inhere in the concept of knowledge, to show that these can be explained in dispositional or other terms more closely related to actual behavior (whether this is possible even in the case of ability$_1$, the normal sense, is another question). But nothing of the sort is achieved by this departure from ordinary usage; the

dispositional or other terms: *the motive, as in Ryle's knowing how or Quine's dispositions (p. 31): knowledge seems to carry justification and awareness, which a disposition to behave doesn't.*

problems remain, exactly as before, now embedded in terminological confusion. The task of determining the nature of our knowledge (= ability$_2$), and accounting for its origins and use, remains exactly as challenging as before, despite the terminological innovations.

Plato's problem and the ability question meet in (14): no data rule out its analogical readings, yet we know they're wrong, and we could produce them, so the bar is knowledge, not ability (p. 12).

Other examples similar to (8)–(10) raise further questions. Consider the following sentences:

John is too stubborn to expect anyone to talk to Bill (11)

John is too stubborn to visit anyone who talked to Bill (12)

Suppose we delete *Bill* from (11) and (12), yielding (13) and (14), respectively:

John is too stubborn to expect anyone to talk to (13)

John is too stubborn to visit anyone who talked to (14)

Sentence (13) is structurally analogous to (10), and is understood in the same manner: it means that John is so stubborn that an arbitrary person would not expect anyone to talk to him (John). "By analogy," then, we would expect sentence (14) to mean that John is so stubborn that an arbitrary person would not visit anyone who talked to him (John). But it does not have that meaning; in fact, it is gibberish. Here we have a double failure of analogy. Sentence (14) is not understood "on the analogy" of (4), (5), (6), (9), and (12) (hence meaning that John is so stubborn that he (John) would not visit anyone who talked to some arbitrary person), nor is it understood "on the analogy" of (7), (10), and (13); rather, it has no interpretation at all. And while the status of (11), (12), and (14) is immediately obvious, it takes some thought or preparation to see that (13) has the interpretation it does have, and thus to determine the consequences of our knowledge in this case.

a double failure of analogy: *analogy offers (14) two readings, one from each family of examples, and English allows neither. Analogy can't even predict that a string means nothing, since every string resembles something heard.*

some thought or preparation: *the asymmetry of (10) again: judging (11), (12) and (14) is instant, finding what (13) means is slow. Knowledge fixes the result, ability to use it governs the speed, the distinction of pp. 9 to 10.*

Again, these are facts that we know, however difficult it may be to determine that our system of knowledge has these consequences. We know these facts without instruction or even

direct evidence, surely without correction of error by the speech community. It would be absurd to try to teach such facts as these to people learning English as a second language, just as no one taught them to us or even presented us with evidence that could yield this knowledge by any generally reliable procedure. This is knowledge without grounds, without good reasons or support by reliable procedures in any general or otherwise useful sense of these notions. Were we to insist that knowledge is a kind of ability, we would have to claim that we lack the ability to understand "John is too stubborn to talk to" as meaning "John is too stubborn to talk to someone or other" (on the analogy of "John ate an apple"—"John ate"), and that we lack the ability to understand (14) on the analogy of "John ate an apple"—"John ate" (so that it means that John is too stubborn to visit anyone who talked to someone or other) or on the analogy of "John is too stubborn to talk to," with the "inversion strategy" that we somehow use in this case (so that (14) means that John is too stubborn for someone or other to visit anyone who talked to him, John). But these would be odd claims, to say the least. These are not failures of ability. It is not that we are too weak, or lack some special skill that could be acquired. We are perfectly capable of associating the sentence (14), for example, with either of the two meanings that would be provided "by analogy" (or others), but we know that these are not the associations that our knowledge of the language provides; ability is one thing, knowledge something quite different. The system of knowledge that has somehow developed in our minds has certain consequences, not others; it relates sound and meaning and assigns structural properties to physical events in certain ways, not others.

knowledge without grounds: *no teaching, no evidence and no reliable method produced it, so it fails every usual test for what makes a belief knowledge. Whether 'know' is still the word is put off to Chapter 4, p. 266.*

"inversion strategy": *the name for how (7) is read: the missing object isn't 'someone', the sentence is turned round and John becomes the one talked to. On (14) it would give the second analogical reading, also refused.*

We are perfectly capable: *the reversal: if knowing were an ability, not reading (14) by analogy would be a lack of skill. But we can do it, so the 'can't' is 'that isn't what it means', a fact of correctness, not capacity.*

It seems that there is little hope in accounting for our knowledge in terms of such ideas as analogy, induction, association, reliable procedures, good reasons, and justification in any generally useful sense, or in terms of "generalized learning mechanisms" (if such exist). And it seems that we should follow normal usage in distinguishing clearly between knowledge and ability to use that knowledge. We should, so it appears, think of knowledge of language as a certain state of the mind/brain, a relatively stable element in transitory mental states once it is attained; furthermore, as a state of some distinguishable

Chapter 1's conclusion: the facts of pp. 7 to 12 leave little hope for analogy, induction or general learning, and knowledge isn't ability. So a state of the distinct faculty of p. 4, SL on p. 22.

faculty of the mind—the language faculty—with its specific properties, structure, and organization, one "module" of the mind.[10]

one "module" of the mind: *Fodor's term (The Modularity of Mind, 1983): a special-purpose system with its own principles, closed to general reasoning. Note 10 on p. 14 says why this faculty isn't a Fodorian input module.*

NOTES

1. On these and many other discussions, primarily in the seventeenth–nineteenth centuries, see Chomsky (1966). For discussion of some misinterpretation of this work, see Bracken (1984).

2. The alleged *a priorism* of work in this tradition has often been exaggerated. See Chomsky (1966) and more recent work for discussion of this point.

3. The tradition, in this case, is a different one, represented in its most advanced form in the early work of the Indian grammarians 2,500 years ago. See Kiparsky (1982). A modern counterpart is Bloomfield (1939), which was radically different in character from the work of the period and inconsistent with his own theories of language, and remained virtually without influence or even awareness despite Bloomfield's great prestige.

4. See Newmeyer (1980) for one view of the history of this period prior to the second major conceptual shift; and for some more personal comments, the introduction to Chomsky (1975a), a somewhat abbreviated version of a 1956 revision of a 1955 manuscript, both unpublished. See Lightfoot (1982) and Hornstein and Lightfoot (1981) for discussion of the general backgrounds for much current work, and Radford (1981) for an introduction to the work that led to the second conceptual shift. See Chomsky (1981) for a more technical presentation of some of the ideas that entered into this conceptual shift and van Riemsdijk and Williams (1985) for an introductory study of this current work.

5. Although basically adopting this point of view, W.V. Quine, however, argued that there is a very severe, in fact, insuperable problem of underdetermination affecting all aspects of language and grammar, and much of psychology more generally (Quine, 1960, 1972). I do not think that he succeeded in showing that some novel form of indeterminacy affects the study of language beyond the normal underdetermination of theory by evidence; his own formulations of the thesis furthermore involve internal inconsistency (see Chomsky, 1975b, 1980b). There seems no reason on these grounds, then, to distinguish linguistics or psychology in principle from the natural sciences in accordance with what Hockney (1975) calls Quine's "bifurcation

thesis." A similar conclusion is reached by Putnam (1981) in his abandonment of metaphysical realism on Quinean grounds. His step also abandons the bifurcation thesis, although in the opposite direction.

6. See Chomsky (1975a). See Crain and Nakayama (1984) for empirical study of this question with 3–5-year-old children.

7. The reaction to such phenomena, also unnoticed until recently, again illustrates the difference of outlook of structuralist-descriptive and generative grammar. For some practitioners of the former, the statement of the facts, which is straightforward enough once they are observed, is the answer—nothing else is necessary; for the latter, the statement of the facts poses the problem to be solved. Cf. Ney (1983), particularly, his puzzlement about the "peculiar view of grammar [that] unnecessarily complicates the whole matter" by seeking an explanation for the facts. Note that there is no question of right or wrong here, but rather of topic of inquiry.

8. In early work, such facts were used to motivate an analysis of intransitives such as *eat* as derived from corresponding transitives by a system of ordered rules that excluded the unwanted cases; see Chomsky (1962).

9. On structures of this type, and problems of binding theory, more generally, see Higginbotham (1938a), among much other work.

10. See Fodor (1983). But it is too narrow to regard the "language module" as an input system in Fodor's sense, if only because it is used in speaking and thought. We might consider supplementing this picture by adding an "output system," but plainly this must be linked to the input system; we do not expect a person to speak only English and understand only Japanese. That is, the input and output systems must each access a fixed system of knowledge. The latter, however, is a central system which has essential problems of modularity, a fact that brings the entire picture into question. Furthermore, even regarded as an input system, the language module does not appear to have the property of rapidity of access that Fodor discusses, as indicated by (8)–(14). Note also that even if Fodor is right in believing that there is a sharp distinction between modules in his sense and "the rest," which is holistic in several respects, it does not follow that the residue is unstructured. In fact, this seems highly unlikely, if only because of the "epistemic boundedness" that he notes. Many other questions arise concerning Fodor's very intriguing discussion of these issues, which I will not pursue here.

2 Concepts of Language

2.1 THE COMMONSENSE CONCEPT AND DEPARTURES FROM IT

What every science of language drops from the everyday notion: political borders and the learner's goal, for an idealized uniform community (p. 16). Then E-language (p. 20) against I-language (p. 22).

Let us turn now to the questions of (1) of Chapter 1. To begin with, let us distinguish the intuitive, pretheoretic commonsense notion of language from various technical concepts that have been proposed with the intent of developing an eventual science of language. Let us call the latter "scientific approaches" to language, with an eye directed more toward a possible future than a present reality, some might argue. The scientific approaches, I believe without exception, depart from the commonsense notion in several ways; these departures also affect the concepts of knowledge or understanding of language, use of language, rule of language, rule-guided linguistic behavior, and others.

In the first place, the commonsense notion of language has a crucial sociopolitical dimension. We speak of Chinese as "a language," although the various "Chinese dialects" are as diverse as the several Romance languages. We speak of Dutch and German as two separate languages, although some dialects of German are very close to dialects that we call "Dutch" and are not mutually intelligible with others that we call "German." A standard remark in introductory linguistics courses is that a language is a dialect with an army and a navy (attributed to Max Weinreich). That any coherent account can be given of "language" in this sense is doubtful; surely, none has been offered or even seriously attempted. Rather, all scientific approaches have simply abandoned these elements of what is called "language" in common usage.[1]

Max Weinreich: *the Yiddish linguist, who reported the saying in 1945. Which dialects count as one language is settled by states and borders, which no theory of the mind can capture, so science drops it.*

normative-teleological element: *we describe the learner by a goal he hasn't reached, English, not by the language he has, as if he fell short of an outside standard. The I-language credits him with a perfect system of his own.*

The commonsense notion also has a normative-teleological element that is eliminated from scientific approaches. I do not refer here to prescriptive grammar but to something else. Consider the way we describe a child or a foreigner learning English. We have no way of referring directly to what that person knows: It is not English, nor is it some other language that resembles English. We do not, for example, say that the person has a perfect knowledge of some language L, similar to English but still different from it. What we say is that the child or foreigner has a "partial knowledge of English," or is "on his or her way" toward acquiring knowledge of English, and if they reach the goal, they will then know English. Whether or not a coherent account can be given of this aspect of the commonsense terminology, it does not seem to be one that has any role in an eventual science of language.

I will follow standard practice in disregarding these aspects of the commonsense notions of language and the associated notions of rule-following and so forth, although the departure should be noted, and one may ask whether it is entirely innocent.

entirely innocent: *what might not be: a rule one can follow or break needs a standard beyond one's present state, and once the goal is dropped the I-language supplies none. The rule-following of p. 32 inherits this.*

Modern linguistics commonly avoided these questions by considering an idealized "speech community" that is internally consistent in its linguistic practice.[2] For Leonard Bloomfield, for example, a language is "the totality of utterances that can be made in a speech community," regarded as homogeneous (Bloomfield, 1928/1957). In other scientific approaches, the same assumption enters in one or another form, explicitly or tacitly, in identification of the object of inquiry. No attempt is made to capture or formulate any concept with the sociopolitical or normative-teleogical aspects of informal usage of the term "language." The same is true of approaches that understand language to be a social product in accordance with the Saussurean concept of "langue."

"langue": *Saussure's system of signs held in common by a community, against parole, individual acts of speech. It counts here as E-language, a social fact, though p. 31 allows it points toward I-language.*

Of course, it is understood that speech communities in the Bloomfieldian sense—that is, collections of individuals with the same speech behavior[3]—do not exist in the real world. Each individual has acquired a language in the course of complex social interactions with people who vary in the ways in which they speak and interpret what they hear and in the internal representations that underlie their use of language. Structural linguistics abstracted from these facts in its attempts at theory construction; we also abstract from these facts in posing ques-

tions (1) of Chapter 1, considering only the case of a person presented with uniform experience in an ideal Bloomfieldian speech community with no dialect diversity and no variation among speakers.

The idealizations under which (1) is asked: uniform experience (p. 16), a 'pure' instance of UG, one faculty for the species (p. 18). Defended on p. 17 as isolating a real property of mind, P.

We should also make note of a more subtle theory-internal assumption: The language of the hypothesized speech community, apart from being uniform, is taken to be a "pure" instance of UG in a sense that must be made precise, and to which we will return. We exclude, for example, a speech community of uniform speakers, each of whom speaks a mixture of Russian and French (say, an idealized version of the nineteenth-century Russian aristocracy). The language of such a speech community would not be "pure" in the relevant sense, because it would not represent a single set of choices among the options permitted by UG but rather would include "contradictory" choices for certain of these options.

options permitted by UG: *principles and parameters: UG fixes the principles and leaves a few switches open, a language being one setting of them. Verb before object is one (p. 27), Pierre's adverb rule another (p. 266).*

Questions (1) of Chapter 1, then, arise initially under these idealizations, and the same is true, in effect, of other approaches to language, although the fact is often not explicitly recognized and may even sometimes be denied.

The legitimacy of these idealizations has sometimes been questioned, but on dubious grounds.[4] Indeed, they seem indispensable. Surely there is some property of mind P that would enable a person to acquire a language under conditions of pure and uniform experience, and surely P (characterized by UG) is put to use under the real conditions of language acquisition. To deny these assumptions would be bizarre indeed: It would be to claim either that language can be learned only under conditions of diversity and conflicting evidence, which is absurd, or that the property P exists—there exists a capacity to learn language in the pure and uniform case—but the actual learning of language does not involve this capacity. In the latter case, we would ask why P exists; is it a "vestigial organ" of some sort? The natural approach, and one that I think is tacitly adopted even by those who deny the fact, is to attempt to determine the real property of mind P, and then ask how P functions under the more complex conditions of actual linguistic diversity. It seems clear that any reasonable study of the nature, acquisition, and use of language in real life circumstances must accept these assumptions and then proceed on the basis of some tentative characterization of the property of mind P. In short, the ideali-

To deny these assumptions would be bizarre indeed: *two ways to deny them, both bad. Either language is learnt only from conflicting data, absurd, or P exists but is idle in real learning, a vestigial organ. So P is real and used, and UG describes it.*

zations made explicit in more careful work are hardly controversial; they isolate for examination a property of the language faculty the existence of which is hardly in doubt, and which is surely a crucial element in actual language acquisition.

By making these idealizations explicit and pursuing our inquiry in accordance with them, we do not in any way prejudice the study of language as a social product. On the contrary, it is difficult to imagine how such studies might fruitfully progress without taking into account the real properties of mind that enter into the acquisition of language, specifically, the properties of the initial state of the language faculty characterized by UG.

Note also that the study of language and UG, conducted within the framework of individual psychology, allows for the possibility that the state of knowledge attained may itself include some kind of reference to the social nature of language. Consider, for example, what Putnam (1975) has called "the division of linguistic labor." In the language of a given individual, many words are semantically indeterminate in a special sense: The person will defer to "experts" to sharpen or fix their reference. Suppose, for example, that someone knows that yawls and ketches are sailing vessels but is unsure of the exact reference of the words "yawl" and "ketch," leaving it to specialists to fix this reference. In the lexicon of this person's language, the entries for "yawl" and "ketch" will be specified to the extent of his or her knowledge, with an indication that details are to be filled in by others, an idea that can be made precise in various ways but without going beyond the study of the system of knowledge of language of a particular individual. Other social aspects of language can be regarded in a like manner—although this is not to deny the possibility or value of other kinds of study of language that incorporate social structure and interaction. Contrary to what is sometimes thought, no conflicts of principle or practice arise in this connection.

division of linguistic labor: *Putnam 1975: I mean different trees by 'elm' and 'beech' though I can't tell them apart, since experts fix the reference. So a social fact can sit in one speaker's lexicon. Used again, p. 267.*

We are also assuming another idealization: That the property of mind described by UG is a species characteristic, common to all humans. We thus abstract from possible variation among humans in the language faculty. It is plausible to suppose that apart from pathology (potentially an important area of inquiry), such variation as there may be is marginal and can be safely ignored across a broad range of linguistic investigation. Again, the assumption is conventional, though generally explicit, in scientific

in scientific approaches. Weaker assumptions than strict identity would suffice for the discussion below, but this stronger assumption seems a reasonable one, to a very good approximation, and I will keep to it here.

2.2 EXTERNALIZED LANGUAGE

The rival, E-language (p. 20), argued against to p. 32: if a language is outside the mind, a grammar only describes it, and of two that yield the same set neither is the correct one (p. 20).

Scientific approaches to language, in the sense of the term used earlier, have developed various technical notions of language to replace the commonsense notion. The term "grammar" has also been used in a variety of ways. In conventional usage, a grammar is a description or theory of a language, an object constructed by a linguist. Let us keep to this usage. Then associated with the various technical notions of language there are corresponding notions of grammar and of universal grammar (UG).

Let us keep to this usage: *so 'grammar' means the linguist's theory, 'language' the thing it's a theory of. Chomsky's earlier practice used 'grammar' for the theory and for what's in the head, which he regrets on p. 29.*

Structural and descriptive linguistics, behavioral psychology, and other contemporary approaches tended to view a language as a collection of actions, or utterances, or linguistic forms (words, sentences) paired with meanings, or as a system of linguistic forms or events. In Saussurean structuralism, a language (*langue*) was taken to be a system of sounds and an associated system of concepts; the notion of sentence was left in a kind of limbo, perhaps to be accommodated within the study of language use. For Bloomfield, as noted earlier, a language is "the totality of utterances that can be made in a speech community." The American variety of structural-descriptive linguistics that was heavily influenced by Bloomfield's ideas furthermore concentrated primarily on sound and word structure, apart from various proposals, notably those of Zellig Harris, as to how larger units (phrases) could be constructed by analytic principles modelled on those introduced for phonology and morphology.[5] Many researchers today adopt a position of the sort lucidly developed by David Lewis, who defines a language as a pairing of sentences and meanings (the latter taken to be set-theoretic constructions in terms of possible worlds) over an infinite range, where the language is "used by a population" when certain regularities "in action or belief" hold among the population with reference to the language, sustained by an interest in communication.[6]

David Lewis: *'Languages and Language' (1975): a language is a function from sentences to meanings, used by a population when a convention of truthfulness and trust in it holds there. The reply is on pp. 30 to 31.*

Let us refer to such technical concepts as instances of "externalized language" (E-language), in the sense that the construct is understood independently of the properties of the mind/brain. Under the same rubric we may include the notion of language as a collection (or system) of actions or behaviors of some sort. From a point of view such as this, a grammar is a collection of descriptive statements concerning the E-language, the actual or potential speech events (perhaps along with some account of their context of use or semantic content). In technical terms, the grammar may be regarded as a function that enumerates the elements of the E-language. Sometimes, grammar has been regarded as a property of E-language, as in Bloomfield's remark that a grammar is "the meaningful arrangement of forms in a language" (Bloomfield, 1933). Despite appearances, the problem of accounting for the unbounded character of the E-language and the person's knowledge of language including this fundamental property is not squarely addressed in such approaches, a matter to which we will return.

(E-language): *external because defined without reference to any mind: a set of utterances, of sentences, or of sentence-meaning pairs. Everything on pp. 16 to 19, Bloomfield, Saussure and Lewis, falls under it.*

The E-language is now understood to be the real object of study. Grammar is a derivative notion; the linguist is free to select the grammar one way or another as long as it correctly identifies the E-language. Apart from this consideration, questions of truth and falsity do not arise. Quine, for example, has argued that it is senseless to take one grammar rather than another to be "correct" if they are extensionally equivalent, characterizing the same E-language, for him a set of expressions (Quine, 1972). And Lewis doubts that there is any way "to make objective sense of the assertion that a grammar G is used by a population P whereas another grammar G′, which generates the same language as G, is not."

extensionally equivalent: *generating the same set of sentences. Quine (1972) adds that behaviour can 'fit' both grammars while 'guided' by neither, so there's no fact about which one is followed. Answered on pp. 23 and 31.*

The notion of E-language is familiar from the study of formal systems, as is the conclusion just cited: In the case of the "language of arithmetic," for example, there is no objective sense to the idea that one set of rules that generates the well-formed formulas is correct and another wrong.

As for UG, to the extent that such a study was recognized as legitimate, this theory would consist of statements that are true of many or all human languages, perhaps a set of conditions satisfied by the E-languages that count as human languages. Some appeared to deny the possibility of the enterprise, for example, Martin Joos, who put forth what he called the "Boas-

ian" view that "languages could differ from each other without limit and in unpredictable ways," echoing William Dwight Whitney's reference to "the infinite diversity of human speech" and Edward Sapir's notion that "language is a human activity that varies without assignable limit."[7] Such statements reflect a fairly broad consensus of the time. Although they could hardly have been intended literally, they did express a relativistic impulse that denigrated the study of UG. More precisely, it cannot be that human language varies without assignable limit, although it might be true that it is "infinitely diverse"; it is an empirical question of some interest whether UG permits an infinite variety of possible languages (or a variety that is infinite in more than structurally trivial respects, say, with no bound on vocabulary), or only a finite diversity.[8]

Nevertheless, significant contributions were made to UG in our sense within these traditions. For example, the theory of distinctive features in phonology, which greatly influenced structuralist studies in other fields, postulated a fixed inventory of "atomic elements" from which phonological systems could be drawn, with certain general laws and implicational relations governing the choice. And it was generally assumed that such notions as topic and comment, or subject and predicate, were universal features of language, reflecting the fact that a declarative sentence is about something and says something about it. Later, important work on linguistic universals was conducted by Joseph Greenberg and others, yielding many generalizations that require explanation, for example, the fact that if a language has subject-object-verb order, it will tend to have postpositions rather than prepositions, and so on.

Along these lines, then, we may develop a certain technical concept of language (E-language), and an associated concept of grammar and UG, as a basis for a scientific study of language. Many different specific ideas fall roughly within this general framework.

2.3 INTERNALIZED LANGUAGE

The alternative, from Jespersen: a 'notion of structure' in the speaker, named I-language on p. 22. The reply to Quine and Lewis is built on it.

A rather different approach was taken, for example, by Otto Jespersen, who held that there is some "notion of structure" in the mind of the speaker "which is definite enough to guide

him in framing sentences of his own," in particular, "free expressions" that may be new to the speaker and to others.[9] Let us refer to this "notion of structure" as an "internalized language" (I-language). The I-language, then, is some element of the mind of the person who knows the language, acquired by the learner, and used by the speaker-hearer.

"free expressions": *Jespersen, Philosophy of Grammar (1924): formulas are learnt whole, like 'how do you do', free expressions built fresh on a pattern. Only a 'notion of structure' in the speaker explains the second.*

Taking language to be I-language, the grammar would then be a theory of the I-language, which is the object under investigation. And if, indeed, such a "notion of structure" exists, as Jespersen held, then questions of truth and falsity arise for grammar as they do for any scientific theory. This way of approaching the questions of language is radically different from the one sketched above and leads to a very different conception of the nature of the inquiry.

truth and falsity arise for grammar: *the contrast with p. 20: there a grammar was a description of a given set, none more correct than another. If the I-language is real, a grammar can be false yet generate the right sentences.*

Let us return now to the point of view outlined in Chapter 1. Knowing the language L is a property of a person H; one task of the brain sciences is to determine what it is about H's brain by virtue of which this property holds. We suggested that for H to know the language L is for H's mind/brain to be in a certain state; more narrowly, for the language faculty, one module of this system, to be in a certain state S_L.[10] One task of the brain sciences, then, is to discover the mechanisms that are the physical realization of the state S_L.

state SL: *the state of H's language faculty when H knows L. Two more states come on pp. 24 to 25: S0, the initial state UG describes, and Ss, the steady state reached, of which the I-language is a component.*

Suppose we analyze the notion "H knows language L" in relational terms, that is, as involving a relation R (knowing, having, or whatever) holding between H and an abstract entity L. One might question this move; we speak of a person as knowing U.S. history without assuming that there is an entity, U.S. history, that the person knows, or knows in part. Let us, however, assume the move to be legitimate in this case. The assumption will be justified to the extent that this move contributes to providing insight into the questions that primarily concern us, those of (1) of Chapter 1; this would be the case, for example, if there are significant principles governing the set of postulated entities L. Suppose that we proceed further to regard talk of mind as talk about the brain undertaken at a certain level of abstraction at which we believe, rightly or wrongly, that significant properties and explanatory principles can be discovered. Then statements about R and L belong to the theory of mind, and one task of the brain sciences will be to explain what

R (knowing, having, or whatever): *the move that gives a grammar something to be about: if knowing is a relation, L is an object a grammar can be true of, the I-language of p. 23. This R is a relation, not the rule R of p. 266.*

it is about H's brain (in particular, its language faculty) that corresponds to H's knowing L, that is, by virtue of which R(H, L) holds and the statement that R(H, L) is true.

It is natural to take L to be I-language, Jespersen's "notion of structure," regarding this as an entity abstracted from a state of the language faculty, the latter being one component of the mind. Then, for H to know L is for H to have a certain I-language. The statements of a grammar are statements of the theory of mind about the I-language, hence statements about structures of the brain formulated at a certain level of abstraction from mechanisms. These structures are specific things in the world, with their specific properties. The statements of a grammar or the statement that R(H, L) are similar to statements of a physical theory that characterizes certain entities and their properties in abstraction from whatever may turn out to be the mechanisms that account for these properties: say, a nineteenth-century theory about valence or properties expressed in the periodic table. Statements about I-language or the statement that R(H, L) (for various choices of H and L) are true or false, much in the way that statements about the chemical structure of benzene, or about the valence of oxygen, or about chlorine and fluorine being in the same column of the periodic table are true or false. The I-language L may be the one used by a speaker but not the I-language L′, even if the two generate the same class of expressions (or other formal objects) in whatever precise sense we give to this derivative notion; L′ may not even be a possible human I-language, one attainable by the language faculty.

abstraction from mechanisms: *a grammar is about the brain, but at a level that says nothing about neurons, as valence theory was true of atoms before anyone knew their structure. Its truth doesn't wait on neuroscience.*

but not the I-language L′: *the answer to Quine and Lewis (p. 20): two I-languages generating the same sentences are different states, so a speaker has one and not the other as a matter of fact. Output settles nothing.*

UG now is construed as the theory of human I-languages, a system of conditions deriving from the human biological endowment that identifies the I-languages that are humanly accessible under normal conditions. These are the I-languages L such that R(H, L) may be true (for normal H, under normal conditions).[11]

Of course, there is no guarantee that this way of approaching the problems of (1) in Chapter 1 is the correct one. This approach may turn out to be thoroughly misguided, even if it achieves substantial success—just as a theory of valence, etc. might have turned out to be completely off the track, despite its substantial success in nineteenth-century chemistry. It is always reasonable to consider alternative approaches, if they can be

devised, and this will remain true no matter what successes are achieved. The situation does not seem different in principle from what we find in other areas of empirical inquiry. I will suggest directly that in certain fundamental respects early ideas about I-language were misguided and should be replaced by a rather different conception, although one formulated in the same general framework. The reasons, however, do not derive from any incoherence or flaw in the general approach but rather from empirical considerations of description and explanation.

early ideas about I-language were misguided: *the second conceptual shift, Chapter 3: the I-language as a system of rules gives way to principles with parameters set by experience. That's why R, a rule, yields to R', a parameter value, on p. 266.*

2.4 THE SHIFT OF FOCUS FROM E-LANGUAGE TO I-LANGUAGE

2.4.1 On the Reasons for the Shift of Focus

The shift restated as E-language to I-language, with the first answers to (1): knowledge of a rule system, S0 turned by experience into Ss, a parser for use. The whole picture is summed up on p. 25.

In Chapter 1, we saw that the study of generative grammar shifted the focus of attention from actual or potential behavior and the products of behavior to the system of knowledge that underlies the use and understanding of language, and more deeply, to the innate endowment that makes it possible for humans to attain such knowledge. The shift in focus was from the study of E-language to the study of I-language, from the study of language regarded as an externalized object to the study of the system of knowledge of language attained and internally represented in the mind/brain. A generative grammar is not a set of statements about externalized objects constructed in some manner. Rather, it purports to depict exactly what one knows when one knows a language: that is, what has been learned, as supplemented by innate principles. UG is a characterization of these innate, biologically determined principles, which constitute one component of the human mind—the language faculty.

as supplemented by innate principles: *the two components of p. 26 in advance: the learnt part, English rather than Japanese, and the innate part UG describes. The grammar depicts both, since what one knows includes what was never learnt.*

With this shift of focus, we at once face the questions (1) of Chapter 1. In the earliest work, the answer to (1i) was taken to be that knowledge of language is knowledge of a certain rule system; the answer to (1ii), that this knowledge arises from an initial state S_0 that converts experience to a "steady state" S_s, which incorporates an I-language. Acquisition of language is,

then, a matter of adding to one's store of rules, or modifying this system, as new data are processed. Question (1iii) breaks down into two parts: a "perception problem" and a "production problem." The perception problem would be dealt with by construction of a parser that incorporates the rules of the I-language along with other elements: a certain organization of memory and access (perhaps a deterministic pushdown structure with buffer of a certain size; see Marcus, 1980), certain heuristics, and so forth. A parser should not map expressions into their structures in the way that these are associated by the I-language. For example, a parser should fail to do so in the case of so-called "garden-path sentences"[12] or sentences that overload memory for left-to-right pass, it should mirror the difficulties experienced with sentences such as (8)–(14) of Chapter 1 and so forth. The production problem is considerably more obscure; we will return to that.

a parser: *how a hearer assigns structure in real time, under memory limits. It needn't match the I-language: 'the horse raced past the barn fell' is grammatical, yet the parser stumbles on it at first.*

The E-language that was the object of study in most of traditional or structuralist grammar or behavioral psychology is now regarded as an epiphenomenon at best. Its status is similar to that of other derivative objects, say, the set of rhyming pairs, which is also determined by the I-language that constitutes the system of knowledge attained. One might argue that the status of the E-language is considerably more obscure than that of the set of rhyming pairs, since the latter is determined in a fairly definite way by the I-language whereas the bounds of E-language can be set one way or another, depending on some rather arbitrary decisions as to what it should include.

an epiphenomenon at best: *a by-product that does no explaining. The set of rhymes is fixed by the I-language too, yet no one would take it for what a speaker knows. The E-language is no better placed.*

Summarizing, then, we have the following general picture. The language faculty is a distinct system of the mind/brain, with an initial state S_0 common to the species (to a very close first approximation, apart from pathology, etc.) and apparently unique to it in essential respects.[13] Given appropriate experience, this faculty passes from the state S_0 to some relatively stable steady state S_s, which then undergoes only peripheral modification (say, acquiring new vocabulary items). The attained state incorporates an I-language (it is the state of having or knowing a particular I-language). UG is the theory of S_0; particular grammars are theories of various I-languages. The I-languages that can be attained with S_0 fixed and experience varying are the attainable human languages, where by "language" we now

UG is the theory of S0: *'UG' names the theory, S0 the state it's a theory of, as a grammar is the theory of an I-language. On p. 29 he regrets having used 'UG' for both.*

mean I-language. The steady state has two components that can be distinguished analytically, however, they may be merged and intertwined: a component that is specific to the language in question and the contribution of the initial state. The former constitutes what is "learned"—if this is the appropriate concept to employ in accounting for the transition from the initial to the mature state of the language faculty; it may well not be.[14]

if this is the appropriate concept: *why 'learned' may be wrong: setting a switch UG supplies, triggered by a few sentences, is more like the growth of an organ (p. 2) than like learning facts.*

The system of knowledge attained—the I-language—assigns a status to every relevant physical event, say, every sound wave. Some are sentences with a definite meaning (literal, figurative, or whatever). Some are intelligible with, perhaps, a definite meaning, but are ill-formed in one way or another ("the child seems sleeping"; "to whom did you wonder what to give?" in some dialects; "who do you wonder to whom gave the book?" in all dialects). Some are well formed but unintelligible. Some are assigned a phonetic representation but no more; they are identified as possible sentences of some language, but not mine. Some are mere noise. There are many possibilities. Different I-languages will assign status differently in each of these and other categories. The notion of E-language has no place in this picture. There is no issue of correctness with regard to E-languages, however characterized, because E-languages are mere artifacts. We can define "E-language" in one way or another or not at all, since the concept appears to play no role in the theory of language.

assigns a status: *the point against sets: an E-language takes a string in or leaves it out, the I-language grades, ill-formed but clear, well-formed but senseless, foreign, noise. So the set's boundary is arbitrary.*

Why the shift was right. Realism: S0 and Ss are real brain states, the I-language abstracted directly from Ss, E-languages artefacts (to p. 27). And closeness to the everyday notion (p. 27).

The shift of focus from E- to I-language, reviving and modifying much older traditions, was very much in order. The technical concept of E-language is a dubious one in at least two respects. In the first place, as just observed, languages in this sense are not real-world objects but are artificial, somewhat arbitrary, and perhaps not very interesting constructs. In contrast, the steady state of knowledge attained and the initial state S_0 are real elements of particular mind/brains, aspects of the physical world, where we understand mental states and representations to be physically encoded in some manner. The I-language is abstracted directly as a component of the state attained. Statements about I-language, about the steady state, and about the initial state S_0 are true or false statements about something real and definite, about actual states of the mind/

physically encoded in some manner: *the realism: the state is in the brain however it's coded, so statements about I-language have something to be true of, statements about E-language only a convention.*

brain and their components (under the idealizations already discussed). UG and theories of I-languages, universal and particular grammars, are on a par with scientific theories in other domains; theories of E-languages, if sensible at all, have some different and more obscure status because there is no corresponding real-world object. Linguistics, conceived as the study of I-language and S_0, becomes part of psychology, ultimately biology. Linguistics will be incorporated within the natural sciences insofar as mechanisms are discovered that have the properties revealed in these more abstract studies; indeed, one would expect that these studies will be a necessary step toward serious investigation of mechanisms.[15] To put it differently, E-language, however construed, is further removed from mechanisms than I-language, at a higher order of abstraction. Correspondingly, the concept raises a host of new problems, and it is not at all clear whether they are worth addressing or trying to solve, given the artificial nature of the construct and its apparent uselessness for the theory of language.

a higher order of abstraction: *I-language is one step from the brain, abstracted from a state. E-language is a second step, generated from the I-language under a convention. So it's E-language that is remote from anything real.*

The shift of focus is also, arguably, a shift toward the commonsense notion of language. This matter is less important than the move toward realism and also much less clear, because, as noted, all of these approaches deviate from the commonsense concept in certain respects. But it seems that when we speak of a person as knowing a language, we do not mean that he or she knows an infinite set of sentences, or sound-meaning pairs taken in extension, or a set of acts or behaviors; rather, what we mean is that the person knows what makes sound and meaning relate to one another in a specific way, what makes them "hang together," a particular characterization of a function, perhaps. The person has "a notion of structure" and knows an I-language as characterized by the linguist's grammar. When we say that it is a rule of English that objects follow verbs, as distinct from the rule of Japanese that verbs follow objects, we are not saying that this is a rule of some set of sentences or behaviors, but rather that it is a rule of a system of rules, English, an I-language. The rules of the language are not rules of some infinite set of formal objects or potential actions but are rules that form or constitute the language, like Articles of the Constitution or rules of chess (not a set of moves, but a game, a particular rule

objects follow verbs: *constitutive, as in chess: a rule of chess isn't a generalisation over games but part of the game. A rule of English is likewise part of the system in the head, so John follows R unawares (p. 266).*

system). Of the various technical notions that have been developed in the study of language, the concept of I-language seems closer to the commonsense notion than others.

The shift of perspective from the technical concept E-language to the technical concept I-language taken as the object of inquiry is therefore a shift toward realism in two respects: toward the study of a real object rather than an artificial construct, and toward the study of what we really mean by "a language" or "knowledge of language" in informal usage (again, abstracting from sociopolitical and normative-teleological factors).

Of these two considerations, the first is the clearer and more important. It is not to be expected that the concepts that are appropriate for the description and understanding of some system of the physical world (say, I-language and S_0) will include the sometimes similar concepts of normal discourse, just as the physicist's concepts of energy or mass are not those of ordinary usage. Furthermore, many questions arise about the usage of the intuitive concepts that have no obvious relevance to the inquiry into the nature of the real objects, I-language and S_0. Suppose, for example, that a Martian with a quite different kind of mind/brain were to produce and to understand sentences of English as we do, but as investigation would show, using quite different elements and rules—say, without words, the smallest units being memorized phrases, and with a totally different rule system and UG. Would we then say that the Martian is speaking the same language? Within what limits would we say this? Similar questions arise as to whether an artificial system is exhibiting some form of intelligence or understanding. These may be reasonable questions concerning the intuitive concepts of language and the like in colloquial usage, but it is not clear that they have much bearing on the inquiry into the real-world objects, I-language and the initial state S_0.[16]

a Martian: *the case pulls output and language apart: the Martian and we would share an E-language, every sentence and its meaning, yet have different I-languages. Only the second is a real object.*

The conceptual shift from E-language to I-language, from behavior and its products to the system of knowledge that enters into behavior, was in part obscured by accidents of publishing history, and expository passages taken out of context have given rise to occasional misunderstanding.[17] Some questionable terminological decisions also contributed to misun-

Why the shift was misunderstood, to p. 30: misleading terms, and the model of a formal language, a given set with rules chosen at will. That model is behind Quine and Lewis (p. 20), answered on p. 30.

derstanding. In the literature of generative grammar, the term "language" has regularly been used for E-language in the sense of a set of well-formed sentences, more or less along the lines of Bloomfield's definition of "language" as a "totality of utterances." The term "grammar" was then used with systematic ambiguity, to refer to what we have here called "I-language" and also to the linguist's theory of the I-language; the same was true of the term UG, introduced later with the same systematic ambiguity, referring to S_0 and the theory of S_0. Because the focus of attention was on I-language, E-language being a derivative and largely artificial construct, we find the paradoxical situation that in work devoted to language, the term "language" barely appears. In my 1965 book *Aspects of the Theory of Syntax*, for example, there is no entry for "language" in the index, but many entries under "grammar," generally referring to I-language.

systematic ambiguity: *'grammar' named both the system in the head and the linguist's theory of it, so 'knows the grammar' sounded like knowing a theory. Much philosophical resistance to 'knowledge of grammar' fed on that.*

It would have been preferable to use the term "language" in something closer to the intuitive sense of informal usage; that is, to use the term "language" as a technical term in place of "(generative) grammar" (in the sense of I-language) while adopting some technical term (perhaps "E-language") for what was called "language." The term "(generative) grammar" would then have naturally been used for the linguist's theory of the (I-) language, along the lines of the preceding discussion. Much confusion might have been spared in this way. I suspect that the debate in past years over the alleged problems concerning the concepts grammar and knowledge of grammar may in part be traced to these unfortunate terminological choices, which reinforced inappropriate analogies to the formal sciences and gave rise to the erroneous idea that the study of grammar poses new, complex, and perhaps intractable philosophical issues compared with the study of E-language.[18]

the debate in past years: *the 1970s controversy, answered in Rules and Representations (1980), over whether a speaker can be said to know a grammar, as if he knew a theory. Chapter 4 (p. 266) takes it up with John and Pierre.*

The misleading choice of terms was, in part, a historical accident. The study of generative grammar developed from the confluence of two intellectual traditions: traditional and structuralist grammar, and the study of formal systems. Although there are important precursors, it was not until the mid-1950s that these intellectual currents truly merged, as ideas adapted from the study of formal systems came to be applied to the far more complex systems of natural language in something

approaching their actual richness, and in subsequent years, their actual variety, thus making it possible, really for the first time, to give some substance to Humboldt's aphorism that language involves "the infinite use of finite means," the "finite means" being those that constitute the I-language.

Humboldt's aphorism: *Humboldt, 1836: language 'makes infinite use of finite means'. The finite means are now the I-language, the infinite use the unbounded set of expressions it generates, beyond any list of habits.*

But the study of formal languages was misleading in this regard. When we study, say, the language of arithmetic, we may take it to be a "given" abstract object: an infinite class of sentences in some given notation. Certain expressions in this notation are well-formed sentences, others are not. And of the well-formed sentences, some express arithmetical truths, some do not. A "grammar" for such a system is simply some set of rules that specifies exactly the well-formed sentences. In this case, there is no further question of the correct choice of grammar, and there is no truth or falsity to the matter of choosing among such grammars. Much the same is true of alternative axiomatizations, although in this case we know that none of them will capture exactly the truths. It is easy to see how one might take over from the study of formal languages the idea that the "language" is somehow given as a set of sentences or sentence-meaning pairs, while the grammar is some characterization of this infinite set of objects, hence, it might be thought, a construct that may be selected one way or another depending on convenience or other extraneous concerns. The move is understandable, but misguided, and it has engendered much pointless discussion and controversy.

alternative axiomatizations: *Gödel's theorem: no consistent axiom system that can be listed by rule proves all the arithmetical truths. So none captures its set exactly, yet choosing among them is still a matter of convention.*

Recall Quine's conclusion, cited above (p. 20), that it is senseless to take one grammar rather than another to be "correct" if they are extensionally equivalent, and Lewis's doubts that there is any way "to make objective sense of the assertion that a grammar G is used by a population P whereas another grammar G′, which generates the same language as G, is not." It is quite true that for every E-language, however we choose to define this notion, there are many grammars (i.e., many grammars, each of which is a theory of a particular I-language that, under some convention that one has adopted, determines this E-language). But this is a matter of no consequence. In the case of some formal system, say arithmetic (presumably the model in mind), we assume the class of well-formed formulas in some notation to be "given," and we select the "grammar" (the rules of forma-

The reply to Quine and Lewis (to p. 31): many grammars fit one E-language, granted, and it's of no consequence, since no one is given the E-language. The child gets data and builds an I-language.

tion) as we please. But the E-language is not "given." What is "given" to the child is some finite array of data, on the basis of which the child's mind (incorporating S_0) constructs an I-language that assigns a status to every expression, and that we may think of as generating some E-language under one or another stipulated convention (or we may dispense with this apparently superfluous step). What is given to the linguist are finite arrays of data from various speech communities, including much data not available to the language learner, on the basis of which the linguist will attempt to discover the nature of S_0 and of the particular I-languages attained. The account presented by Quine, Lewis, and others has the story backwards: E-languages are not given, but are derivative, more remote from data and from mechanisms than I-languages and the grammars that are theories of I-languages; the choice of E-language therefore raises a host of new and additional problems beyond those connected with grammar and I-language. Whether it is worthwhile addressing or attempting to solve these problems is not at all clear, because the concept of E-language, however construed, appears to have no significance. The belief that E-language is a fairly clear notion whereas I-language or grammar raises serious, perhaps intractable philosophical problems, is quite mistaken. Just the opposite is true. There are numerous problems concerning the notions I-language and grammar, but not the ones raised in these discussions.

has the story backwards: *why the order matters: if the set came first, any grammar fitting it would do, Quine's point on p. 20. If the I-language comes first, the set is its by-product, and fitting it settles nothing.*

It should be noted that familiar characterizations of "language" as a code or a game point correctly toward I-language, not the artificial construct E-language. A code is not a set of representations but rather a specific system of rules that assigns coded representations to message-representations. Two codes may be different, although extensionally identical in the message-code pairings that they provide. Similarly, a game is not a set of moves but rather the rule system that underlies them. The Saussurean concept of *langue,* although far too narrow in conception, might be interpreted as appropriate in this respect. The same is true of Quine's definition of a language as a "complex of present dispositions to verbal behavior" insofar as it focuses on some internal state rather than E-language, although it is unacceptable for other reasons: Thus, two individuals who speak the same language may differ radically in

extensionally identical: *the code analogy answers Quine: two codebooks can pair every message with the same cipher and still be different codes, so the pairing doesn't fix the code. Nor does its output fix the I-language.*

dispositions to verbal behavior: *Quine, Word and Object (1960). Two faults: speakers of a language differ in dispositions, and response probabilities don't explain how new sentences are understood. Still, it points at an inner state.*

their dispositions to verbal behavior, and if dispositions are characterized in terms of probability of response under given conditions, then it is impossible to identify languages in these terms; and again, the fundamental question of the use and understanding of new sentences is left without any explanation. Perhaps the clearest account is Jespersen's in terms of the "notion of structure" that guides the speaker "in framing sentences of his own...," these being "free expressions."

Back to the analogy story of p. 7, the E-language account of new sentences. Chapter 1 showed it gives wrong readings, here it's shown to presuppose the I-language. The case closes here.

As we have seen, these ideas became the focus of attention in the study of generative grammar, although not without controversy. Saussurean structuralism had placed Jespersen's observation about "free expressions" outside of the scope of the study of language structure, of Saussure's *langue*. Bloomfield (1933) held that when a speaker produces speech forms that he has not heard, "we say that he utters them *on the analogy* of similar forms which he has heard," a position later adopted by Quine, C.F. Hockett, and the few others who even attempted to deal with the problem. This idea is not wrong but rather is vacuous until the concept of analogy is spelled out in a way that explains why certain "analogies" are somehow valid whereas others are not, a task that requires a radically different approach to the whole question. Why, for example, are sentences (6) and (7) of Chapter 1 (p. 8) not understood "on the analogy" of (4) and (5)? Why is sentence (14) not understood "on the analogy" of *any* of the earlier examples, in fact given no interpretation at all? We can give substance to the proposal by explaining "analogy" in terms of I-language, a system of rules and principles that assigns representations of form and meaning to linguistic expressions, but no other way to do so has been proposed; and with this necessary revision in the proposal, it becomes clear that "analogy" is simply an inappropriate concept in the first place.

vacuous: *to say which analogies hold you must state the rules of the I-language, and then analogy does no work. Overlearning has no account of (6), (7) or (14).*

I have been freely using various commonsense notions such as "knowledge," "rule-following," and so forth in this account. Various questions have been raised about the legitimacy of this usage. I will put these questions off for now, returning to them in Chapter 4, but meanwhile continuing to use the terms. I think the usage here is reasonably in accord with common usage, but nothing of great moment is at stake, and one could

"knowledge," "rule-following": *the questions about them are deferred to Chapter 4, which follows this reading, pp. 266 to 267: does John, who follows the adverb rule R, know it, know its source R' and the binding principles?*

Chapter 4, 'Questions about Rules', p. 266, mid-list of facts a speaker knows. Does he know the rules behind them too, R, its source R', the binding principles? Yes, or 'cognize' serves (p. 267).

word *pin*, /p/ is aspirated, whereas in *spin* it is not; that the pronoun may be referentially dependent on *the men* in (9i), but not in the identical phrase in (9ii), and so forth:

(9) (i) I wonder who [the men expected to see them]
(ii) [the men expected to see them]

If these are not instances of knowledge, it is hard to see what is. In this case, the person who knows the language knows these facts, and cognizes them.

cognizes them: *coined on p. 265: whatever we know we cognize, and we also cognize the rules and principles the facts derive from, down to the innate ones. 'Know' stretched to the whole system, awareness dropped.*

Suppose that R is a rule of English grammar that states that verbs cannot be separated from their objects by adverbs, so that in accordance with R, the sentence "I read often the newspaper on Sunday" is unacceptable; rather, we say "I often read the newspaper on Sunday." Suppose that John, a speaker of English, follows the rule, but Pierre, who is learning English, does not and regularly produces and accepts the sentences R marks unacceptable, as in his native French. What we would say, in this case, is that John knows that verbs cannot be separated from their objects by adverbs, but Pierre has not yet learned this and does not know it. Thus John knows that R, but Pierre does not know that R. Of course, we cannot assume that John knows that R holds, obtains, is a rule of his language. John quite probably does not know this, although some linguist may. In other words, there is no legitimate "semantic ascent" from "John knows that R" to "John knows that R holds."

no legitimate "semantic ascent": *Quine's phrase for passing from things to words. Here, from John's knowing that verbs aren't split from objects, which he has, to knowing that R is a rule of English, which only a linguist has.*

Suppose, however, that R is not itself a rule of English but rather a consequence of the rule R′ that states that Case assignment in English observes a strict adjacency requirement: The value for the Case assignment parameter in English is *strict adjacency*. This seems a plausible conclusion, as we have seen. Would it then be proper to say that John knows that R′, but Pierre does not—that is, that John knows, but Pierre has not yet learned and does not know that the value for the Case assignment parameter in English is *strict adjacency*? I think many people would be reluctant to say this of John and Pierre, although I would not. The reason for the difference between attributing knowledge of R and R′ does not lie in the nature of the state of knowledge of John or Pierre, but rather in the familiarity of the notions verb, adverb, and object, which enter into R, as com-

strict adjacency: *Case theory: every noun phrase needs an abstract Case, which a verb gives its object, in English only if adjacent. So 'read often the newspaper' leaves the object Caseless. French is set loosely.*

pared with the unfamiliarity of the notions Case assignment and adjacency parameter, which enter into R′. But this is irrelevant to the description of John's or Pierre's state of knowledge: These states are what they are, independently of our knowledge of linguistic theory.

independently of our knowledge: *our reluctance to credit John with R′ reflects which words we know, not his state. So either 'know' tracks his state and he knows R′, or it tracks our vocabulary and is the wrong word for the theory.*

In fact, there is nothing about the case that is specific to knowledge of language. Suppose that the sun emits light because of processes of fusion internal to the sun. The statement that this is so is not a meaningful statement of my language if I do not know the meaning of the word "fusion," unless we allow, as perhaps we should, that the meaning of the term for me is expressed in terms of the knowledge of others in my speech community (see p. 18). In this case, however, I think we should attribute knowledge that R′ to John, but not Pierre, on the same grounds.

the word "fusion": *the parallel: I can know the sun shines by fusion while 'fusion' gets its meaning for me from the physicists, the division of labour of p. 18. So knowing R′ needn't require the linguist's concepts.*

Assuming this to be correct, let us return to examples (9). Suppose our best theory asserts that speakers know the facts of referential dependence in these cases because their language provides the representations (10i) and (10ii) for (9i) and (9ii), respectively, with the interpretation discussed earlier, and contains the principles assumed earlier—in particular the principles of the binding theory and those determining the presence and character of empty categories in mental representations:

empty categories: *silent parts of structure: e is the trace who leaves when it moves, PRO an infinitive's unspoken subject, here the men. So in (10ii) them shares a clause with a subject meaning the men, and is barred.*

(i) I wonder [who the men expected [*e* to see them]] (10)
(ii) the men expected [PRO to see them]

As we have seen, there is good reason to suppose that something like this is correct. Should we then say that the person who "has" this language "knows the binding theory principles" and so forth? The case is very much like that of R′ in the preceding example, and a positive answer seems consistent with normal usage.

"know" departs from "cognize": *the upshot: whichever way ordinary 'know' goes, the theory needs the term that's about John, not us. If 'know' won't stretch to R′, it's dropped for 'cognize', so the question isn't very important.*

If this reasoning is accepted, then "know" is very much like "cognize." If not, then "know" departs from "cognize" in this respect and is not an appropriate term for the theory of knowledge, because the correct account of the state of knowledge attained by John and Pierre should be independent of what we know; it is a statement about them, not us. I do not think that the question is a very important one. If we continue to use the

REFLECTIONS ON LANGUAGE

NOAM CHOMSKY



PANTHEON BOOKS

A DIVISION OF RANDOM HOUSE, NEW YORK

might be discovered by scientific inquiry over the course of generations, with the intervention of individual genius, with explicit articulation of principles and careful experimentation. This would be possible if the language happened to fall within the bounds of the "science-forming" component of human cognitive capacity. But discovery of the grammar of this language would not be comparable to language learning, just as inquiry in physics is qualitatively different from language learning.

UG will specify properties of sound, meaning, and structural organization. We may expect that in all of these domains, UG will impose conditions that narrowly restrict the variety of languages. For familiar reasons, we cannot conclude from the highly restrictive character of UG that there is a translation procedure of any generality or significance, even in principle (cf. Chomsky, 1965). And quite obviously, nothing is implied about the possibility of translating actual texts, since a speaker or writer naturally presupposes a vast background of unspecified assumptions, beliefs, attitudes, and conventions. The point is perhaps worth noting, since there has been much confusion about the matter. For some discussion, see Keyser (1975).

Reflections, ch. 1: a scientist builds a learning theory for language. The example, to p. 33, shows how a UG principle is found: a rule the child uses that no data chose.

We can gain some insight into UG, hence LT(H,L), whenever we find properties of language that can reasonably be supposed not to have been learned. To make the discussion more concrete, consider a familiar example, perhaps the simplest one that is not entirely trivial. Think of the process of forming questions in English. Imagine again our neutral scientist, observing a child learning English. Suppose that he discovers that the child has learned to form such questions as those of (A), corresponding to the associated declaratives:

our neutral scientist: *Reflections p. 13, before this excerpt: an ideal inquirer with no prior view of the mind, studying the child as any organism. If even he must posit innate structure, innateness isn't a prejudice.*

(A) the man is tall—is the man tall?
the book is on the table—is the book on the table?
etc.

Observing these facts, the scientist might arrive at the following tentative hypothesis as to what the child is doing, assuming now that sentences are analyzed into words:

> *Hypothesis 1:* The child processes the declarative sentence from its first word (i.e., from "left to right"), continuing until he reaches the first occurrence of the word "is" (or others like it: "may," "will," etc.); he then preposes this occurrence of "is," producing the corresponding question (with some concomitant modifications of form that need not concern us).

This hypothesis works quite well. It is also extremely simple. The scientist has every right to be satisfied, and will be able to find a great deal of evidence to support his tentative hypothesis. Of course, the hypothesis is false, as we learn from such examples as (B) and (C):

(B) the man who is tall is in the room—is the man who is tall in the room?

(C) the man who is tall is in the room—is the man who tall is in the room?

Our scientist would discover, surely, that on first presentation with an example such as "the man who is tall is in the room," the child unerringly forms the question (B), not (C) (if he can handle the example at all). Children make many mistakes in language learning, but never mistakes such as exemplified in (C). If the scientist is reasonable, this discovery will surprise him greatly, for it shows that his simple hypothesis 1 is false, and that he must construct a far more complex hypothesis to deal with the facts. The correct hypothesis is the following, ignoring complications that are irrelevant here:

> *Hypothesis 2:* The child analyzes the declarative sentence into abstract phrases; he then locates the first occurrence of "is" (etc.) that follows the first noun

if he can handle the example: *the hedge matters: a young child may not manage relative clauses yet. The claim is only that whenever he does form the question it's (B), never (C). The error the simple rule predicts is never made.*

a far more complex hypothesis: *the point isn't the surprise but that the data don't force the complex rule: all the child heard fits the simple one too. So something other than the data must explain the child's choice, p. 32.*

phrase; he then preposes this occurrence of "is," forming the corresponding question.

Hypothesis 1 holds that the child is employing a "structure-independent rule"—that is, a rule that involves only analysis into words and the property "earliest" ("leftmost") defined on word sequences. Hypothesis 2 holds that the child is employing a "structure-dependent rule," a rule that involves analysis into words and phrases, and the property "earliest" defined on sequences of words analyzed into abstract phrases. The phrases are "abstract" in the sense that neither their boundaries nor their categories (noun phrase, verb phrase, etc.) need be physically marked. Sentences do not appear with brackets, intonation boundaries regularly marking phrases, subscripts identifying the type of phrase, or anything of the sort.

"structure-dependent rule": *a rule stated over phrases. In (B) the first noun phrase is all of 'the man who is tall', so the 'is' after it is the second 'is'. Nothing in the sound marks phrases, so the child supplies them.*

By any reasonable standards, hypothesis 2 is far more complex and "unlikely" than hypothesis 1. The scientist would have to be driven by evidence, such as (B), (C), to postulate hypothesis 2 in place of the simpler and more elementary hypothesis 1. Correspondingly, the scientist must ask why it is that the child unerringly makes use of the structure-dependent rule postulated in hypothesis 2, rather than the simpler structure-independent rule of hypothesis 1. There seems to be no explanation in terms of "communicative efficiency" or similar considerations. It is certainly absurd to argue that children are trained to use the structure-dependent rule, in this case. In fact, the problem never arises in language learning. A person may go through a considerable part of his life without ever facing relevant evidence, but he will have no hesitation in using the structure-dependent rule, even if all of his experience is consistent with hypothesis 1. The only reasonable conclusion is that UG contains the principle that all such rules must be structure-dependent. That is, the child's mind (specifically, its component LT(H,L)) con-

far more complex and "unlikely": *the scientist's own standard: a simpler hypothesis is kept until evidence forces the complex one. The child never meets that evidence yet picks the complex rule, so the preference isn't his.*

the problem never arises in language learning: *'the problem' is the scientist's, choosing between hypotheses 1 and 2 on evidence like (B). The child never faces it, and may go years without such a sentence, yet always takes 2.*

Construct a structure-dependent rule: *an instruction to the learner, not a rule of English: it fixes which rules can be learnt, as the eye's structure fixes what can be seen without being seen. Hence 'a condition for learning'.*

LT(H,L): *defined on Reflections p. 14, before this excerpt: the learning theory for humans (H) in the domain of language (L), what takes a child from data to a grammar. UG is part of it.*

tains the instruction: Construct a structure-dependent rule, ignoring all structure-independent rules. The principle of structure-dependence is not learned, but forms part of the conditions for language learning.

To corroborate this conclusion about UG (hence LT(H,L)), the scientist will ask whether other rules of English are invariably structure-dependent. So far as we know, the answer is positive. If a rule is found that is not structure-dependent, the scientist will be faced with a problem. He will have to inquire further into UG, to discover what additional principles differentiate the two categories of rules, so that the child can know without instruction that one is structure-dependent and the other not. Having gotten this far, the scientist will conclude that other languages must have the same property, on the assumption that humans are not specifically designed to learn one rather than another language, say English rather than Japanese. On this reasonable assumption, the principle of structure-dependence (perhaps, if necessary, qualified as indicated above) must hold universally, if it holds for English. Investigating the consequences of his reasoning, the scientist would discover (so far as we know) that the conclusion is correct.

must hold universally: *the inference to all languages: the principle wasn't learnt from English data, so it comes from the learner, who isn't built for English. A prediction that could fail, so it corroborates UG.*

More complex examples can be produced, but this simple one illustrates the general point. Proceeding in this way, the scientist can develop some rich and interesting hypotheses about UG, hence LT(H,L). Thus, learning theory for humans in the domain of language incorporates the principle of structure-dependence along with other more intricate (and, I should add, more controversial) principles like it. I will return to some of these in the third chapter.

"innateness hypothesis": *that the learner brings innate structure. At issue is only whether some of it is specific to language.*

Keeping this single example of a principle of UG in mind, let us return now to the "innateness hypothesis." Recall that there is no issue as to the necessity for such a hypothesis, only as to its character.