LANGUAGE AND MIND:
CHALLENGES AND PROSPECTS

Noam Avram Chomsky

Inevitably, I suppose, on an occasion such as this, one is led to reminisce and to speculate, to reflect on the course of the thought and efforts of a community of inquiry over many years, the challenges and prospects, the false starts and fallacies, the partial solutions and glimmerings of understanding, the problems that remain unsolved and the mysteries that seem to lie beyond our intellectual grasp. A retrospective assessment is, of course, very different from intellectual history. And an attempt to picture what the future might hold is sure to be partial at best, more likely seriously misleading. Speaking of the most advanced sciences, Edward Witten of the Princeton Institute for Advanced Studies observes that, “the progress of physics has always been such that the level of understanding for which one generation aims wasn’t even dreamed of a generation or two earlier.” Something similar is true of the cognitive sciences, which are only beginning to enter into something like a Galilean revolution, to place the most hopeful cast on the developments of the past generation of research. In the study of language, for example, my students are now addressing problems that could not even be formulated, and were not envisaged, when I began to study these fields 40 years ago. Recognizing the pitfalls, I will, nonetheless, attempt a retrospective assessment and some speculations as to prospects and limits of rational inquiry into problems of language and mind.

How did the project of understanding the world appear to a student reaching intellectual maturity in the immediate postwar period? In the early 20th century, the concepts of fundamental physics had been radically modified, leading to hopes for authentic unity of science. Quantum theory explained “most of physics and all of chemistry” so that “physics and chemistry have been fused into complete oneness.,” Paul Dirac and Werner Heisenberg observed. If only the apparent conflicts of general relativity and quantum theory could be overcome, so physicists informed us, we might even go forward to what some today call a “theory of everything.” The discoveries of early genetics were accommodated within known biochemistry, eliminating the last vestige of vitalism from scientific biology and offering the hope that the evolution and growth of living organisms might fall within the compass of the unified natural sciences as well. The next scientific frontier was naturally assumed to be the human mind and its manifestations in thought and action, judgment and evaluation, creation and understanding.

There was, at the time, considerable optimism about the prospects for this next leap into the unknown. Many felt that with Claude Shannon's insights, information theory would provide unifying concepts for the study of language and mind. The computer age was dawning, and these and other technological marvels provided a stimulus to the scientific imagination, much as the automata of the 17th and 18th centuries suggested approaches to mind, physiology and behavior. It seemed possible that ideas developing in the computer sciences would provide the models of mind required to understand its mysteries. The behavioral sciences were in the ascendant. Many felt the most complex and intricate human capacities could be resolved into systems of habit and skill, explained in terms of paradigms of conditioning that were believed to account for animal behavior. Particularly in Cambridge Massachusetts, where I was then a graduate student, there was a degree of euphoria about all of these matters.

All of this might be seen in a broader sociopolitical context. The war years had produced atrocities and horrors of a scale and character beyond the worst imaginings of the most gloomy prophets. From the ashes of a world devastated or severely wounded, the United States emerged in a position of wealth and power that had no historical analogue, in the first truly global system. There were awesome threats, but also it seemed, endless prospects for "a world restored" under American leadership. I think there is little doubt that the general mood infected the scientific and scholarly communities as well.

Some were less sanguine. My own feeling at the time was one of considerable skepticism about prevailing ideas in the behavioral sciences and the structuralist tendencies that developed in uneasy relation to them. And about the hopes placed in the new technology and information sciences, significant as these developments undoubtedly were. In particular, I was one of those who felt the attempt to account for human capacities within the framework of habits, skills, conditioning and reinforcement, or even in terms of the more abstract models of the information sciences, were misguided at their core and would not succeed; and that the hopes for developing the unified science that all sought lie along quite different paths. In my opinion, subsequent work showed that this skepticism was in order, though the issues are far from settled, and have arisen again in the context of study of connectionist models of mind.

The so-called “cognitive revolution” of the mid-1950s has been variously understood by those who participated in it. As I understand it, this “revolution” offered a shift of perspective with regard to the problems of language and mind: from behavior and the products of behavior (for example, utterances, discourses and texts), to the inner mechanisms of mind that underlie behavior and determine its specific form and character, and how it is interpreted and understood. Behavior and its products are no longer taken to be the object of investigation. Rather, they provide data, which may serve as one kind of evidence, alongside of others, for the study of what really concerns us: the inner mechanisms of mind.

With regard to language, an alternative was counterposed to the standard conception in terms of habits, abilities, dispositions, skills, patterns, and structures. The alternative was a computational-representational theory of mind: the mind, using its internal mechanisms, forms and manipulates representations, and uses them in determining and executing actions and in interpreting experience. The avowed mentalism should be understood as a step towards integrating the study of language and other aspects of psychology within the natural sciences; a step towards the unification of the sciences.

We might think of 19th century chemistry as the study of the properties of as yet unknown physical mechanisms, expressing its principles and descriptions in terms of such abstract notions as chemical elements, valence, the structure of organic molecules, the Periodic Table, and so on. This abstract study set the stage for the subsequent inquiry into “more fundamental” entities that exhibit the properties formulated at the abstract level of inquiry. The same may be said of early genetics. Correspondingly, the study of computational-representational theories of mind and their role in action and understanding should serve as a guide for the emerging brain sciences, providing them with an analysis of the conditions that the mechanisms sought must satisfy.

Suppose in the course of time, as we hope and anticipate, neural mechanisms are discovered that exhibit the properties and satisfy the principles formulated in the cognitive sciences in terms of such entities as the rules and representations of language. We will not conclude that these entities do not exist. Any more than the unification of chemistry, parts of biology and physics shows that there are no chemical elements and ions, genes and alleles, tables and chairs, continents and galaxies, plants and animals, or persons understood in the highly abstract terms in which we conceive them with the curious identity conditions they satisfy on separation from environment, spatio-temporal contiguity, and psychic persistence. Some of these theoretical constructions may be shown to be misguided with the unification of science, going the way of phlogiston and vital forces; others may be sharpened and modified in the course of this unification. But generally, we expect, progress in establishing the links among various levels of rational inquiry will provide deeper understanding of entities of which we have only a partial grasp when we are limited to the more abstract level of theory construction and explanation of observed phenomena.

The "cognitive revolution" of the 1950s was not as novel as many of its practitioners assumed it to be. In significant respects, it recapitulated leading ideas of what we might call "the first cognitive revolution" of the 17th century. Descartes and his followers also developed a computational-representational theory of mind, extended in interesting ways into the 19th century. This was particularly the case in the areas of vision and language, just those that have progressed most rapidly within the more recent revival of these ideas.

The significance of the work of these cognitive scientists—as we might call them—was not appreciated at the outset of the "second cognitive revolution" in the mid-1950s, and is still misunderstood. Without expanding on their actual contributions, we might nevertheless identify three crucial problems that were raised, which I will call "Plato's problem," "Descartes's problem," and "Humboldt's problem."

By "Plato's problem" I refer to the problem posed in the *Meno*, as Socrates demonstrates that a slave boy with no prior instruction nevertheless knows the principles of geometry. The general question raised is how we can know so much, given that we have so little basis for this knowledge in experience; it is a problem that is far more serious than often believed. Plato's answer was that we know so much because we remember it from an earlier existence. The answer requires a mechanism, and for those unsatisfied with the concept of an immortal soul, Plato's answer must be "purged of the error of preexistence," as Leibniz phrased the task. In more modern terms, our version of Plato's answer would be we know so much because our mind/brain, by virtue of the human genetic endowment, is constructed in such a way as to develop certain cognitive systems, not others. In particular, the human biological endowment determines the basic properties of subsystems of mind that we may call "the language faculty" and the conceptual faculty. These permit a certain range of possible linguistic and conceptual systems, not others, and determine how they may be linked. The problem faced by the child is to determine, from presented data, which of these systems is the one of the community, a task that will be feasible if the range of permissible options is not too large and varied. Having performed this task, the child's language faculty incorporates a computational system that forms and modifies representations of utterances, providing the basis for speech and perception and linked to a conceptual system that provides the means for thought, construction of experience, and interpretation of it.

"Descartes's problem" has to do with the use of language, and with human action more generally. Descartes was properly impressed by the fact that the normal use of language has a "creative aspect": it is unbounded, neither determined by stimuli nor random, coherent and appropriate to situations, evoking thoughts that the hearer might have expressed in a similar way. Descartes argued that these properties lie beyond the bounds of "machines," understood to be objects subject to the principles of what we might call "contact mechanics": pushing, pulling, etc., through direct contact. Hence, the creative aspect of language use required some new principle, outside of mechanics, so conceived. Within Cartesian metaphysics, this required postulation of a second substance, a *res cagitans,* standing apart from body, which is subject to contact mechanics. The two substances must then interact in some manner (that was debated in subsequent years).

In these terms, the Cartesians proposed tests for the existence of other minds, generally formulated in terms of the creative aspect of language use: thus if an experiment convinces us that another organism exhibits these properties, it would only be reasonable to assume that it has a mind like ours (as for ourselves, introspection provides the relevant evidence for existence of mind as distinct from body). Similar ideas were reinvented in more recent years, particularly by the British mathematician Alan Turing, and are now called "the Turing test," a criterion to determine whether a machine "exhibits intelligence."

What I will call "Humboldt's problem" is based on observations similar to Descartes's. Wilhelm von Humboldt observed that language is a system that provides for infinite use of finite means. We may take these finite means to constitute a particular language; to know the language is to have these finite means represented in the mind/brain. Crucially, Humboldt regarded language not as a set of constructed objects—expressions, utterances or speech acts—but rather as a process of generation. Language is *eine Erzeugung* (a process of generation), not *ein todtes Erzeugtes* (the "dead" objects generated). With a bit of interpretive license, we might understand him to be saying that a language is a generative procedure that enables articulated, structured expressions of thought to be freely produced and understood.

Notice there is interpretive license in this account. In the early 19th century, one could not clearly distinguish between, on the one hand, an abstract generative procedure that assigns structural descriptions to all expressions, and on the other, the actual "work of the mind" ("Arbeit des Geistes") that brings thought to expression in linguistic performance. There are passages in Humboldt's writings that suggest one or the other interpretation, sometimes with fair explicitness, but to attempt to determine which notion he had in mind is an error, since the two concepts were not clearly distinguished, and could hardly have been, the relevant concepts lacking. A century later, progress in the formal sciences made it possible to formulate Humboldt's problem quite explicitly, as a problem for substantive and productive inquiry.

The conception of generative grammar that developed in the 1950s crucially, and properly, distinguishes the process of production of speech from the abstract process of generation that constitutes linguistic knowledge. We thus distinguish performance from competence (in the sense of possession of knowledge), construing knowledge of language as incorporation in the mind/brain of a generative procedure taken in the abstract sense. The "finite means." This constitutes the generative procedure of what Otto Jespersen, in work generally neglected by professional linguists, called the "notion of structure" that guides our behavior in forming "free expressions" we have never heard and that may never have been produced, as in normal language use. As reconstructed within the "second cognitive revolution," the primary goal of linguistic research is to identify and precisely characterize these finite means represented in the mind/brain. More deeply, it is concerned to discover the principles of the language faculty that determine the form and character of acquired knowledge. To the extent that there further tasks can be accomplished, we have the basis for approaching Plato's problem in this domain, and we will at least understand the mechanisms that enter into the human actions that give rise to Descartes's problem. Similar conclusions hold in other cognitive domains.

Humboldt's and Plato's problems can be addressed within the framework of "the second cognitive revolution," and inquiry into these problems has achieved a measure of success. Descartes's problem, however, remains as mysterious as when it was formulated. It is a special case of problems about human action more generally, as the Cartesians emphasized. When its parts are fixed in a particular manner in a given environment, the behavior of a machine is determined or random. Under similar conditions, the Cartesians held, humans may be "incited and inclined" to act in a certain way, and may often or even always act in the way they are "incited and inclined" to act, but they are not "compelled" to do so; and within the limits of their capacities, they may choose to act in some different way. The problem remains beyond our comprehension, and may be beyond the grasp of our intelligence, as Descartes sometimes speculated.

Notice that this would not be a very surprising discovery, if true. If humans are part of the physical world, not angels, then they will have certain mental capacities, and not others, the capacities they have result from the richness of initial structure of the mind/brain, and this same richness of structure entails that other problems will lie beyond the scope of human intelligence, either in principle, or for reasons of complexity (understood not in absolute terms, but relative to a specific organism). We know this is true of other organisms, and it is plainly true of physical growth: a rich biological endowment enables the human embryo to grow to a complex and highly structured mature adult, and by the same token, prevents it from becoming a bird. It is a point of logic that scope and limits are related in the manner just indicated.

The same point of logic holds of human cognitive growth and development. A rich biological endowment enables the language faculty of the mind/brain to mature to a complex and highly articulated language, largely shared with others, and by the same token, prevents access to innumerable other systems that can be imagined and even constructed by other faculties of the mind. In empirical inquiry, humans have been capable of formulating and dealing with questions of certain types at great depth, making use of the innate capacities of the human mind. The very same initial endowment will block access to other possible domains. We evidently have capacities to construct theories that we regard as intelligible in certain problem situations, taking a problem situation to be determined by some state of understanding, some array of phenomena subjected to inquiry, and some questions formulated about them. Since these capacities have definite structure—otherwise they would achieve nothing in any problem situation—they will have scope and limits, and there is little reason *a priori* to expect these limits will include all matters we might hope to subject to inquiry. Looking at a rat from our point of view, we can readily understand why it is incapable of solving a maze that requires turning right at every prime number option, or even far simpler mazes; it simply lacks the relevant concepts, in principle. Similarly, knowing something about the human language faculty, we can readily design "languages" that will be unattainable by the language faculty, which will always make the wrong guesses. An intelligence constituted differently from ours might be able to draw similar conclusions about human science, observing our stumbling failures; and we might even be able to do so ourselves, without contradiction.

For those willing to adopt realist assumptions, the attainable sciences should be regarded as a kind of chance convergence of properties of our intelligence and the world as it is—a *chance* convergence. Because contrary to many speculations, there is little reason to suppose that evolution provided a *deus ex machina* to guarantee that human intelligence is capable of solving the problems posed in inquiry; ability to solve problems in mathematics or quantum theory, for example, was not a factor in human evolution. Karl Popper observed long ago that it "is clearly mistaken" to suppose that "our quest for knowledge must necessarily succeed," that it must be possible for us to explain the world. This conclusion should not appear controversial. And of course, we fortunately have many ways to come to understand aspects of the world apart from our science-forming faculties, whatever their character may be.

To clarify where we stand today, we might observe that these problems look fundamentally different to us than they did to Cartesians. The reason is that although Descartes's representational theory of mind was not seriously challenged, and has been partially resurrected in the second cognitive revolution, his theory of body quickly collapsed as Newton demonstrated the inadequacy of Cartesian "contact mechanics" for the study of simple interactions among bodies, the motion of the planets, for example. Following much the same logic as Descartes in his inquiry into the creative aspect of language use, Newton therefore proposed a principle that escaped the bounds of Cartesian mechanics: a principle of attraction among bodies that allowed "action at a distance." There is evidence that Newton found this "occult property" (as he sometimes called it) unsatisfying, as did many continental physicists. Nevertheless, in subsequent years it became part of the core of the sciences.

In retrospect we may say that Newton's discoveries effectively destroyed the classical notion of "body," though it took time for the insight to be absorbed; until early in this century it was often held that a true explanation, as contrasted with a predictive mathematical theory, must be "mechanical" in something not entirely unlike the Cartesian sense. But this conception is now understood to be quite wrong. Under a more reasonable and now current interpretation, we have no fixed and determinate notion of body, and have had no such notion since Newton undermined the common sense ideas articulated in Cartesian mechanics. Rather, the material world is whatever science determines it to be. If the material world contains fields of force, massless particles, curved space, strings vibrating in 10-dimensional space, or whatever the physicist will concoct tomorrow, then so be it: that is the nature of matter, of "body." Lacking any definite conception of "body," we cannot even pose the classical mind-body or man-machine problems. The Cartesians, in contrast, could pose the problems, since they had a fixed conception of "body" and "machine." Lacking such a conception, we cannot ask intelligibly whether something lies beyond its scope. If some phenomena are discovered, we can try to understand them in terms of an explanatory theory framed in terms of concepts appropriate to the task, and we can then turn to the problem of discovering how this theory is related to core areas of the natural sciences. But we never escape the bounds of "body" or "the material world," since there are no such bounds, there being no determinate concept of "body" or "matter" as distinct from what is discovered in empirical inquiry. When we speak of "mind," we are simply referring to properties of physical mechanisms, which remain to be discovered as the cognitive sciences are integrated into the "more fundamental" natural sciences in the manner of chemistry and genetics in earlier years.

I think our best current understanding is that the human mind includes specific faculties, the language faculty among them. These provide mechanisms for representation and computation, interpretation and understanding, thought and behavior, and social interaction. In these terms we can also, perhaps, hope to develop a version of the classical aesthetic doctrine that true creativity involves constructions undertaken within a framework of rule established by the mind, in terms of its own inner resources. There is every reason to suppose that the same is true of moral judgment, though little is understood about these matters as yet. This is, if true, a hopeful prospect. It suggests that moral discourse should be able to find a common ground in values and commitments that may be obscure to us and that we may have to discover by experience and experiment, much in the way that we learn about ourselves in other domains. History provides some reason to believe this might be true. The debate over slavery, for example, was not simply a matter of "I believe this" and "you believe that." Rather, arguments were offered on both sides, in terms of shared moral values. With the progress of civilization, it came to be understood it is a fundamental human right to be free from arbitrary constraint. The scope of this insight is one that is constantly enlarging, and that is far from having reached its limits.

Much of the fascination and significance of the study of language, I believe, lies in the fact that in this particular domain, we can formulate the issues that arise throughout the study of mind with considerable clarity, and can provide substantive and productive answers to Humboldt's and Plato's problem, and at least certain aspects of Descartes's. While these results do not bear directly on other domains, they do help clarify the issues and, I think, suggest a course that inquiry might fruitfully pursue.

It is important to recognize that until recently, the problems now on the agenda were not recognized as serious ones within the dominant tendencies in the fields concerned with language and other aspects of psychology. This was a significant regression from the insights of an earlier era in Western thought. The regression was associated with dramatic progress in certain narrower domains. Structural linguistics, in both its European and American varieties, achieved much progress and insight in specific areas, but was quite impoverished in general conception as compared with earlier thought; it is of some interest that these narrow and restricted concepts of language have had such enormous influence in other domains of intellectual culture. Insofar as Humboldt's concerns, or Jespersen's, were recognized at all, it was assumed that some process of "analogy" would suffice to account for the phenomena. Given the impoverished conception of language structure, the force of Plato's problem was also not recognized.

We see here a failure of imagination, not unlike that of the pre-Galilean era in the natural sciences. It is important to gain the capacity to be surprised by simple and apparently obvious things. If we are content to say that a stone falls to the ground because that is its natural place, and to dismiss the rate of fall as an uninteresting triviality, science cannot progress. Similarly, there will be no serious understanding of language if we dismiss the fact that normal linguistic behavior regularly involves expressions that have no close analogue in experience, or the fact that constructions of natural language and the elements of lexical-conceptual structure have the properties and meet the curious invariant conditions that we discover wherever we inquire seriously.

In retrospect, I think we can detect two major conceptual changes in the study of language during the period of the second cognitive revolution. The first was its inception, with the construction—or better, reconstruction—of a computational-representational theory of mind, and a serious concern for Humboldt's and Plato's problems. The second is more recent, more theory-internal, but quite radical in its implications. Throughout thousands of years of rich and productive study, a language has typically been regarded as some kind of rule system, where the rules are "learned" in the course of language acquisition. There is reason to believe this conception is inaccurate. Recent work suggests that human languages all adhere to a fixed framework of invariant principles that are quite different from the "rules of grammar" of traditional or modern generative theories. These principles are different, first, in that they are not specific to particular grammatical constructions; and second, in that they are genuinely invariant, biologically determined properties of the language faculty on a par with the structures that compel the embryo to develop arms, not wings, under appropriate conditions. Thus, there is no rule for formation of noun phrases, or questions, or relative clauses, or passive constructions. Rather, the traditional grammatical constructions appear to be epiphenomena, with no real existence. Their properties result through the interaction of principles of much broader scope and generality, formulated in much more abstract terms.

As for the variety of languages, this appears to derive from the fact that the invariant principles have associated with them certain limited parameters of variation. To illustrate with a simple case, the principles determining the formation of phrases require each phrase be a "projection" of a certain category of elements drawn from the lexicon, the "head" of the phrase: noun phrases are a projection of nouns as heads, verb phrases of verbal heads, and so on. The nature of the projections appears to be largely invariant across languages, but languages do differ in such matters as whether the head is initial, as in English, or final, as in Japanese; the result then is that Japanese and English appear to be mirror-images in certain structural respects. To acquire a language, a child must determine how these parameters are set. It is as if the child, prior to experience, already knows the general form of language down to rich detail and specificity, and approaches the task of language acquisition with a "questionnaire" consisting of a series of simple questions to be answered. For example, is this language head-first or head-last? Given the well-established fact that knowledge of language can be acquired on the basis of very fragmentary data, it must be that simple data suffice to answer each question on the questionnaire, as in the example just mentioned. Once the answers to these questions are determined, the language is fully fixed, in all of its richness and variety. It is then available for expression of thought, interpretation, communication, and other special uses. Much the same seems to be true of the structure of the lexicon, which draws from the resources of the invariant conceptual system in ways that allow extraordinarily rapid acquisition of lexical items, yielding a rich array of concepts of remarkable intricacy, with semantic connections among them that provide a determinate framework for thought, belief, understanding and the growth of knowledge. On the other hand, it appears that the phonological component of the language, which associates structural forms with representations of sound, does have the character of a rule system, much as assumed in linguistic research from Panini in ancient India to contemporary generative phonology.

If these conceptions are correct, then many questions of linguistics appear in a new guise. Language change will be the change in one or more parameter values; note that change in even a single parameter might have wide-ranging phenomenal consequences, as its effects filter through the fixed network of principles, something that has often been observed in the study of the history of language. Typology reduces to the study of variation in parameters, and it is no longer surprising to discover, as we commonly do, that unrelated languages are often remarkably similar in the most subtle ways; the reason would be that they happen to have set parameters in the same way.

Problems of language use also merit reconsideration from this point of view. I mentioned that Descartes's problem remains beyond the range of inquiry, possibly in principle, for human science. But aspects of the problem of language use can be formulated and addressed, in particular, by what is called "the parsing problem," a special case of the problem of perception and interpretation. The parsing problem abstracts from situation, context, shared assumptions and understanding, and asks how, under these idealized conditions, the mind assigns to an utterance a representation of its structure and lexical content. This is a problem that has been of much interest for technological applications such as man-machine communication.

Most approaches to the parsing problem are "rule-based"; it is assumed that parsing involves, in effect, checking the elements of a rule-system that is programmed into the parser. But if the principles-and-parameters theory is correct, then this approach may be misguided. Rather, one might suspect that the principles of the language faculty are an invariant element of the parser, hard-wired in effect, and that computation involves reference to the choice of parameters, or to parsing strategies and "shortcuts" that may be inherent to the parser itself or may be constructed by this faculty of mind.

These conclusions also are suggestive with regard to other aspects of perception and cognition, though, again, caution is in order. There is no reason to believe we will find the same mechanisms and principles in different components of the mind, and our limited current understanding suggests that indeed we do not.

One general question that arises in this connection has to do with the relation of form and function. It is commonly assumed that the design of language must be such as to render it readily usable. This assumption is less obvious than it may seem. Evolutionary biology, for example, gives us little reason to believe that the system of language that developed in the human mind/brain is somehow "designed for use." We must not succumb to what the evolutionary biologists Stephen Gould and Richard Lewontin call the "Panglossian fallacy," the assumption that each trait of an organism is selected so as to yield near-optimal adaptation to the environment. Many factors beyond survival value enter into biological variation and evolution."In some cases at least, the forms of living things, and of the parts of living things, can be explained by physical consideration," D'Arcy Thompson observed in his classic work, and such ideas might reach quite far towards an explanation of the properties of complex systems of nature. What we learn from evolutionary biology is that parts of language should be expected to be usable; those parts will be used, others not. And indeed, that is exactly what we find.

Considerations of this sort bear on the possibility of providing so-called "functional explanations" for properties of language. If we construct two systems at random, one a generative procedure that strongly generates structural descriptions, the other a parsing system, we are likely to find some respects in which the two are well adapted to one another, others in which they are not. If the generative procedure is incorporated in the parser, which has access to it for performance, then the parser will be able to make use of the information provided by the generative procedure to the extent that the two systems are mutually adapted. It would be a mistake to conclude that the generative procedure was designed for use by the parser just on the basis of the fact that there is a domain of adaptation. One would have to show this domain goes beyond what might be expected on other grounds, not an easy task. Such questions arise whenever functional accounts are offered.

It is well-known that design features of language render large classes of expressions unusable, though their form and meaning is fixed and determinate. One basic feature of natural language is that it permits embedding of phrases within other phrases, where the containing phrase may exhibit a dependency across the embedding. Thus in the sentence "the man who you met is tall," the phrase "who you met" is embedded within the dependent phrases "the man" and "is tall," each of which must be singular in number. It is readily demonstrated that this core design feature of natural language yields simple constructions that are unusable, though they can often be deciphered with sufficient attention, time, effort and external memory. Similarly, it is a familiar fact that language design yields sentences assigned the wrong interpretation, or no intelligible interpretation, because parsing proceeds along the wrong path. These normal properties of language are not at all surprising. Nor do they impede communication particularly; the speaker keeps to those aspects of language that are usable, and these are the only ones that the interpreter can process.

Further analysis of language structure reveals more subtle illustration of the incompatibility of form and function. Though we now enter into areas of ongoing controversy, there is, I believe, good reason to suppose that a language is a self-contained system of representation and computation that "interfaces" with other systems at three points, yielding three basic representations for each linguistic expression: (1) the phonetic representation is the interface with the external articulatory and perceptual systems; (2) the "logical form" representation is the interface with the conceptual systems that enter into semantic interpretation and conditions of use; (3) a representation sometimes called "D-structure" ("deep structure") is a direct reflection of properties of lexical structure. For a given utterance, there will be, then, three representations; the form of each is determined by properties of the interface. These three representations must be related. The design of language requires that the relation is only indirect: each of the basic representations is related directly to a fourth level of representation, called "S-structure," which is a "derivative" in the sense that it satisfies the conditions imposed "externally" on each of the basic levels and the very restrictive conditions that the language faculty permits for expressing the interlevel relations. S-structure may also have to meet additional conditions of its own. It is, in effect, the solution to a certain "system of equations." Furthermore, there is reason to believe the relations among these levels are "directional mappings." Certain operations form S-structure from D-structure and map S-structure to phonetic and logical forms, independently.

Consider then the problem of parsing an utterance. The parser is presented with a phonetic form. If it follows the structure of language design, it must "guess" the D-structure, then forming the S-structure, checking its compatibility with phonetic form, and deriving the logical form representation from the S-structure. This is a very difficult computational problem, far from optimal. But it appears to be the nature of language design.

Natural languages also differ in design features from artificial languages designed for ease of use, as in quantificational structure, for example, again suggesting form-function incompatibility.

Another such feature is the prevalence of displaced phrases, appearing in the utterance in a position different from the one in which they are interpreted. For example, in the sentence "which men did they expect to hurt themselves," the question phrase *which men* is understood to be the subject of *hurt* and the antecedent of the reflexive form *themselves*. Furthermore, there is strong reason to believe that an "empty category," syntactically real but lacking a phonetic realization, appears in the position in which the displaced phrase is interpreted; in the sentence "which men did they expect to hurt themselves," an empty category linked to "which men" appears as the subject of *hurt*, thus preventing the hearer from interpreting the reflexive *themselves* as referring to the physically closest noun phrase, *they*, as it would be in "they expect to hurt themselves," with *which men* deleted. The empty category serves as a variable bound by the quantificational expression "which men."

Other phrases can also be displaced. The noun phrase "that boy" is understood to be the object of "injured" in the sentence "that boy seems to have been injured," though it is the subject of "seem" and is unrelated formally to "injure"; there is good evidence that it is related formally to an empty category that is indeed the object of "injure," behaving in the manner of a bound variable. Similarly quantificational expressions can be displaced, again leaving bound empty categories in the position of normal interpretation. Thus, in the sentence "one translator each seems to have been assigned to the visiting diplomats," the quantifier *each* is associated with *the visiting diplomats*. But in the very similar sentence "one translator each wanted to be assigned to the visiting diplomats," no such interpretation is possible. The same is true of the numerical quantifiers of Japanese, a matter studied in forthcoming work by Shigeru Miyagawa.

These "displacements" are only of a limited variety, in fact, a variety satisfying certain narrow conditions on the relation of the S-structure to the underlying D-structure from which it derives. Displaced phrases and empty categories also impose computational problems, but these are widespread phenomena in natural language.

There is, furthermore, some recent work that suggests that language design satisfies certain overarching conditions that have a kind of "least effort" flavor to them. Specifically, there appear to be conditions requiring that the computations relating the various representations of an utterance be “minimal” in a certain formal sense, and that the representations themselves be “minimal,” lacking superfluous symbols. Once made precise, these conditions have wide-ranging empirical consequences. Superficially, such “minimality” conditions seem to provide form-function compatibility, but that is a mistake. The conditions are “global,” affecting the entire computation of the set of representations that constitute the structure of a sentence that we produce and understand. It follows that this set cannot be constructed in terms of local relations among representations and their parts. It is well-known that such “global” properties yield computational intractability for production and parsing. Correspondingly, such principles render language highly “unusable”—though particular parts may be usable, and certain computational “tricks” might be available to the parser to overcome the deep incompatibility of form and function.

Notice that these “minimality conditions,” with their “least effort” flavor have a kind of generality that is lacking in the specific principles of language structure. Nevertheless, the actual formulation of the conditions appears to be highly specific to the language faculty. The generality is, furthermore, more a matter of elegance than utility; it is the kind of property that one seeks in core areas of the natural sciences, for example, in searching for conservation principles, symmetry, and the like.

The general conclusion that seems to come to the fore is that language is designed as a system that is “beautiful,” but in general unusable. It is designed for elegance, not for use, though with features that enable to it be used sufficiently for the purposes of normal life. These are properties of language that have been observed in other respects as well. Thus, it has often proven to be a productive guiding intuition in research that if some property of language is “overdetermined” by proposed principles, then probably the principles are wrong, and some way should be found to reconstruct them so as to avoid this redundancy. Insofar as this is true, the system is elegant, but badly designed for use. Typically, biological systems are not like this at all. They are highly redundant, for reasons that have a plausible functional account. Redundancy offers protection against damage, and might facilitate overcoming problems that are computational in nature. In these respects, then, language seems rather different from other biological systems. We must, of course, take into account the possibility that all such conclusions might be a kind of artifact, a result of our methods of investigation and theory construction, not properties of the real object of the real world that we are investigating. There is fairly good evidence, however, that they are basically accurate.

We might suppose these properties, if indeed they are real, are related to other features of language that are unusual among biological organisms. Language is, at its core, a system that is both digital and infinite. Such systems are rare, though not unknown; another obvious case is the number system, also a unique human possession it appears, and quite probably, a derivative from the language faculty. As to how this faculty with its unusual features and its form-function incompatibility arose in the human species, we can only speculate. As I mentioned, speculations about natural selection are no more plausible than many others; perhaps these are simply emergent physical properties of a brain that reaches a certain level of complexity under the specific conditions of human evolution. Here, we move to questions that are, at the moment, intractable.

Reviewing where we now stand, with regard to the language faculty, a reasonable position seems to me to be something like this. For unknown reasons, the human mind/brain developed the faculty of language, a computational-representational system based on digital computation with many specific design properties. The system appears to be surprisingly elegant, possibly observing conditions of nonredundancy, global "least effort" conditions, and so on. It also seems to have many properties, including some deeply rooted in its basic design, that make it dysfunctional, unusable, although adequate for actual use over a sufficient range because of other special properties, a fact that might be relevant to its persistence and development in the human species and to human biological success. The properties of the language faculty seem to be unique to humans in interesting respects and distinct from other subsystems of the mind/brain. The mind, then, is not a system of general intelligence as has been assumed over a very broad spectrum of traditional and contemporary thought. Rather, the mind has distinct subsystems such as the language faculty; this is a cognitive system, a system of knowledge, not an input or output system.

This faculty, furthermore, is internally highly modularized, with separate subsystems of principles governing sound, meaning, structure and interpretation of linguistic expressions. These can be used, to a sufficient degree, in thought and its expression, and in specific language functions such as communication; language is not intrinsically a system of communication, nor is it the only system used for communication. The language faculty is based on fixed principles with limited options of parametric variation as the system is “tuned” to a specific environment, yielding a finite number of languages apart from lexicon, also sharply constrained. It might turn out that such variation is limited to lexical properties, in which case there is only one language, apart from properties of the lexicon.

A rich system of knowledge develops over a broad infinite range. This system provides examples of propositional knowledge—knowledge that so-and-so—as well as knowledge-how, knowledge-why, and so on. Contrary to what is often supposed, knowledge-that knowledge-how, in particular, cannot be understood in terms of ability; rather, ability to use a system of knowledge must be clearly distinguished from possession of this knowledge. Furthermore, such knowledge does not satisfy the standard conditions of epistemology; it is not obtained by general principles, is not based on good reasons or justified, or anything of the sort. Rather, what we come to know and understand is determined by our biological nature in quite substantial ways, which we can sketch out with some degree of specificity in a range of interesting cases. We have knowledge of some aspect of the world only when the systems that develop in the mind/brain, and our modes of interpreting data as experience, conform to a sufficient degree with elements of the world around us.

I have spoken only of the study of language, one of a few domains of cognitive psychology with rather far-reaching results. But I think it would hardly be surprising if the truth of the matter were qualitatively similar in other domains, where far less is known. As far as I can see, only ancient prejudice makes this prospect appear to many to be unlikely.