D.E.L.T.A., Vol. 13, Nº Especial, 1997 (1-20)

New Horizons in the Study of Language * **

Noam Chomsky

The study of language is one of the oldest branches of systematic inquiry, tracing back to classical India and Greece, with a rich and fruitful history of achievement. From a different point of view, it is quite young. The major research enterprises of today took shape only about 40 years ago, when some of the leading ideas of the tradition were revived and reconstructed, opening the way to what has proven to be very productive inquiry.

That language should have exercised such fascination over the years is not surprising. The human faculty of language seems to be a true "species property", varying little among humans and without significant analogue elsewhere. Probably the closest analogues are found in insects, at an evolutionary distance of a billion years. There is no serious reason today to challenge the Cartesian view that the ability to use linguistic signs to express freely-formed thoughts marks "the true distinction between man and animal" or machine, whether by "machine" we mean the automata that captured the imagination of the 17th and 18th century, or those that are providing a stimulus to thought and imagination today.

Furthermore, the faculty of language enters crucially into every aspect of human life, thought, and interaction. It is largely responsible for the fact that alone in the biological world, humans have a history, cultural evolution and diversity of any complexity and richness, even biological success in the technical sense that their numbers are huge.

A Martian scientist observing the strange doings on Earth could hardly fail to be struck by the emergence and significance of this apparently unique form of intellectual organization. It is even more natural that the topic, with its

* Talk given at the Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil, on November, 18th, 1996 (Promotion: Linguistics Department of the School of Lettres, UFRJ) and at the Federal University of Pará (UFPa) in November, 29th, 1996 (Promotion: Museu Goeldi/CNPq e BASA. Live-transmission by TV Cultura).

** Thanks are also due to CNPq for financial help supplied for the event, to COPPE/UFRJ for joining the Linguistics Deparment in the organization of the two lectures that took place in our University, to many colleagues in this Department who invested lots of energy in the planning and organization (in a very special way, Carlota Rosa, as well as Edione Trindade Azevedo and Lucinda Brito), to the Director of the Faculdade de Letras Carlos Tannus, for his priceless help.

many mysteries, should have stimulated the curiosity of those who seek to understand their own nature and their place within the wider world.

Human language is based on an elementary property that also seems to be biologically isolated: the property of discrete infinity, which is exhibited in its purest form by the natural numbers 1, 2, 3,... Children do not learn this property; unless the mind already possesses the basic principles, no amount of evidence could provide them. Similarly, no child has to learn that there are three and four word sentences, but no three-and-a half word sentences, and that they go on forever; it is always possible to construct a more complex one, with a definite form and meaning. Such knowledge must come to us from "the original hand of nature", in David Hume's phrase, as part of our biological endowment.

This property intrigued Galileo, who regarded the discovery of a means to communicate our "most secret thoughts to any other person with 24 little characters" as the greatest of all human inventions. The invention succeeds because it reflects the discrete infinity of the language that these characters are used to represent. Shortly after, the authors of the Port Royal Grammar were struck by the "marvelous invention" of a means to construct from a few dozen sounds an infinity of expressions that enable us to reveal to others what we think and imagine and feel — from a contemporary standpoint, not an "invention" but no less "marvelous" as a product of biological evolution, about which virtually nothing is known, in this case.

The faculty of language can reasonably be regarded as a "language organ" in the sense in which scientists speak of the visual system, or immune system, or circulatory system, as organs of the body. Understood in this way, an organ is not something that can be removed from the body, leaving the rest intact. It is a subsystem of a more complex structure.

We hope to understand the full complexity by investigating parts that have distinctive characteristics, and their interactions. Study of the faculty of language proceeds in the same way.

We assume further that the language organ is like others in that its basic character is an expression of the genes. How that happens remains a distant prospect for inquiry, but we can investigate the genetically-determined "initial state" of the language faculty in other ways. Evidently, each language is the

result of the interplay of two factors: the initial state and the course of experience. We can think of the initial state as a "language acquisition device" that takes experience as "input" and gives the language as an "output" — an "output" that is internally represented in the mind/brain. The input and the output are both open to examination: we can study the course of experience and the properties of the languages that are acquired. What is learned in this way can tell us quite a lot about the initial state that mediates between them. Furthermore, there is strong reason to believe that the initial state is common to the species: if my children had grown up in Tokyo, they would speak Japanese, like other children there. That means that evidence about Japanese bears directly on the assumptions concerning the initial state for English. In such ways, it is possible to establish strong empirical conditions that the theory of the initial state must satisfy, and also to pose several problems for the biology of language: How do the genes determine the initial state, and what are the brain mechanisms involved in the initial state and the later states it assumes? These are extremely hard problems, even for much simpler systems where direct experiment is possible, but some may be at the horizons of inquiry.

To proceed we should be more clear about what we mean by "a language". There has been much impassioned controversy about the right answer to this question, and more generally, to the question of how languages should be studied. The controversy is pointless, because there is no right answer. If we are interested in how bees communicate, we will try to learn something about the internal nature of bees, their social arrangements, and their physical environment. These approaches are not in conflict; they are mutually supportive. The same is true of the study of human language: it can be investigated from a biological point of view, and from numerous others: sociolinguistics, language and culture, historical, and so on. Each approach defines the object of its inquiry in the light of its special concerns; and if rational, each will try to learn what it can from other approaches. Why such matters arouse great passion in the study of humans is perhaps an interesting question, but I will put it aside for now.

The approach I have been outlining is concerned with the faculty of language: its initial state, and the states it assumes. Suppose that Peter's language organ is in state L. We can think of L as Peter's "internalized language". When I speak of a language here, that is what I mean. So understood, a language is something like "the way we speak and understand", one traditional conception of language.

Adapting a traditional term to a new framework, we call the theory of Peter's language the "grammar" of his language. Peter's language determines an infinite array of expressions, each with its sound and meaning. In technical terms, Peter's language "generates" the expressions of his language. The theory of his language is therefore called a generative grammar. Each expression is a complex of properties, which provide "instructions" for Peter's performance systems: his articulatory apparatus, his modes of organizing his thoughts, and so on. With his language and the associated performance systems in place, Peter has a vast amount of knowledge about the sound and meaning o expressions, and a corresponding capacity to interpret what he hears, express his thoughts, and use his language in a variety of other ways.

Generative grammar arose in the context of what is often called "the cognitive revolution" of the 1950s, and was an important factor in its development. Whether or not the term "revolution" is appropriate, there was an important change of perspective: from the study of behavior and its products (such as texts), to the inner mechanisms that enter into thought and action.

The cognitive perspective regards behavior and its products not as the object of inquiry, but as data that may provide evidence about the inner mechanisms of mind and the ways these mechanisms operate in executing actions and interpreting experience. The properties and patterns that were the focus of attention in structural linguistics find their place, but as phenomena to be explained along with innumerable others, in terms of the inner mechanisms that generate expressions. The approach is "mentalistic", but in what should be an uncontroversial sense. It is concerned with "mental aspects of the world", which stand alongside its mechanical, chemical, optical and other aspects. It undertakes to study a real object in the natural world — the brain, its states and functions — and thus to move the study of the mind towards eventual integration with the biological sciences.

The "cognitive revolution" renewed and reshaped many of the insights, achievements, and quandaries of what we might call "the first cognitive revolution" of the 17th and 18th century, which was part of the scientific revolution that so radically modified our understanding of the world. It was recognized at the time that language involves "the infinite use of finite means", in von Humboldt's phrase; but the insight could be developed only in limited ways, because the basic ideas remained vague and obscure. By mid-20th

century, advances in the formal sciences had provided appropriate concepts in a very sharp and clear form, making it possible to give a precise account of the computational principles that generate the expressions of a language, and thus to capture, at least partially, the idea of "infinite use of finite means". Other advances also opened the way to investigation of traditional questions with greater hope of success. The study of language change had registered major achievements. Anthropological linguistics provided a far richer understanding of the nature and variety of languages, also undermining many stereotypes. And certain topics, notably the study of sound systems, had been much advanced by the structural linguistics of the 20th century.

The earliest attempts to carry out the program of generative grammar quickly revealed that even in the best studied languages, elementary properties had passed unrecognized, that the most comprehensive traditional grammars and dictionaries only skim the surface.

The basic properties of languages are presupposed throughout, unrecognized and unexpressed. That is quite appropriate if the goal is to help people to learn a second language, to find the conventional meaning and pronunciation of words, or to have some general idea of how languages differ. But if our goal is to understand the language faculty and the states it can assume, we cannot tacitly presuppose "the intelligence of the reader". Rather, this is the object of inquiry.

The study of language acquisition leads to the same conclusion. A careful look at the interpretation of expressions reveals very quickly that from the earliest stages, the child knows vastly more than experience has provided. That is true even of simple words. At peak periods of language growth, a child is acquiring words at a rate of about one an hour, with extremely limited exposure under highly ambiguous conditions. The words are understood in delicate and intricate ways that are far beyond the reach of any dictionary, and are only beginning to be investigated. When we move beyond single words, the conclusion becomes even more dramatic. Language acquisition seems much like the growth of organs generally; it is something that happens to a child, not that the child does. And while the environment plainly matters, the general course of development and the basic features of what emerges are predetermined by the initial state. But the initial state is a common human possession. It must be, then, that in their essential properties and even down to fine detail, languages are cast to the same mold. The Martian scientist might

reasonably conclude that there is a single human language, with differences only at the margins.

For our lives, the slight differences are what matter, not the overwhelming similarities, which we unconsciously take for granted. But if we want to understand what kind of creature we are, we have to adopt a very different point of view, basically that of the Martian studying humans. That is, in fact, the point of view we adopt when we study other organisms, or even humans apart from their mental aspects — humans "below the neck", metaphorically speaking. There is every reason to study what is above the neck the same way.

As languages were more carefully investigated from the point of view of generative grammar, it became clear that their diversity had been underestimated as radically as their complexity and the extent to which they are determined by the initial state of the faculty of language. At the same time, we know that the diversity and complexity can be no more than superficial appearance.

These were surprising conclusions, paradoxical but undeniable. They pose in a stark form what has become the central problem of the modern study of language: How can we show that all languages are variations on a single theme, while at the same time recording faithfully their intricate properties of sound and meaning, superficially diverse? A genuine theory of human language has to satisfy two conditions: "descriptive adequacy" and "explanatory adequacy". The grammar of a particular language satisfies the condition of descriptive adequacy insofar as it gives a full and accurate account of the properties of the language, of what the speaker of the language knows. To satisfy the condition of explanatory adequacy, a theory of language must show how each particular language can be derived from a uniform initial state under the "boundary conditions" set by experience. In this way, it provides an explanation of the properties of languages at a deeper level.

There is a serious tension between these two research tasks. The search for descriptive adequacy seems to lead to ever greater complexity and variety of rule systems, while the search for explanatory adequacy requires that language structure must be invariant, except at the margins. It is this tension that has largely set the guidelines for research. The natural way to resolve the tension is to challenge the traditional assumption, carried over to early generative grammar, that a language is a complex system of rules, each specific to particular languages and particular grammatical constructions: rules for forming relative clauses in Hindi, verb phrases in Bantu, passives in Japanese, and so on.

Considerations of explanatory adequacy indicate that this cannot be correct.

The central problem was to find general properties of rule systems that can be attributed to the faculty of language itself, in the hope that the residue will prove to be more simple and uniform.

About 15 years ago, these efforts crystallized in an approach to language that was a much more radical departure from the tradition than earlier generative grammar had been. This "Principles and Parameters" approach, as it has been called, rejected the concept of rule and grammatical construction entirely: there are no rules for forming relative clauses in Hindi, verb phrases in Bantu, passives in Japanese, and so on. The familiar grammatical constructions are taken to be taxonomic artifacts, useful for informal description perhaps but with no theoretical standing. They have something like the status of "terrestrial mammal" or "household pet". And the rules are decomposed into general principles of the faculty of language, which interact to yield the properties of expressions.

We can think of the initial state of the faculty of language as a fixed network connected to a switch box; the network is constituted of the principles of language, while the switches are the options to be determined by experience. When the switches are set one way, we have Bantu; when they are set another way, we have Japanese. Each possible human language is identified as a particular setting of the switches — a setting of parameters, in technical terminology. If the research program succeeds, we should be able literally to deduce Bantu from one choice of settings, Japanese from another, and so on through the languages that humans can acquire. The empirical conditions of language acquisition require that the switches can be set on the basis of the very limited information that is available to the child. Notice that small changes in switch settings can lead to great apparent variety in output, as the effects proliferate through the system. These are the general properties of language that any genuine theory must capture somehow.

This is, of course, a program, far from a finished product. The conclusions tentatively reached are unlikely to stand in their present form; and, needless to say, one can have no certainty that the whole approach is on the right track. As a research program, however, it has been highly successful, leading to a real explosion of empirical inquiry into languages of a very broad typological range, to new questions that could never even have been formulated before, and to

many intriguing answers. Questions of acquisition, processing, pathology, and others also took new forms, which have proven very productive as well. Furthermore, whatever its fate, the program suggests how the theory of language might satisfy the conflicting conditions of descriptive and explanatory adequacy. It gives at least an outline of a genuine theory of language, really for the first time.

Within this research program, the main task is to discover and clarify the principles and parameters and the manner of their interaction, and to extend the framework to include other aspects of language and its use. While a great deal remains obscure, there has been enough progress to at least consider, perhaps to pursue, some new and more far-reaching questions about the design of language. In particular, we can ask how good is the design. How close does language come to what some super-engineer would construct, given the conditions that the language faculty must satisfy?

The questions have to be sharpened, and there are ways to proceed. The faculty of language is embedded within the broader architecture of the mind/brain. It interacts with other systems, which impose conditions that language must satisfy if it is to be usable at all. We might think of these as "legibility conditions", in the sense that other systems must be able to "read" the expressions of the language and use them as "instructions" for thought and action. The sensorimotor systems, for example, have to be able to read the instructions having to do with sound, the "phonetic representations" generated by the language. The articulatory and perceptual apparatus have specific design that enables them to interpret certain phonetic properties, not others. These systems thus impose legibility conditions on the generative processes of the faculty of language, which must provide expressions with the proper phonetic form. The same is true of conceptual and other systems that make use of the resources of the faculty of language: they have their intrinsic properties, which require that the expressions generated by the language have certain kinds of "semantic representations", not others. We may therefore ask to what extent language is a "good solution" to the legibility conditions imposed by the external systems with which it interacts. Until quite recently, this question could not seriously be posed, even formulated sensibly. Now it seems that it can, and there are even indications that language faculty maybe close to "perfect" in this sense, a surprising conclusion, if true.

What has come to be called "the minimalist program" is an effort to

explore these questions. It is too soon to offer a firm judgment about the project. My own judgment is that the questions can now profitably be placed on the agenda, and that early results are promising. I would like to say a few words about the ideas and the prospects, and then to return to some problems that remain at the horizons.

The minimalist program requires that we subject conventional assumptions to careful scrutiny. The most venerable of these is that language has sound and meaning. In current terms, that translates in a natural way to the thesis that the faculty of language engages other systems of the mind/brain at two "interface levels", one related to sound, the other to meaning. A particular expression generated by the language contains a phonetic representation that is legible to the sensorimotor systems, and a semantic representation that is legible to conceptual and other systems of thought and action.

If this is correct, we next have to ask just where the interface is located. On the sound side, it has to be determined to what extent, if any, sensorimotor systems are language-specific, hence within the faculty of language; there is considerable disagreement about the matter. On the meaning side, the questions have to do with the relations between the faculty of language and other cognitive systems – the relations between language and thought. On the sound side, the questions have been studied intensively with sophisticated technology for half a century, but the problems are hard, and understanding remains limited. On the meaning side, the questions are much more obscure. Far less is known about the language-external systems; much of the evidence about them is so closely linked to language that it is notoriously difficult to determine when it bears on language, when on other systems (insofar as they are distinct). And direct investigation of the kind possible for sensorimotor systems is in its infancy. Nonetheless, there is a huge amount of data about how expressions are used and understood in particular circumstances, enough so that natural language semantics is one of the liveliest areas of study of language, and we can make at least some plausible guesses about the nature of the interface level and the legibility conditions it must meet.

With some tentative assumptions about the interface, we can proceed to further questions. We ask how much of what we are attributing to the faculty of language is really motivated by empirical evidence, and how much is a kind of technology, adopted in order to present data in a convenient form while covering up gaps of understanding. Not infrequently, accounts that are

offered in technical work turn out on investigation to be of roughly the order of complexity of what is to be explained, and involve assumptions that are not independently very well-grounded. That is not problematic as long as we do not mislead ourselves into thinking that useful and informative descriptions, which may provide stepping stones for further inquiry, are something more than that.

Such questions are always appropriate in principle, but often not worth posing in practice; they may be premature, because understanding is just too limited. Even in the hard sciences, in fact even mathematics, questions of this kind have commonly been put to the side. But the questions are nevertheless real, and with a more plausible concept of the general character of language at hand, perhaps worth exploring.

Let us turn to the question of optimality of language design: How good a solution is language to the general conditions imposed by the architecture of the mind/brain? This question too might be premature, but unlike the problem of distinguishing between principled assumptions and descriptive technology, it might have no answer at all. There is no strong reason to believe that a biological system should be well-designed in anything like this sense. To the extent that it is, the conclusion is surprising, therefore interesting, perhaps another curious respect in which the faculty of language is biologically isolated.

Despite the initial implausibility, let us tentatively assume that both of these questions are appropriate ones, in practice as well as principle. We now proceed to subject postulated principles of language to close scrutiny to see if they are empirically justified in terms of legibility conditions. I will mention a few examples, apologizing in advance for the use of some technical terminology, which I'll try to keep to a minimum, but have no time here to explain in any satisfactory way.

One question is whether there are levels other than the interface levels: Are there levels "internal" to the language, in particular, the levels of deep and surface structure that have been postulated in modern work? The minimalist program seeks to show that everything that has been accounted for in terms of these levels has been misdescribed, and is as well or better understood in terms of legibility conditions at the interface: for those of you who know the technical literature, that means the projection principle, binding theory, Case theory, the chain condition, and so on.

We also try to show that the only computational operations are those that are unavoidable on the weakest assumptions about interface properties. One such assumption is that there are word-like units: the external systems have to be able to interpret such items as "Peter" and "tall". Another is that these items are organized into larger expressions, such as "Peter is tall". A third is that the items have properties of sound and meaning: the word "Peter" begins with closure of the lips and is used to refer to persons. The language therefore involves three kinds of elements: the properties of sound and meaning, called "features"; the items that are assembled from these properties, called "lexical items"; and the complex expressions constructed from these "atomic" units. It follows that the computational system that generates expressions has two basic operations: one assembles features into lexical items, the second forms larger syntactic objects out of those already constructed, beginning with lexical items.

We can think of the first operation as essentially a list of lexical items. In traditional terms, this list, called the lexicon, is the list of "exceptions", arbitrary associations of sound and meaning and particular choices among the inflectional properties made available by the faculty of language that determine how we indicate that nouns and verbs are plural or singular, that nouns have nominative or accusative case, and so on. These inflectional features turn out to play a central role in computation.

Optimal design would introduce no new features in the course of computation. There should be no indices or phrasal units and no bar levels (hence no phrase structure rules or X-bar theory). We also try to show that no structural relations are invoked other than those forced by legibility conditions or induced in some natural way by the computation itself. In the first category we have such properties as adjacency at the phonetic level, and argument-structure and quantifier-variable relations at the semantic level. In the second category, we have very local relations between features, and elementary relations between two syntactic objects joined together in the course of computation: the relation holding between one of these and the parts of the other is the relation of c-command, as Samuel Epstein has pointed out, a notion that plays a central role throughout language design and has been regarded as highly unnatural, though it falls into place in a natural way from this perspective. But we exclude government, binding relations internal to the derivation of expressions, and a variety of other relations and interactions.

As anyone familiar with recent work will be aware, there is ample empirical evidence to support the opposite conclusion throughout. Worse yet, a core assumption of the work within the Principles-and-Parameters framework, and its fairly impressive achievements, is that everything I have just proposed is false — that language is highly "imperfect" in these respects, as might well be expected. So it is no small task to show that such apparatus is eliminable as unwanted descriptive technology; or even better, that descriptive and explanatory force are extended if such "excess baggage" is shed. Nevertheless, I think that work of the past few years suggests that these conclusions, which seemed out of the question a few years ago, are at least plausible, quite possibly correct.

Languages plainly differ, and we want to know how. One respect is in choice of sounds, which vary within a certain range. Another is in the association of sound and meaning, essentially arbitrary. These are straightforward and need not detain us. More interesting is the fact that languages differ in inflectional systems: case systems, for example. We find that these are fairly rich in Latin, even more so in Sanskrit or Finnish, but minimal in English and invisible in Chinese. Or so 0it appears; considerations of explanatory adequacy suggest that here too appearance may be misleading, and in fact, recent work indicates that these systems vary much less than appears to be the case from the surface forms. Chinese and English, for example, may have the same case system as Latin, but the phonetic realization is different. Furthermore, it seems that much of the variety of language can be reduced to properties of inflectional systems. If this is correct, then language variation is located in a narrow part of the lexicon.

Legibility conditions impose a three-way division among the features assembled into lexical items:

(1) semantic features, interpreted at the semantic interface
(2) phonetic features, interpreted at the phonetic interface
(3) features that are not interpreted at either interface

Independently, features are subdivided into the "formal features" that are used by the syntactic operations and others that are not. A natural principle that would sharply restrict language variation is that only inflectional properties are formal features. That seems to be correct, an important matter that I will have to leave to the side.

In a perfectly designed language, each feature would be semantic or phonetic, not merely a device to create a position or to facilitate computation. If so, there are no uninterpretable formal features. That is too strong a requirement, it seems. Such prototypical formal features as structural case — Latin nominative and accusative, for example — have no interpretation at the semantic interface, and need not be expressed at the phonetic level. We can therefore consider a weaker requirement approaching optimal design: each feature is interpreted at the semantic interface or is accessible to the component of the grammar that assigns phonetic form to a syntactic object, the phonological component, which may (and sometimes does) use the feature in question to determine the phonetic representation. Let's assume that weaker condition to hold.

In the syntactic computation, there seems to be a second and more dramatic imperfection in language design, at least an apparent one: the "displacement property" that is a pervasive aspect of language: phrases are interpreted as if they were in a different position in the expression, where similar items sometimes do appear and are interpreted in terms of natural local relations. Take the sentence "Clinton seems to have been elected". We understand the relation of "elect" and "Clinton" as we do when they are locally related in the sentence "It seems that they elected Clinton": "Clinton" is the direct object of "elect", in traditional terms, though "displaced" to the position of subject of "seems": the subject and verb agree in inflectional features in this case, but have no semantic relation; the semantic relation of the subject is to the remote verb "elect".

We now have two "imperfections": uninterpretable formal features, and the displacement property. On the assumption of optimal design, we would expect them to be related, and that seems to be the case: uninterpretable formal features are the mechanism that implements the displacement property.

The displacement property is never built into the symbolic systems that are designed for special purposes, called "languages" or "formal languages" in a metaphoric usage: "the language of arithmetic", or "computer languages", or "the languages of science". These systems also have no inflectional systems, hence no uninterpreted formal features. Displacement and inflection are special properties of human language, among the many that are ignored when symbolic systems are designed for other purposes, which may disregard the legibility

conditions imposed on human language by the architecture of the mind/brain.

The displacement property of human language is expressed in terms of grammatical transformations or by some other device, but it is always expressed somehow. Why language should have this property is an interesting question, which has been discussed for almost 40 years without resolution.

My suspicion is that part of the reason has to do with phenomena that have been described in terms of surface structure interpretation, many of them familiar from traditional grammar: topic-comment, specificity, new and old information, the agentive force that we find even in displaced position, and so on. If that is correct, then the displacement property is indeed forced by legibility conditions: it is motivated by interpretive requirements that are externally imposed by our systems of thought, which have these special properties, so the study of language use indicates. These questions are currently being investigated in interesting ways, which I cannot go into here.

From the origins of generative grammar, the computational operations were assumed to be of two kinds: phrase structure rules that form larger syntactic objects from lexical items, and transformational rules that express the displacement property. Both have traditional roots, but it was quickly found that they differ substantially from what had been supposed, with unsuspected variety and complexity. The research program sought to show that the complexity and variety are only apparent, and that the two kinds of rules can be reduced to simpler form. A "perfect" solution to the problem of variety of phrase structure rules would be to eliminate them entirely in favor of the irreducible operation that takes two objects already formed and attaches one to the other, forming a larger object with just the properties of the target of attachment: the operation we can call Merge. That goal may well be attainable, recent work indicates.

The optimal computational procedure, then, consists of the operation Merge and operations to construct the displacement property: transformational operations or some counterpart. The second of the two parallel endeavors sought to reduce the transformational component to the simplest form, though unlike phrase structure rules, it seems to be ineliminable. The end result was the thesis that for a core set of phenomena, there is just a single operation Move — basically, move anything anywhere, with no properties specific to languages or particular constructions. How it applies is determined by general principles

interacting with the specific parameter choices — switch settings — that determine a particular language. The operation Merge takes two distinct objects X and Y and attaches Y to X. The operation Move takes a single object X and an object Y that is part of X, and merges Y to X. The syntactic object formed in this way includes what is called a CHAIN consisting of the two occurrences of Y; the occurrence in the original position is called the TRACE.

The next problem is to show that it is indeed the case that uninterpretable formal features are the mechanism that implements the displacement property, so that the two basic imperfections of the computational system reduce to one. And if it turns out that the displacement property is motivated by legibility conditions imposed by external systems of thought, as I just suggested, then the imperfections are eliminated completely and language design turns out to be optimal after all: uninterpreted formal features are required as a mechanism to satisfy a legibility condition imposed by the general architecture of the mind/brain.

The way this unification proceeds is quite simple, but to explain it coherently would go beyond the scope of these remarks. The basic intuitive idea is that uninterpretable formal features have to be erased to satisfy the interface condition, and erasure requires a local relation between the offending feature and a matching feature that can erase it.

Typically these two features are remote from one another for reasons having to do with the way semantic interpretation proceeds. For example, in the sentence "Clinton seems to have been elected", semantic interpretation requires that "elect" and "Clinton" be locally related in the phrase "elect Clinton" for the construction to be properly interpreted, as if the sentence were actually "seems to have been elected Clinton". The main verb of the sentence, "seems", has inflectional features that are uninterpretable: it is singular/third person/masculine, properties that add nothing independent to the meaning of the sentence, since they are already expressed in the noun phrase that agrees with it, and are ineliminable there. These offending features of "seems" therefore have to be erased in a local relation, an explicit version of the traditional descriptive category of "agreement". To achieve this result, the matching features of the agreeing phrase "Clinton" are attracted by the offending features of the main verb "seems", which are then erased under local matching. But now the phrase "Clinton" is displaced.

Note that only the FEATURES of "Clinton" are attracted; the full phrase moves for reasons having to do with the sensorimotor system, which is unable to "pronounce" or "hear" isolated features separated from the phrase in which they belong. But if the sensorimotor system is inactivated for some reason, then the features alone raise, and alongside of such sentences as "an unpopular candidate seems to have been elected", with overt displacement, we have sentences of the form "seems to have been elected an unpopular candidate"; here the remote phrase "an unpopular candidate" agrees with the verb "seems", which means that its features have been attracted to a local relation with "seem" while leaving the rest of the phrase behind. The reason is that the sensorimotor system has been inactivated in this case, called "covert movement", a phenomenon with quite interesting properties. In many languages, Spanish for example, there are such sentences. English has them too, though it is necessary for other reasons to introduce the semantically empty element "there", giving the sentence "there seems to have been elected an unpopular candidate"; and also, for quite interesting reasons, to carry out an inversion of order, so it comes out "there seems to have been an unpopular candidate elected". These properties follow from specific choices of parameters, which have effects through the languages generally and interact to give acomplex array of phenomena, only superficially distinct. In the case we are looking at, all reduce to the simple fact that uninterpretable formal features must be erased in a local relation with a matching feature, yielding the displacement property required for semantic interpretation at the interface.

Combining these various ideas, some still highly speculative, we can envisage both a motivation and a trigger for the displacement property. Note that these have to be distinguished. An embriologist studying the development of the eye may take note of the fact that for an organism to survive it would be helpful for the lens to contain something that protects it from damage and something that refracts light, and looking further would discover that crystalline proteins have both these properties and also seem to be ubiquitous components of the lens of the eye, showing up on independent evolutionary paths. The first property has to do with "motivation" or "functional design", the second with the trigger that yields the right functional design. There is an indirect and important relation between them, but it would be an error to confound them. Thus, a biologist accepting all of this would not offer the functional design property as the mechanism of embriological development of the eye.

Similarly, we do not want to confound functional motivations for

properties of language with the specific mechanisms that implement them. Similarly, we do not want to confound the fact the displacement property is required by external systems at the interface level with the mechanisms of the operations ATTRACT and its reflex.

There is a fair amount of handwaving in this brief description. Filling in the blanks yields a rather interesting picture, with many ramifications in typologically different languages. But to go on would take us well beyond the scope of these remarks.

I'd like to finish with at least brief reference to other issues, having to do with the ways the internalist study of language relates to the external world. For simplicity, let's keep to simple words. Suppose that "book" is a word in Peter's lexicon. The word is a complex of properties, phonetic and semantic. The sensorimotor systems use the phonetic properties for articulation and perception, relating them to external events: motions of molecules, for example. Other systems of mind use the semantic properties of the word when Peter talks about the world and interprets what others say about it.

There is no far-reaching controversy about how to proceed on the sound side, but on the meaning side there are profound disagreements, or so it appears; some at least may dissolve on closer inspection. Empirically-oriented studies seem to me to approach problems of meaning rather in the way they study sound, as in phonology and phonetics. They try to find the semantic properties of the word "book": that it is nominal not verbal, used to refer to an artifact not a substance like water or an abstraction like health, and so on. One might ask whether these properties are part of the meaning of the word "book" or of the concept associated with the word; on current understanding, there is no good way to distinguish these proposals, but perhaps some day an empirical issue will be unearthed. Either way, some features of the lexical item "book" that are internal to it determine modes of interpretation of the kind just mentioned.

Investigating language use, we find that words are interpreted in terms of such factors as material constitution, design, intended and characteristic use, institutional role, and so on. Things are identified and assigned to categories in terms of such properties, which I am taking to be semantic features, on a par with phonetic features that determine its sound. The use of language can attend in various ways to these semantic features. Suppose the library has two copies of Tolstoy's WAR AND PEACE, Peter takes out one, and John the

other. Did Peter and John take out the same book, or different books? If we attend to the material factor of the lexical item, they took out different books; if we focus on its abstract component, they took out the same book. We can attend to both material and abstract factors simultaneously, as when we say that "the book that he is planning will weigh at least five pounds if he ever writes it", or "his book is in every store in the country". Similarly, we can paint the door white and walk through it, using the pronoun "it" to refer ambiguously to figure and ground. We can report that the bank was blown up after it raised the interest rate, or that it raised the rate to keep from being blown up. Here the pronoun "it" and the "empty category" that is the subject of "being blown up", simultaneously adopt both the material and institutional factors.

The same is true if my house is destroyed and I re-build it, perhaps somewhere else; it is not the same house, even if I use the same materials, though I RE-built IT. The referential terms "re" and "it" cross the boundary. Cities are still different. London could be destroyed by fire and IT could rebuilt somewhere else, from completely different materials, and still be London. Carthage could be rebuilt today, and still be Carthage. Suppose I tell you that I used to think that Constantinople and Istanbul were different cities, but now I know they are the same, and then add that Istanbul will have to be moved somewhere else, so that Constantinople won't have an Islamic character; IT will have to be moved, and RE-built elsewhere, though remaining the SAME city, somehow. Such usage is perfectly intelligible; I have found even stranger examples in ordinary speech and writing, and these comments barely touch the surface of what we find when we begin to look closely at the meanings of words.

The facts about such matters are often clear, but not trivial. Thus referentially dependent elements, even the most narrowly constrained, observe some distinctions but ignore others, in ways that vary for different types of words in curious ways. Such properties can be investigated in many ways: language acquisition, generality among languages, invented forms, etc. What we discover is surprisingly intricate; and not surprisingly, known in advance of any evidence, hence shared among languages. There is no a priori reason to expect that human language will have such properties; Martian could be different. The symbolic systems of science and mathematics surely are. No one knows to what extent the specific properties of human language are a consequence of general biochemical laws applying to objects with general features of the brain, another important problem at a still distant horizon. An

approach to semantic interpretation in similar terms was developed in interesting ways in 17th-18th century philosophy, often adopting Hume's principle that the "identity which we ascribe" to things is "only a fictitious one", established by the human understanding. Hume's conclusion is very plausible. The book on my desk does not have these strange properties by virtue of its internal constitution; rather, by virtue of the way people think, and the meanings of the terms in which these thoughts are expressed. The semantic properties of words are used to think and talk about the world in terms of the perspectives made available by the resources of the mind, rather in the way phonetic interpretation seems to proceed.

Contemporary philosophy of language follows a different course. It asks to what a word refers, giving various answers. But the question has no clear meaning. The example of "book" is typical. It makes little sense to ask to what THING the expression "Tolstoy's WAR AND PEACE" refers, when Peter and John take identical copies out of the library. The answer depends on how the semantic features are used when we think and talk, one way or another. The observations extend to the simplest referential and referentially dependent elements (pronouns, empty categories, "same" etc.). And to proper names, which have rich semantic-conceptual properties. Something is named as a person, a river, a city, with the complexity of understanding that goes along with these categories. Language has no logically proper names, stripped of such properties, as Oxford philosopher Peter Strawson pointed out many years ago. In general, a word, even of the simplest kind, does not pick out an entity of the world, or of our "belief space" — which is not to deny, of course, that there are books and banks, or that we are talking about something if we discuss the fate of the earth and conclude that IT is grim. But we should follow the good advice of the 18th century philosopher Thomas Reid and his modern successors, Ludwig Wittgenstein and others, and not draw unwarranted conclusions from common usage.

We can, if we like, say that the word "book" refers to books, "sky" to the sky, "health" to health, and so on. Such conventions basically express lack of interest in how words are used to talk about things, and about their semantics. The conventions raise other problems and involve what seem to me dubious assumptions, another important topic that I cannot hope to go into here.

I mentioned that modern generative grammar has sought to address concerns that animated the tradition, in particular, the Cartesian idea that "the

true distinction" between humans and other creatures or machines is the ability to act in the manner they took to be most clearly illustrated in the ordinary use of language: without any finite limits, influenced but not determined by internal state, appropriate to situations but not caused by them, coherent and evoking thoughts that the hearer might have expressed, and so on. The goal of the work I have been discussing is to unearth some of the factors that enter into such normal practice. Only SOME of these, however.

Generative grammar seeks to discover the mechanisms that are used, thus contributing to the study of HOW they are used in the creative fashion of normal life. How they are used is the problem that intrigued the Cartesians, and it remains as mysterious to us as it was to them, even though far more is understood today about the mechanisms that are involved.

In this respect, the study of language is again much like that of other organs. Study of the visual and motor systems has uncovered mechanisms by which the brain interprets scattered stimuli as a cube and the arm reaches for a book on the table. But these branches of science do not raise the question of how people decide to look at a book on the table or to pick it up, and speculations about the use of the visual or motor systems, or others, amount to very little. It is these capacities, manifested most strikingly in language use, that are at the heart of traditional concerns: for Descartes, they are "the noblest thing we can have" and all that "truly belongs" to us. Half a century before Descartes, the Spanish philosopher-physician Juan Huarte observed that this "generative faculty" of ordinary human understanding and action is foreign to "beasts and plants", though it is a lower form of understanding that falls short of true exercise of the creative imagination. Even the lower form lies beyond our theoretical reach, apart from the study of mechanisms that enter into it.

In a number of areas, language included, a lot has been learned in recent years about these mechanisms. The problems that can now be faced are hard and challenging, but many mysteries still lie beyond the reach of the form of human inquiry we call "science", a conclusion that we should not find surprising if we consider humans to be part of the organic world, and perhaps one we should not find distressing either.