

Language and the Brain

The right way to address the announced topic would be to review the fundamental principles of language and the brain and to show how they can be unified, perhaps on the model of chemistry and physics sixty-five years ago, or the integration of parts of biology within the complex a few years later. But that course I am not going to try to attempt. One of the few things I can say about this topic with any confidence is that I do not begin to know enough to approach it in the right way. With less confidence I suspect it may be fair to say that current understanding falls well short of laying the basis for the unification of the sciences of the brain and higher mental faculties, language among them, and that many surprises may lie along the way to what seems a distant goal—which would itself come as no surprise if the classical examples I mentioned are indeed a realistic model.

This somewhat skeptical assessment of current prospects differs from two prevalent but opposing views. The first holds that the skepticism is unwarranted, or more accurate, profoundly in error, because the question of unification does not even arise. It does not arise for psychology as the study of mind, because the topic does not fall within biology, a position taken to define the “computer model of mind”;¹ nor for language, because language is an extra-human object, the standard view within major currents of philosophy of mind and language, and also put forth recently by prominent figures in neuroscience and ethology. At least that is what the words seem to imply; the intentions may be different. I will return to some prominent current examples.

This chapter first appeared in *On Nature and Language* (Cambridge: Cambridge University Press, 2002), 61–91.

A contrasting view holds that the problem of unification does arise, but that the skepticism is unwarranted. Unification of the brain and cognitive sciences is an imminent prospect, overcoming Cartesian dualism. This optimistic assessment is expressed forthrightly by evolutionary biologist E. O. Wilson in a recent publication of the American Academy of Arts and Sciences devoted to the brain, summarizing the state of the art, and seems to be shared rather broadly: “Researchers now speak confidently of a coming solution to the brain-mind problem.”² Similar confidence has been expressed for half a century, including announcements by eminent figures that the brain-mind problem has been solved.

We can, then, identify several points of view with regard to the general problem of unification:

1. There is no issue: language and higher mental faculties generally are not part of biology.
2. They belong to biology in principle, and any constructive approach to the study of human thought and its expression, or of human action and interaction, relies on this assumption, at least tacitly.

Category (2), in turn, has two variants: (A) unification is close at hand; (B) we do not currently see how these parts of biology relate to one another, and suspect that fundamental insights may be missing altogether.

The last point of view, (2B), seems to me the most plausible. I will try to indicate why, and to sketch some of the terrain that should be covered in a careful and comprehensive overview of these topics.

As a framework for the discussion, I would like to select three theses that seem to me generally reasonable, and have for a long time. I will quote current formulations by leading scientists, however, not my own versions from past years.

The first thesis is articulated by neuroscientist Vernon Mountcastle, introducing the American Academy study I mentioned. A guiding theme of the contributions, and the field generally, he observes, is that “Things mental, indeed minds, are emergent properties of brains,” though “these emergences are not regarded as irreducible but are produced by principles that control the interactions between lower level events—principles we do not yet understand.”

The second thesis is methodological. It is presented clearly by ethologist Mark Hauser in his comprehensive study *Evolution of Communication*.³ Following Tinbergen, he argues, we should adopt four perspectives in studying “communication in the animal kingdom, including human language.” To understand some trait, we should:

1. Seek the mechanisms that implement it, psychological and physiological; the mechanistic perspective.
2. Sort out genetic and environmental factors, which can also be approached at psychological or physiological levels; the ontogenetic perspective.
3. Find the “fitness consequences” of the trait, its effects on survival and reproduction; the functional perspective.
4. Unravel “the evolutionary history of the species so that the structure of the trait can be evaluated in light of ancestral features”; the phylogenetic perspective.

The third thesis is presented by cognitive neuroscientist C. R. Gallistel:⁴ the “modular view of learning,” which he takes to be “the norm these days in neuroscience.” According to this view, the brain incorporates “specialized organs,” computationally specialized to solve particular kinds of problems, as they do with great facility, apart from “extremely hostile environments.” The growth and development of these specialized organs, sometimes called “learning,” is the result of internally directed processes and environmental effects that trigger and shape development. The language organ is one such component of the human brain.

In conventional terminology, adapted from earlier usage, the language organ is the faculty of language (FL); the theory of the initial state of FL, an expression of the genes, is universal grammar (UG); theories of states attained are particular grammars; the states themselves are internal languages, “languages” for short. The initial state is, of course, not manifested at birth, as in the case of other organs, say the visual system.

Let us now look more closely at the three theses—reasonable I think, but with qualifications—beginning with the first: “Things mental, indeed minds, are emergent properties of brains.”

The thesis is widely accepted, and is often considered a distinctive and exciting contribution of the current era, if still highly controversial. In the past few years it has been put forth as an “astonishing hypothesis,” “the bold assertion that mental phenomena are entirely natural and caused by the neurophysiological activities of the brain” and “that capacities of the human mind are in fact capacities of the human brain”; or as a “radical new idea” in the philosophy of mind that may at last put an end to Cartesian dualism, though some continue to believe that the chasm between body and mind cannot be bridged.

The picture is misleading, and it is useful to understand why. The thesis is not new, and it should not be controversial, for reasons understood centuries ago. The thesis was articulated clearly in the eighteenth century, and for compelling reasons—though controversially then, because of affront to religious

doctrines. By 1750, David Hume casually described thought as a “little agitation of the brain.”⁵ A few years later the thesis was elaborated by the eminent chemist Joseph Priestley: “the powers of sensation or perception and thought” are properties of “a certain organized system of matter”; properties “termed mental” are “the result [of the] organical structure” of the brain and “the human nervous system” generally. Equivalently: “Things mental, indeed minds, are emergent properties of brains” (Mountcastle). Priestley of course could not say how this emergence takes place, nor can we do much better after two hundred years.

I think the brain and cognitive sciences can learn some useful lessons from the rise of the emergence thesis in early modern science, and the ways the natural sciences have developed since right up to the mid twentieth century, with the unification of physics–chemistry–biology. Current controversies about mind and brain are strikingly similar to debates about atoms, molecules, chemical structures and reactions, and related matters, which were very much alive well into the twentieth century. Similar, and in ways that I think are instructive.

The reasons for the eighteenth-century emergence thesis, recently revived, were indeed compelling. The modern scientific revolution, from Galileo, was based on the thesis that the world is a great machine, which could in principle be constructed by a master artisan, a complex version of the clocks and other intricate automata that fascinated the seventeenth and eighteenth centuries, much as computers have provided a stimulus to thought and imagination in recent years; the change of artifacts has limited consequences for the basic issues, as Alan Turing demonstrated sixty years ago. The thesis—called “the mechanical philosophy”—has two aspects: empirical and methodological. The factual thesis has to do with the nature of the world: it is a machine constructed of interacting parts. The methodological thesis has to do with intelligibility: true understanding requires a mechanical model, a device that an artisan could construct.

This Galilean model of intelligibility has a corollary: when mechanism fails, understanding fails. For this reason, when Galileo came to be disheartened by apparent inadequacies of mechanical explanation, he finally concluded that humans will never completely understand even “a single effect in nature.” Descartes, in contrast, was much more optimistic. He thought he could demonstrate that most of the phenomena of nature could be explained in mechanical terms: the inorganic and organic world apart from humans, but also human physiology, sensation, perception, and action to a large extent. The limits of mechanical explanation were reached when these human functions are mediated by thought, a unique human possession based on a principle that escapes mechanical explanation: a “creative” principle that underlies

acts of will and choice, which are “the noblest thing we can have” and all that “truly belongs” to us (in Cartesian terms). Humans are only “incited and inclined” to act in certain ways, not “compelled” (or random), and in this respect are unlike machines—that is, the rest of the world. The most striking example for the Cartesians was the normal use of language: humans can express their thoughts in novel and limitless ways that are constrained by bodily state but not determined by it, appropriate to situations but not caused by them, and that evoke in others thoughts that they could have expressed in similar ways—what we may call “the creative aspect of language use.”

It is worth bearing in mind that these conclusions are correct, as far as we know.

In these terms, Cartesian scientists developed experimental procedures to determine whether some other creature has a mind like ours—elaborate versions of what has been revived as the Turing test in the past half century, though without some crucial fallacies that have attended this revival, disregarding Turing’s explicit warnings, an interesting topic that I will put aside.⁶ In the same terms, Descartes could formulate a relatively clear mind-body problem: having established two principles of nature, the mechanical and mental principles, we can ask how they interact, a major problem for seventeenth-century science. But the problem did not survive very long. As is well-known, the entire picture collapsed when Newton established, to his great dismay, that not only does mind escape the reach of the mechanical philosophy, but so does everything else in nature, even the simplest terrestrial and planetary motion. As pointed out by Alexander Koyré, one of the founders of the modern history of science, Newton showed that “a purely materialistic or mechanistic physics is impossible.”⁷ Accordingly, the natural world fails to meet the standard of intelligibility that animated the modern scientific revolution. We must accept the “admission into the body of science of incomprehensible and inexplicable ‘facts’ imposed upon us by empiricism,” as Koyré puts the matter.

Newton regarded his refutation of mechanism as an “absurdity,” but could find no way around it despite much effort. Nor could the greatest scientists of his day, or since. Later discoveries introduced still greater “absurdities.” Nothing has lessened the force of David Hume’s judgment that by refuting the self-evident mechanical philosophy, Newton “restored Nature’s ultimate secrets to that obscurity in which they ever did and ever will remain.”

A century later, in his classic history of materialism, Friedrich Lange pointed out that Newton effectively destroyed the materialist doctrine as well as the standards of intelligibility and the expectations that were based on it: scientists have since “accustomed ourselves to the abstract notion of forces, or rather to a notion hovering in a mystic obscurity between abstraction and

concrete comprehension,” a “turning-point” in the history of materialism that removes the surviving remnants of the doctrine far from those of the “genuine Materialists” of the seventeenth century, and deprives them of much significance.

Both the methodological and the empirical theses collapsed, never to be reconstituted.

On the methodological side, standards of intelligibility were considerably weakened. The standard that inspired the modern scientific revolution was abandoned: the goal is intelligibility of theories, not of the world—a considerable difference, which may well bring into operation different faculties of mind, a topic some day for cognitive science, perhaps. As the preeminent Newton scholar I. Bernard Cohen put the matter, these changes “set forth a new view of science” in which the goal is “not to seek ultimate explanations,” rooted in principles that appear to us self-evident, but to find the best theoretical account we can of the phenomena of experience and experiment. In general, conformity to commonsense understanding is not a criterion for rational inquiry.

On the factual side, there is no longer any concept of body, or matter, or “the physical.” There is just the world, with its various aspects: mechanical, electromagnetic, chemical, optical, organic, mental—categories that are not defined or delimited in an a priori way, but are at most conveniences: no one asks whether life falls within chemistry or biology, except for temporary convenience. In each of the shifting domains of constructive inquiry, one can try to develop intelligible explanatory theories, and to unify them, but no more than that.

The new limits of inquiry were understood by working scientists. The eighteenth-century chemist Joseph Black observed that “chemical affinity must be accepted as a first principle, which we cannot explain any more than Newton could explain gravitation, and let us defer accounting for the laws of affinity until we have established such a body of doctrine as Newton has established concerning the laws of gravitation.” That is pretty much what happened. Chemistry proceeded to establish a rich body of doctrine; “its triumphs [were] built on no reductionist foundation but rather achieved in isolation from the newly emerging science of physics,” a leading historian of chemistry observes.⁸ In fact, no reductionist foundation was discovered. What was finally achieved by Linus Pauling sixty-five years ago was unification, not reduction. Physics had to undergo fundamental changes in order to be unified with basic chemistry, departing even more radically from commonsense notions of “the physical”: physics had to “free itself” from “intuitive pictures” and give up the hope of “visualizing the world,” as Heisenberg put

it,⁹ yet another long leap away from intelligibility in the sense of the scientific revolution of the seventeenth century.

The early modern scientific revolution also brought about what we should properly call “the first cognitive revolution”—maybe the only phase of the cognitive sciences to deserve the name “revolution.” Cartesian mechanism laid the groundwork for what became neurophysiology. Seventeenth- and eighteenth-century thinkers also developed rich and illuminating ideas about perception, language, and thought that have been rediscovered since, sometimes only in part. Lacking any conception of body, psychology could then—and can today—only follow the path of chemistry. Apart from its theological framework, there has really been no alternative to John Locke’s cautious speculation, later known as “Locke’s suggestion”: God might have chosen to “superadd to matter a faculty of thinking” just as he “annexed effects to motion which we can in no way conceive motion able to produce”—notably the property of attraction at a distance, a revival of occult properties, many leading scientists argued (with Newton’s partial agreement).

In this context the emergence thesis was virtually inescapable, in various forms:

For the eighteenth century: “the powers of sensation or perception and thought” are properties of “a certain organized system of matter”; properties “termed mental” are “the result [of the] organical structure” of the brain and “the human nervous system” generally.

A century later, Darwin asked rhetorically why “thought, being a secretion of the brain,” should be considered “more wonderful than gravity, a property of matter.”¹⁰

Today, the study of the brain is based on the thesis that “Things mental, indeed minds, are emergent properties of brains.”

Throughout, the thesis is essentially the same, and should not be contentious: it is hard to imagine an alternative in the post-Newtonian world.

The working scientist can do no better than to try to construct “bodies of doctrine” for various aspects of the world, and seek to unify them, recognizing that the world is not intelligible to us in anything like the way the pioneers of modern science hoped, and that the goal is unification, not necessarily reduction. As the history of the sciences clearly reveals, one can never guess what surprises lie ahead.

It is important to recognize that Cartesian dualism was a reasonable scientific thesis, but one that disappeared three centuries ago. There has been no mind-body problem to debate since. The thesis did not disappear because of inadequacies of the Cartesian concept of mind, but because the concept of

body collapsed with Newton's demolition of the mechanical philosophy. It is common today to ridicule "Descartes's error" in postulating mind, his "ghost in the machine." But that mistakes what happened: Newton exorcized the machine; the ghost remained intact. Two contemporary physicists, Paul Davies and John Gribbin, close their recent book *The Matter Myth* by making that point once again, though they misattribute the elimination of the machine: to the new quantum physics. True, that adds another blow, but the "matter myth" had been demolished 250 years earlier, a fact that was understood by working scientists at the time, and has become part of the standard history of the sciences since. These are issues that merit some thought, I believe.

For the rejuvenated cognitive science of the twentieth century, it is also useful, I think, to pay close attention to what followed the unification of a virtually unchanged chemistry with a radically revised physics in the 1930s, and what preceded the unification. The most dramatic event that followed was the unification of biology and chemistry. This was a case of genuine reduction, but to a newly created physical chemistry; some of the same people were involved, notably Pauling. This genuine reduction has sometimes led to the confident expectation that mental aspects of the world will be reduced to something like the contemporary brain sciences. Maybe so, maybe not. In any event, the history of science provides little reason for confident expectations. True reduction is not so common in the history of science, and need not be assumed automatically to be a model for what will happen in the future.

Still more instructive is what was taking place just before the unification of chemistry and physics. Prior to unification, it was commonly argued by leading scientists that chemistry is just a calculating device, a way to organize results about chemical reactions, sometimes to predict them. In the early years of the last century, molecules were regarded the same way. Poincaré ridiculed the belief that the molecular theory of gases is more than a mode of calculation; people fall into that error because they are familiar with the game of billiards, he said. Chemistry is not about anything real, it was argued: the reason is that no one knew how to reduce it to physics. In 1929, Bertrand Russell—who knew the sciences well—pointed out that chemical laws "cannot at present be reduced to physical laws";¹¹ not false, but misleading in an important way. It turned out that the phrase "at present" was out of place. Reduction was impossible, as was soon discovered, until the conception of physical nature and law was (radically) revised.

It should now be clear that the debates about the reality of chemistry were based on fundamental misunderstanding. Chemistry was "real" and "about the world" in the only sense of these concepts that we have: it was part of the best conception of how the world works that human intelligence had been able to contrive. It is impossible to do better than that.

The debates about chemistry a few years ago are in many ways echoed in philosophy of mind and cognitive science today—and theoretical chemistry, of course, is hard science, merging indistinguishably with core physics: it is not at the periphery of scientific understanding, like the brain and cognitive sciences, which are trying to study systems that are vastly more complex, and poorly understood. These very recent debates about chemistry, and their unexpected outcome, should be instructive for the brain and cognitive sciences. They suggest that it is a mistake to think of computer models of the mind that are divorced from biology—that is, in principle unaffected by anything that might be discovered in the biological sciences—or Platonistic or other nonbiological conceptions of language, also insulated from important evidence, to their detriment, or to hold that the relation of the mental to the physical is not reducibility but the weaker notion of *supervenience*: any change in mental events or states entails a “physical change,” though not conversely, and there is nothing more specific to say. The reunification debates over chemistry could be rephrased in these terms: those denying the reality of chemistry could have held that chemical properties supervene on physical properties, but are not reducible to them. That would have been an error: the right physical properties had not yet been discovered. Once they were, talk of supervenience became superfluous and we move towards unification. The same stance seems to me reasonable in the study of mental aspects of the world.

In general, it seems sensible to follow the good advice of post-Newtonian scientists, and Newton himself for that matter, and seek to construct “bodies of doctrine” in whatever terms we can, unshackled by commonsense intuitions about how the world must be—we know that it is not that way—and untroubled by the fact that we may have to “defer accounting for the principles” in terms of general scientific understanding, which may turn out to be inadequate to the task of unification, as has regularly been the case for three hundred years. A good deal of discussion of these topics seems to me misguided, perhaps seriously so, for reasons such as these.

There are other similarities worth remembering between reunification chemistry and current cognitive science. The “triumphs of chemistry” provided valuable guidelines for the eventual reconstruction of physics: they provided conditions that core physics would have to meet. In a similar way, discoveries about bee communication provide conditions that have to be met by some future account in terms of cells. In both cases, it is a two-way street: the discoveries of physics constrain possible chemical models, as those of basic biology should constrain models of insect behavior.

There are familiar analogues in the brain and cognitive sciences: the issue of computational, algorithmic and implementation theories emphasized by David Marr, for example. Or Eric Kandel’s work on learning in marine snails,

seeking “to translate into neuronal terms ideas that have been proposed at an abstract level by experimental psychologists,” and thus to show how cognitive psychology and neurobiology “may begin to converge to yield a new perspective in the study of learning.”¹² Very reasonable, though the actual course of the sciences should alert us to the possibility that the convergence may not take place because something is missing—where, we cannot know until we find out.

I have been talking so far about the first of the three theses I mentioned at the outset: the guiding principle that “Things mental, indeed minds, are emergent properties of brains.” That seems correct, but close to truism, for reasons understood by Darwin and by eminent scientists a century earlier, and that followed from Newton’s discovery of “absurdities” that were nonetheless true.

Let us turn to the second: the methodological thesis, quoted from Mark Hauser’s *Evolution of Communication*: to account for some trait we must adopt the ethological approach of Tinbergen, with its four basic perspectives: (1) mechanisms, (2) ontogenesis, (3) fitness consequences, (4) evolutionary history.

For Hauser, as for others, the “Holy Grail” is human language: the goal is to show how it can be understood if we investigate it from these four perspectives, and only that way. The same should be true of vastly simpler systems: the “dance language” of the honeybee, to select the sole example in the animal world that, according to standard (though not uncontroversial) accounts, seems to have at least some superficial similarity to human language: infinite scope, and the property of “displaced reference”—the ability to communicate information about something not in the sensory field. Bees have brains the size of a grass seed, with less than a million neurons; there are related species that differ in mode of communication; there are no restrictions on invasive experiment. But basic questions remain unanswered: questions about physiology and evolution, in particular.

In his review of this topic, Hauser does not discuss mechanisms, and the few suggestions that have been made seem rather exotic; for example, mathematician/biologist Barbara Shipman’s theory that the bee’s performance is based on an ability to map a certain six-dimensional topological space into three dimensions, perhaps by means of some kind of “quark detector.”¹³ On evolution, Hauser has only a few sentences, which essentially formulate the problem. The same is true of other cases he reviews. For example, songbirds, which are “the success story in developmental research,” although there is no “convincing scenario” about selection—or even an unconvincing one, it seems.

It should hardly surprise us, then, that questions about physiological mech-

anisms and phylogenesis remain so mysterious in the incomparably more difficult case of human language.

A closer look at Hauser's study gives some indication of the remoteness of the goal that he and others set—a worthy goal, but we should be realistic about where we stand in relation to it. First, the title of the book is misleading: it is not about the evolution of communication, a topic that receives only passing mention. Rather, it is a comparative study of communication in many species. That is made explicit in the comments in Derek Bickerton's review in *Nature* that are quoted on the jacket cover; and in the final chapter, which speculates about "future directions." The chapter is entitled "Comparative communication," realistically; there is little speculation about evolution, a quite different matter. Rather generally, what Hauser and others describe as the record of natural selection turns out to be an account of the beautiful fit of an organism to its ecological niche. The facts are often fascinating and suggestive, but they do not constitute evolutionary history: rather, they formulate the problem to be solved by the student of evolution.

Second, Hauser points out that this comprehensive study of comparative communication is "irrelevant to the formal study of language" (an overstatement, I think). That is no small point: what he calls the "formal study of language" includes the psychological aspects of the first two perspectives of the ethological approach: (1) the mechanisms of language, and (2) their ontogenesis. And what is irrelevant to psychological aspects is irrelevant to physiological aspects as well, since anything that has bearing on physiological aspects imposes conditions on psychological aspects. Accordingly, the first two perspectives of the recommended approach of Tinbergen are effectively abandoned, for human language. For similar reasons, the comparative study may be "irrelevant," in the same sense, to contemporary inquiry into bee communication, largely a richly detailed variety of "descriptive linguistics." That seems a plausible conclusion: a great deal has been learned about particular species at a descriptive level—insects, birds, monkeys, and others. But little emerges of any generality.

The "irrelevance" to human language is, however, far deeper. The reason is that—as Hauser also observes—language is not properly regarded as a system of communication. It is a system for expressing thought, something quite different. It can of course be used for communication, as can anything people do—manner of walking or style of clothes or hair, for example. But in any useful sense of the term, communication is not the function of language, and may even be of no unique significance for understanding the functions and nature of language. Hauser quotes Somerset Maugham's quip that "if nobody spoke unless he had something to say, . . . the human race would very soon lose the use of speech." His point seems accurate enough, even apart from the

fact that language use is largely to oneself: “inner speech” for adults, monologue for children. Furthermore, whatever merit there may be to guesses about selectional processes that might, or might not, have shaped human language, they do not crucially depend on the belief that the system is an outgrowth of some mode of communication. One can devise equally meritorious (that is, equally pointless) tales of the advantage conferred by a series of small mutations that facilitated planning and clarification of thought; perhaps even less fanciful, since it is unnecessary to suppose that the mutations took place in parallel in the group—not that I am proposing this or any other story. There is a rich record of the unhappy fate of highly plausible stories about what might have happened, once something was learned about what did happen—and in cases where far more is understood.

In the same connection, it is noteworthy that human language does not even appear in Hauser’s “taxonomy of communicative information” (mating, survival, identity of caller). Language can surely be used for alarm calls, identification of speaker, and so on, but to study the functioning of language in these terms would be hopelessly misleading.

A related difficulty is that Hauser restricts the functional perspective to “adaptive solutions.” That sharply limits the study of evolution, a point that Darwin forcefully emphasized and is now much better understood. In fact, Hauser cites case after case of traits that have no adaptive function, so he argues—appearing only in contrived situations with no counterpart in nature.

These matters are barely discussed; what I have cited are scattered remarks, a sentence here and there. But they indicate the immensity of the gaps that we must contemplate if we take the ethological perspective seriously—as of course we should, so I believe, and have been arguing for forty years.¹⁴ Hauser’s speculations about some future inquiry into the evolution of human language highlight the mystery. He refers to the two familiar basic problems: it is necessary to account for (1) the massive explosion of the lexicon, and (2) the recursive system for generating an infinite variety of meaningful utterances. For the latter, no speculation is offered. As for (1), Hauser reports that there is nothing analogous in the animal kingdom, including his own specialty (nonhuman primates). He observes that a precondition for the explosion of the lexicon is an innate human capacity to imitate, which he finds to be fundamentally different from anything in the animal world, perhaps unique. He was able to find only one possible exception: apes subjected to training. His conclusion is that “certain features of the human environment are required for engaging the capacity to imitate in apes,” which, if true, would seem to imply that the capacity is not the result of the adaptive selection to which he and others insist we must restrict ourselves in studying evolution. As for the origins of

the human capacity to imitate, he points out that we know nothing and may never be able to find out when—or for that matter how—it appeared in hominid evolution.

Furthermore, like many others, Hauser seriously underestimates the ways in which the human use of words to refer differs in its essential structural and functional properties from the rare examples of “referential signals” in other species, including some monkeys (possibly some apes, though the evidence, he says, is uncertain), a matter that goes well beyond the issues of displaced and situation-free reference. And he also seriously overstates what has been shown. Thus, citing some of Darwin’s cautious speculations, he writes that “we thus learn two important lessons” about “human language evolution”: that “the structure and function of human language can be accounted for by natural selection,” and that “the most impressive link between human and nonhuman-animal forms of communication lies in the ability to express emotional state.” Similarly, Steven Pinker “shows how a Darwinian account of language evolution is the only possible account, . . . because natural selection is the only mechanism that can account for the complex design features of a trait such as language” (my emphasis). It would be remarkable if something had been “shown” about the evolution of human language, let alone the vastly more ambitious claim cited; or if we could “learn” anything significant from speculations about the topic. Surely nothing so amazing has taken place. Cautious speculation and confident pronouncement do not show anything, and the most that we learn is that there might be a useful path to follow. Perhaps.

That aside, the conclusions that have supposedly been demonstrated make little sense, apart from a charitable reading; uncontroversially, natural selection operates within a space of options determined by natural law (and historical/ecological contingencies), and it would be the sheerest dogmatism to issue a priori proclamations on the role of these factors in what comes to pass. That is true whether we are considering the appearance of the Fibonacci series in nature, or human language, or anything else in the biological world. What has been “shown” or “persuasively argued” is that natural selection is plausibly taken to be a primary factor in evolution, as Darwin argued, and as no one (within the circles that Hauser considers) even questions; why he has decided that I (or anyone) have insisted that “natural selection theory cannot account for the design features of human language,” he does not say (and it is manifestly untrue, under the charitable reading required to grant the statement some meaning). Beyond the generally shared assumptions about natural selection and other mechanisms in evolution, one tries to find out what took place, whether studying the eye, the giraffe’s neck, the bones of the middle

ear, mammalian visual systems, human language, or anything else. Confident pronouncement is not to be confused with demonstration or even persuasive argument.

Though I suppose Hauser would deny this, it seems to me that on a close look, his actual conclusions do not differ much from the extreme skepticism of his Harvard colleague, evolutionary biologist Richard Lewontin, who concludes—forcefully—that the evolution of cognition is simply beyond the reach of contemporary science.¹⁵

The remoteness of the proclaimed goals leads to what seem to me some strange proposals: for example, that “the human brain, vocal tract, and language appear to have co-evolved” for the purposes of linguistic communication. Hauser is borrowing the notion of co-evolution of language and the brain from neuroscientist Terrence Deacon.¹⁶ Deacon argues that students of language and its ontogenesis—the first two perspectives of the ethological approach—are making a serious error when they adopt the standard approach of the neurosciences: seeking to discover a genetically determined component of the mind-brain and the state changes it undergoes through experience and maturation. They have overlooked a more promising alternative: “that the extra support for language learning,” beyond the data of experience, “is vested neither in the brain of the child nor in the brains of parents or teachers, but outside brains, in language itself.” Language and languages are extra-human entities with a remarkable “capacity . . . to evolve and adapt with respect to human hosts.” These creatures are not only extra-human, but apparently outside the biological world altogether.

What are these strange entities, and where did they come from? What they are is left unstated, except that they have evolved to incorporate the properties of language that have been mistakenly attributed to the brain. Their origin is no less mysterious, though once they somehow appeared, “the world’s languages evolved spontaneously,” through natural selection, in a “flurry of adaptation” that has “been going on outside the human brain.” They have thereby “become better and better adapted to people”—like parasites and hosts, or perhaps prey and predator in the familiar cycle of co-evolution; or perhaps viruses provide the best analogy, he suggests. We also derive an account of language universals: they have “emerged spontaneously and independently in each evolving language . . . They are convergent features of language evolution,” like the dorsal fins of sharks and dolphins. Having evolved spontaneously and acquired the universal properties of language by rapid natural selection, one of these extra-human creatures attaches itself to my granddaughter in New England, and a different one to my granddaughter in Nicaragua—actually she is infected by two of these mysterious viruses. It is a mistake to seek an explanation of the outcome in these and all other cases

by investigating the interplay of experience and innate structure of the brain; rather, the right parasites attach themselves to hosts in a particular community in some mystical fashion—by a “magician’s trick,” to borrow Deacon’s term for the ordinary assumptions of naturalistic science—yielding their knowledge of specific languages.

Deacon agrees, of course, that infants are “predisposed to learn human languages” and “are strongly biased in their choices” of “the rules underlying language,” acquiring within a few years “an immensely complex rule system and a rich vocabulary” at a time when they cannot even learn elementary arithmetic. So there is “something special about human brains that enables us to do with ease what no other species can do even minimally without intense effort and remarkably insightful training.” But it is a mistake to approach these predispositions and special structures of the brain the way we do other aspects of nature—the visual system, for example; no one would propose that insect and mammalian visual organs evolved spontaneously by rapid natural selection and now attach themselves to hosts, yielding the visual capacities of bees and monkeys; or that the waggle dance of bees or the calls of vervets are organism-external parasites that have co-evolved to provide the capacities of the host. But in the special case of human language, we are not to pursue the normal course of the natural sciences, seeking to determine the nature of the “predispositions” and “special structures” and the ways they are realized in brain mechanisms (in which case the extra-organic entities that have co-evolved with language vanish from the scene).

Since in this unique case extra-organic “viruses” have evolved that attach themselves to hosts in just the right way, we need not attribute to the child more than a “general theory of learning.” So we discover once we overcome the surprising failure of linguists and psychologists to recognize that the languages of the world—in fact, the possible languages that are as yet unspoken—may have evolved spontaneously, outside of brains, coming to “embody the predispositions of children’s minds” by natural selection.

There is, I think, a sense in which Deacon’s proposals are on the right track. The idea that a child needs no more than a “general theory of learning” to attain language and other cognitive states can be sustained only with quite heroic moves. That is a basic thrust of the third of the framework theses introduced at the outset, to which we return directly. Much the same conclusion is illustrated by the extraordinarily rich innatist and modular assumptions embedded within attempts to implement what are often misleadingly presented as unstructured general learning theories, and the no less extraordinary assumptions about innate structure built into approaches based on speculative evolutionary scenarios that explicitly assume extreme modularity.¹⁷

The only real problem, Deacon argues, is “symbolic reference.” The rest

will somehow fall into place if we account for this in evolutionary terms. How the rest falls into place is not discussed. But perhaps that does not matter, because “symbolic reference” is also left as a complete mystery, in part because of failure to attend to its most elementary properties in human language.

I have been giving quotes, because I have no idea what this means. And understanding is not facilitated by an account of “linguistics” (including views attributed to me) that is unrecognizable, with allusions so vague that it is often hard even to guess what might have been the source of the misunderstanding (sometimes it is easy; e.g., misunderstanding of terminology used in a technical sense, such as “competence”). Whatever the meaning may be, the conclusion seems to be that it is an error to investigate the brain to discover the nature of human language; rather, studies of language must be about the extra-biological entities that co-evolved with humans and somehow “latch on” to them. These proposals have been highly acclaimed by prominent evolutionary psychologists and biologists, but I do not see why. Taken at all seriously, they seem only to reshape standard problems of science as utter mysteries, placing them beyond any hope of understanding, while barring the procedures of rational inquiry that have been taken for granted for hundreds of years.

Returning to the methodological thesis that we should adopt an ethological approach, it is reasonable enough in principle, but the ways it is pursued raise many questions. As far as I can see, the renewed call to pursue this approach, as advocated forty years ago in the critical literature on “behavioral science,” leaves us about where we were. We can study the genetically determined component of the brain—and maybe more than the brain—that is dedicated to the structure and use of language, and the states it attains (the various languages), and we can investigate the process by which the state changes take place (language acquisition). We can try to discover the psychological and physiological mechanisms and principles, and to unify them, standard problems of science. These inquiries constitute the first two perspectives of the ethological approach: the study of mechanisms and ontogenesis. Turning to the third perspective, the functional perspective, we can investigate the use of language by the person who has attained a particular state, though the restriction to effects on survival and reproduction is far too narrow, if we hope to understand much about language. The fourth perspective—phylogenesis—seems a remote prospect at best, and does not seem much advanced by the comparative study of communication, a wholly different matter.

Let us turn finally to the third thesis I mentioned, quoting Gallistel: the substantive thesis that in all animals, learning is based on specialized mechanisms, “instincts to learn” in specific ways; what Tinbergen called “innate dispositions to learn.”¹⁸ These “learning mechanisms” can be regarded as

“organs within the brain [that] are neural circuits whose structure enables them to perform one particular kind of computation,” as they do more or less reflexively apart from “extremely hostile environments.” Human language acquisition is instinctive in this sense, based on a specialized “language organ.” This “modular view of learning” Gallistel takes to be “the norm these days in neuroscience.” He argues that this framework includes whatever is fairly well understood, including conditioning, insofar as it is a real phenomenon. “To imagine that there exists a general purpose learning mechanism in addition to all these problem-specific learning mechanisms . . . is like trying to imagine the structure of a general purpose organ, the organ that takes care of problems not taken care of by adaptively specialized organs like the liver, the kidney, the heart and the lungs,” or a “general purpose sensory organ, which solves the problem of sensing” for the cases not handled by the eye, the ear, and other specialized sensory organs. Nothing like that is known in biology: “Adaptive specialization of mechanism is so ubiquitous and so obvious in biology, at every level of analysis, and for every kind of function, that no one thinks it necessary to call attention to it as a general principle about biological mechanisms.” Accordingly, “it is odd but true that most past and contemporary theorizing about learning” departs so radically from what is taken for granted in the study of organisms—a mistake, he argues.

As far as I know, the approach Gallistel recommends is sound; in the special case of language, it seems to me to be adopted by all substantive inquiry, at least tacitly, even when that is heatedly denied. It is hard to avoid the conclusion that a part of the human biological endowment is a specialized “language organ,” the faculty of language (FL). Its initial state is an expression of the genes, comparable to the initial state of the human visual system, and it appears to be a common human possession to close approximation. Accordingly, a typical child will acquire any language under appropriate conditions, even under severe deficit and in “hostile environments.” The initial state changes under the triggering and shaping effect of experience, and internally determined processes of maturation, yielding later states that seem to stabilize at several stages, finally at about puberty. We can think of the initial state of FL as a device that maps experience into state L attained: a “language acquisition device” (LAD). The existence of such a LAD is sometimes regarded as controversial, but it is no more so than the (equivalent) assumption that there is a dedicated “language module” that accounts for the linguistic development of an infant as distinct from that of her pet kitten (or chimpanzee, or whatever), given essentially the same experience. Even the most extreme “radical behaviorist” speculations presuppose (at least tacitly) that a child can somehow distinguish linguistic materials from the rest of the confusion around it, hence postulating the existence of FL (= LAD);¹⁹ and as discussion of language

acquisition becomes more substantive, it moves to assumptions about the language organ that are more rich and domain specific, without exception to my knowledge. That includes the acquisition of lexical items, which turn out to have rich and complex semantic structure, even the simplest of them. Knowledge of these properties becomes available on very limited evidence and, accordingly, would be expected to be essentially uniform among languages; and is, as far as is known.

Here we move to substantive questions within the first three perspectives of the ethological approach, though again without restricting inquiry into language use to fitness consequences: survival and reproduction. We can inquire into the fundamental properties of linguistic expressions, and their use to express thought, sometimes to communicate, and sometimes to think or talk about the world. In this connection, comparative animal research surely merits attention. There has been important work on the problem of representation in a variety of species. Gallistel introduced a compendium of review articles on the topic a few years ago by arguing that representations play a key role in animal behavior and cognition; here “representation” is understood as isomorphism, a one-to-one relation between mind-brain processes and “an aspect of the environment to which these processes adapt the animal’s behavior”—e.g., when an ant represents the corpse of a conspecific by its odor.²⁰ It is a fair question whether, or how, the results relate to the mental world of humans; in the case of language, to what is called “phonetic” or “semantic representation.”

As noted, from the biolinguistic point of view that seems to me appropriate—and tacitly adopted in substantive work—we can think of a particular language *L* as a state of FL. *L* is a recursive procedure that generates an infinity of expressions. Each expression can be regarded as a collection of information for other systems of the mind-brain. The traditional assumption, back to Aristotle, is that the information falls into two categories, phonetic and semantic; information used, respectively, by sensorimotor systems and conceptual-intentional systems—the latter “systems of thought,” to give a name to something poorly understood. That could well be a serious oversimplification, but let us keep to the convention. Each expression, then, is an internal object consisting of two collections of information: phonetic and semantic. These collections are called “representations,” phonetic and semantic representations, but there is no isomorphism holding between the representations and aspects of the environment. There is no pairing of internal symbol and thing represented, in any useful sense.

On the sound side, this is taken for granted. It would not be false to say that an element of phonetic representation—say the internal element /ba/ in my language—picks out a thing in the world, namely the sound BA. But that

would not be a helpful move, and it is never made. Rather, acoustic and articulatory phonetics seek to understand how the sensorimotor system uses the information in the phonetic representation to produce and interpret sounds, no trivial task. One can think of the phonetic representation as an array of instructions for the sensorimotor systems, but a particular element of the internal representation is not paired with some category of events in the outside world, perhaps a construction based on motions of molecules. Similar conclusions seem to me appropriate on the meaning side. It has been understood at least since Aristotle that even the simplest words incorporate information of many different kinds: about material constitution, design and intended use, origin, gestalt and causal properties, and much more. These topics were explored in some depth during the cognitive revolution of the seventeenth and eighteenth centuries, though much of the work, even including the well-studied British empiricist tradition from Hobbes to Hume, remains little known outside of historical scholarship. The conclusions hold for simple nouns, count and mass—"river," "house," "tree," "water," personal and place names—the "purest referential terms" (pronouns, empty categories), and so on; and the properties become more intricate as we turn to elements with relational structure (verbs, tense and aspect, . . .), and of course far more so as we move on to more complex expressions. As to how early in ontogenesis these complex systems of knowledge are functioning, little is known, but there is every reason to suppose that the essentials are as much a part of the innate biological endowment as the capacity for stereoscopic vision or specific kinds of motor planning, elicited in considerable richness and specificity on the occasion of sense, in the terminology of the early modern scientific revolution.

There seems nothing analogous in the rest of the animal world, even at the simplest level. It is doubtless true that the massive explosion of lexicon, and symbolic representation, are crucial components of human language, but invoking imitation or symbol-thing correspondence does not carry us very far, and even those few steps could well be on the wrong track. When we turn to the organization and generation of representations, analogies break down very quickly beyond the most superficial level.

These properties of language are almost immediately obvious on inspection—which is not to say that they are deeply investigated or well understood; they are not. Moving beyond, we find other properties that are puzzling. The components of expressions—their features, in standard terminology—must be interpretable by the systems that access them; the representations at the interface with sensorimotor and thought systems consist of interpretable features. One would therefore expect that the features that enter computation should be interpretable, as in well-designed artificial symbolic systems: for-

mal systems for metamathematics, computer languages, etc. But it is not true for natural language; on the sound side, perhaps never true. One crucial case has to do with inflectional features that receive no semantic interpretation: structural case (nominative, accusative), or agreement features such as plurality (interpretable on nouns, but not on verbs or adjectives). The facts are not obvious in surface forms, but are reasonably well substantiated. Work of the past twenty years has provided considerable reason to suspect that these systems of uninterpretable features are quite similar among languages, though the external manifestation of the features differs in fairly systematic ways; and that a good deal of the typological variety of language reduces to this extremely narrow subcomponent of language. It could be, then, that the recursive computational system of the language organ is fixed and determinate, an expression of the genes, along with the basic structure of possible lexical items. A particular state of FL—a particular internal language—is determined by selecting among the highly structured possible lexical items and fixing parameters that are restricted to uninterpretable inflectional features and their manifestation. It could be that that is not a bad first approximation, maybe more than that.

It seems that the same uninterpretable features may be implicated in the ubiquitous dislocation property of natural language. The term refers to the fact that phrases are commonly articulated in one position but interpreted as if they were somewhere else, where they can be in similar expressions: the dislocated subject of a passive construction, for example, interpreted as if it were in the object position, in a local relation to the verb that assigns it a semantic role. Dislocation has interesting semantic properties. It may be that the “external” systems of thought (external to FL, internal to the mind-brain) require that FL generate expressions with these properties, to be properly interpreted. There is also reason to believe that the uninterpretable features may be the mechanism for implementing the dislocation property, perhaps even an optimal mechanism for satisfying this externally imposed condition on the language faculty. If so, then neither the dislocation property nor uninterpretable features are “imperfections” of FL, “design flaws” (here using the term “design” metaphorically, of course). These and other considerations raise more general questions of optimal design: could it be that FL is an optimal solution to interface conditions imposed by the systems of the mind-brain in which it is embedded, the sensorimotor and thought systems?

Such questions have been seriously posed only quite recently. They could not be raised before there was a fairly good grasp of the fixed principles of the faculty of language and the restricted options that yield the rich typological variety that we know must be rather superficial, despite appearances, given the empirical conditions on language acquisition. Though naturally partial

and tentative, such understanding has increased markedly in the past twenty years. Now it seems that questions of optimal design can be seriously raised, sometimes answered. Furthermore, the idea that language may be an optimal solution to interface conditions, in nontrivial respects, seems a good deal more plausible than it did a few years ago. Insofar as it is true, interesting questions arise about the theory of mind, the design of the brain, and the role of natural law in the evolution of even very complex organs such as the language faculty, questions that are very much alive in the theory of evolution at elementary levels, in work of the kind pioneered by D'Arcy Thompson and Alan Turing that has been somewhat at the margins until recently. It is conceivable that the comprehensive ethological approach discussed earlier might be enriched in these terms, though that remains a distant prospect.

Still more remote are the fundamental questions that motivated the classical theory of mind—the creative aspect of language use, the distinction between action appropriate to situations and action caused by situations, between being “compelled” to act in certain ways or only “incited and inclined” to do so; and in general, the question of how “members of animal bodies move at the command of the will,” Newton’s phrase in his review of mysteries that remain unresolved, including the causes of interaction of bodies, electrical attraction and repulsion, and other basic issues that remained unintelligible, by the standards of the scientific revolution.

In some domains, inquiry into components of the mind-brain has made dramatic progress. There is justified enthusiasm about the promise of new technologies, and a wealth of exciting work waiting to be undertaken in exploring mental aspects of the world and their emergence. It is not a bad idea, however, to keep in some corner of our minds the judgment of great figures of early modern science—Galileo, Newton, Hume and others—concerning the “obscurity” in which “nature’s ultimate secrets ever will remain,” perhaps for reasons rooted in the biological endowment of the curious creature that alone is able even to contemplate these questions.