Title: Biolinguistic Explorations: Design, Development, Evolution
Professor Noam Chomsky
May 10, 2006
West Hall, Bathish Auditorium, AUB

Biolinguistic Explorations: design, development, evolution

The problem that has virtually defined the serious study of language since its ancient origins has been to identify the specific nature of this distinctive human possession. Fifty years ago, a new approach began to take shape. The new approach regarded the faculty of language as an organ of the body, on a par with the visual system, the system of motor planning, the immune systems, and other subcomponents of the organism that interact in its growth and life. In this case, it is one of the "cognitive organs," like the faculties of reasoning, interpretation, planning, and so on. This "biolinguistic perspective," as it came to be called, focuses on the traditional problem of determining the specific nature of the faculty of language, and reinterprets it as the problem of discovering the genetic endowment that underlies the acquisition and use of language. This component of human biology is of particular significance because the language faculty is a species property: shared among humans with no detectable variation apart from pathology, and unique to humans in its essential properties. It is also surely central to human life in all of its aspects. Within the "biolinguistic perspective," language is understood to be internal language, a state of the language faculty, like a state of the visual system – which, like the language faculty, can vary somewhat within narrow bounds depending on early visual experience. The term "biolinguistics" itself was coined in 1974 as the topic for an international conference that brought together evolutionary biologists, neuroscientists, linguists, and others concerned with language and biology, one of many such initiatives, before and since.

I'd like to say a few words on where I think this enterprise stands today, keeping mostly to some leading ideas, and also straying a little into areas that are not narrowly linguistic, but that have for centuries been considered closely related to language, with an interesting revival in recent years, particularly in domain of moral philosophy.

Inquiry into these two domains proceeds along almost parallel paths: almost parallel because they should meet short of infinity, though where is far from clear. One path seeks to understand more about language and mind. The other is guided by concerns for freedom and justice. There should be some shared elements: in particular, what the co-founder of modern evolutionary theory, Alfred Russel Wallace, called "man's intellectual and moral nature": the human capacities for creative imagination, language and symbolism generally, interpretation and recording of natural phenomena, intricate social practices and the like, a complex that seems to have crystallized fairly recently among a small group in East Africa of which we are all descendants, sometimes called simply "the human capacity." The archaeological record suggests that the crystallization was not only recent but sudden in evolutionary time; some eminent scientists call the events "the great leap forward," which distinguished contemporary humans sharply from other animals. It's commonly assumed that the emergence of language was a core element in the great leap forward, and language is one component of the human capacity that is accessible to study in some depth. The principles of our intellectual and moral nature remain a considerable mystery, but we can hardly doubt that they are rooted in our very nature and play a central role in our lives. This is conventionally denied, but not credibly, in my opinion.

Puzzlement about the human capacity goes back to the origins of reflective thought in every culture we know much about. In the West, there was another kind of "great leap forward" in the course of the intellectual revolutions of the 17th century, which set the stage for modern science, and also included a cognitive revolution that is a precursor of what is commonly called "the cognitive revolution" of the 1950s, better considered a second cognitive revolution, which rediscovered themes that had been developed in thoughtful and instructive ways from which we still have a lot to learn.

One theme, developed by the physician/philosopher Juan Huarte in the16th century, is that the essential property of human intelligence is the capacity of the human mind to “engender within itself, by its own power, the principles on which knowledge rests.” Recognition that the capacity to gain understanding relies on fixed principles of the mind is as old as Plato, who could suggest only that the principles are remembered from an earlier life. The 17th century cognitive revolution sought to purge Plato’s theory of “the error of reminiscence,” as Leibniz put it, which means attributing to the mind itself an invariant innate structure. Today we try to reformulate that in terms of some genetic endowment, which enables us to construct a world of human experience from chaotic stimuli (different from the world constructed by other animals) and to rely on internal principles to develop our conception of the world in which we live, including language, moral codes, and a great deal more. Then further questions arise: questions about how these systems are put to use in our lives, how they are “encoded” in (mostly) the brain, how they evolved, and so on.

For Cartesian scientists, as for Huarte much earlier, the fundamental nature of human intelligence was a creative capacity, the basic distinction between humans and other organisms or machines: called by historians of philosophy/science the “beast-machine hypothesis.” Much like scientists in the past half century, the Cartesians were greatly intrigued by the capacities of machines, in our day computers, for them the intricate and quite remarkable artifacts constructed by skilled artisans. And they sought to understand how much of the world could be understood by regarding it as a machine – one constructed by the most skilled artisan of all. Descartes’s conclusion was that machines included animals, and in fact virtually everything about humans as well, apart from the creative capacity, which for him and his followers was exhibited most clearly in the use of language to express thought: the capacity of every person, but no beast or machine, to construct linguistic expressions without bound, undetermined by external stimuli or internal physical states, appropriate to circumstances but not caused by them – a fundamental distinction -- and eliciting thoughts in others that they could have expressed the same way. Use of language and other human action is “incited and inclined” by external and internal factors, but not “compelled,” in their terms – an observation that we have no reason to doubt, I believe. Cartesian psychologists developed experimental tests to determine whether another creature has these creative capacities, and concluded that if it passes the strictest tests we can devise, it would only be reasonable to conclude that it has a mind like ours. Notice that this is a more sophisticated and scientifically sound version of what has come to be called “the Turing test” in recent years, based, I believe, on a serious misunderstanding of a brief paper of Alan Turing’s, the very prominent mathematician who provided much of the theoretical basis for modern computing, and in fact much more. The 17th century concerns also led to investigation of the nature of concepts, some of them called “common notions” or “innate ideas,” which are part of our nature, developed or refined through experience. They also contributed to investigation of the principles by which linguistic expressions are constructed, a research program that came to be called philosophical grammar, meaning what we would call scientific grammar – the philosophy-science distinction is more recent, from the 19th century.

David Hume recognized that knowledge and belief are grounded in a “species of natural instincts,” part of the “springs and origins” of our inherent mental nature. He recognized that something similar must be true in the domain of moral judgment as well. The reason is that our moral judgments are unbounded in scope and that we constantly apply them in systematic ways to new circumstances. Hence they too must be founded on general principles that are part of our nature though beyond our “original instincts,” the instincts shared with animals. That should lead directly to efforts to develop something like a grammar of moral judgment. That task that was undertaken by the leading American moral and political philosopher of the late 20th century, John Rawls, who modeled his efforts in part on the theories of generative grammar that were being developed as he was writing his classic Theory of Justice in the 1960s. In the last few years these ideas have been revived in important work by John Mikhail, now a law professor at Georgetown University, and have become a lively field of theoretical and empirical inquiry as well.

It’s worth stressing that these questions can be sensibly raised only if we regard the human capacity as an internal property of people – part of their organic nature – a cognitive organ, a “mental organ,” where the term “mental” should be understood as referring simply to certain aspects of the world, to be studied in the same way as chemical, optical, electrical and other aspects.

For Cartesian scientists, in contrast, the domain of the mental was distinct from body, the physical world. That was a reasonable scientific thesis, based on the plausible belief that elements of the human capacity, specifically the creative aspect of language use, cannot be explained in terms of physical mechanisms. That belief was undermined by Newton, in an entirely unexpected way. Newton did not exorcise the ghost in the machine, as commonly believed. Rather he exorcised the machine, leaving the ghost intact. Newton demonstrated, to the consternation of scientists of the day including Newton himself, that we have no coherent notion of machine (or body, physical, material, etc.). We have no insight into nature beyond the best theories of various aspects of the world that we can devise, always looking forward to deeper explanation and eventual unification, though in unpredictable and often quite surprising ways. As historians of science have recognized, this intellectual move “set forth a new view of science” in which the goal is “not to seek ultimate explanations” but to find the best theoretical account we can of the phenomena of experience and experiment (I. Bernard Cohen). The goals of scientific inquiry were significantly lowered. It had been presumed, since Galileo, that the world itself would be intelligible to us, just as a complex artifact is. But Newton showed, to his dismay, that the expectation must be abandoned. The most we can hope for is that theories of the world will be intelligible to us, a very different perspective.

One consequence of the Newtonian revolution was that the mind-body problem in its traditional sense became unformulable, because the notion of body disappeared. It followed that “mental aspects of the world” should be considered on a par with all others. The theological framework aside, there was little alternative to the speculation by John Locke that came to be called “Locke’s suggestion” by historians of philosophy: that 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, again, much to his dismay. To the end of his life, Newton sought to find ways to overcome this conclusion, as did other leading scientist then and later, but always in vain.

In his classic history of materialism, Friedrich Lange, discussing “the real service rendered by Newton,” writes that “We have in our own days so accustomed ourselves to the abstract notion of forces, or rather to a notion hovering in a mystic obscurity between abstraction and concrete comprehension, that we no longer find any difficulty in making one particle of matter act upon another without immediate contact… From such ideas the great mathematicians and physicists of the seventeenth century were far removed. They were all in so far still genuine Materialists in the sense of ancient Materialism, that they made immediate contact a condition of influence. The collision of atoms or the attraction by hook-shaped particles, a mere modification of collision, were the type of all Mechanism and the whole movement of science tended towards Mechanism.” But no longer, with the collapse of the “mechanical philosophy” that regarded the world as a great machine.

All of this came to be pretty well understood through the 18th and 19th centuries, leading to the conclusion that properties “termed mental” reduce somehow to “the organical structure of the brain,” in the formulation of the famous chemist/philosopher Joseph Priestley in the late 18th century. However, it all seems to have been forgotten. Today we regularly read from eminent scientists and philosophers about the “astonishing hypothesis” that the mind is the result of the organization of the brain, a “bold hypothesis” and “radically new idea” in philosophy of mind – in fact, reiterating formulations of centuries ago in virtually the same words.

It is worth noting that the phrase “mind-body problem” is now commonly used for a very different problem, only loosely related to the traditional one, which was really part of normal science: certain phenomena could not be explained by the principles of the mechanical philosophy, the presupposed scientific theory of nature. Therefore a new principle was proposed. In terms of the metaphysics of the day, that meant postulating a new substance, res cogitans, a thinking substance, alongside of material substance. The next task for scientists would be to discover its properties and to try to unify the two substances. That task was undertaken, but was effectively terminated when Newton demolished the notion of material substance.

What is now called "the mind-body problem" is quite different. It is not part of normal science. The new version is based primarily on the distinction between the first person and the third person perspective. The first person perspective yields a view of the world presented by one's own experience – what the world looks like, feels like, sounds like to me, and so on. The third person perspective is the picture of the world developed in its most systematic form in scientific inquiry, which seeks to understand the world from outside any particular personal perspective. Its observations and conclusions are to be verifiable independent of any individual experience.

The new version of the mind-body problem resurrects some observations of Bertrand Russell's 80 years ago, and recently re-invented. Russell asked us to consider a blind physicist who knows all of physics but doesn't know something we know: what it's like to see the color blue. Russell's conclusion was that the natural sciences seek to discover "the causal skeleton of the world." Other aspects of the world of experience lie beyond their reach.

Recasting Russell's insight in naturalistic terms, we might say that like all animals, our internal cognitive capacities reflexively provide us with a world of experience, largely shared in fundamental properties – the human Umwelt, to borrow the term of ethologists. But being reflective creatures, thanks to emergence of the human capacity, we go on to seek to gain a deeper understanding of the phenomena of experience. These exercises are called myth, or magic, or philosophy, or "science" in the sense of that term proposed in the 19th century, distinguishing the pursuit from the rest of philosophy. If humans are part of the organic world, we expect that our capacities of understanding and explanation have fixed scope and limits, like any other natural object -- a truism that is sometimes thoughtlessly derided as "mysterianism." It could be that these innate capacities do not lead us beyond some understanding of Russell's causal skeleton of the world – including the principles that enter into determining conscious experience; there is of course no reason to expect that these are even in principle accessible to consciousness. It is always an open question how much of Russell's "causal skeleton of the world" can be attained. These could become topics of empirical inquiry into the nature of what we might call "the science-forming faculty," another "mental organ." These are interesting topics, in principle part of normal science, and now the topic of some investigation. They should not be confused with the traditional mind-body problem, which evaporated after Newton.

The cognitive revolution of the 17th century also led to inquiry into the nature of concepts, with important contemporary implications, also insufficiently appreciated. Aristotle had recognized that the objects to which we refer in using language cannot be identified by their material substance. A house, he pointed out, is not merely a collection of bricks and wood, but is defined in part by its function and design: a place for people to live and store their possessions, and so on. In Aristotle's terms, a house is a combination of matter and form. Notice that his account is metaphysical: he is defining what a house is, not the word or idea "house." That approach led to hopeless conundrums. The ship of Theseus is a classic case that may be familiar from philosophy courses; Saul Kripke's puzzle about belief is a modern variant. With the cognitive turn of the 17th century these questions were reframed in terms of operations of the mind: what does the word "house" mean, and how do we use it to refer. Pursuing that course we find that for natural language there is no word-object relation, where objects are mind-independent entities. That becomes very clear for Aristotle's example, the word house, when we look into its meaning more closely. Its "form" in the Aristotelian sense is vastly more intricate than he assumed. Furthermore, the conundrums based on the myth of a word-object relation dissolve, when viewed from this perspective, which I believe has ample empirical support.

One may be misled by the fact that houses are artifacts, so let's consider simpler cases that were investigated in the 17th-18th century classics. Take "river," a notion that Thomas Hobbes considered. He suggested that rivers are mentally individuated by origin. But while there is some truth to the observation, it is not really accurate, and it only scratches the surface of our intuitive understanding of the concept. Take the Charles River, which flows near my office. It would remain the very same river under quite extreme changes – among many others, reversing its course, so origin cannot be the defining characteristic. It would still be the Charles River if it were divided into separate streams that converge in some new place, or if any H20 that happens to be in it were replaced by chemicals from an upstream manufacturing plant. On the other hand, under trivial changes it would no longer be a river at all: for example, directing it between fixed boundaries and using it for shipping freight (in which case it is a canal, not a river), or hardening the surface by some almost undetectable physical change, painting a line down the middle, and using to drive to Boston (in which case it is a highway). And on to further intricacies as we can easily determine. Much the same holds of even the simplest concepts: "tree," "water," virtually any one we pick – perhaps any one, apart from invented technical terms, like "neutron," which means what we say it means. Exploring these concepts confirms the conclusion that the world of experience and interpretation is constructed by our rich "cognoscitive powers," as they were called, relying on gestalt properties, causal relations, and other mental notions explored by 17th century British philosophers and their successors.

Summarizing many years of discussion of such topics, David Hume concluded that "the identity we ascribe" to minds, vegetables, animal bodies and other entities is "only a fictitious one" established by the imagination, not a "peculiar nature belonging to this form." His conclusion is I think correct.

A notion of particular significance to Locke and his successors was person. Locke recognized that persons are identified in part by psychic continuity. It follows that a single body could house two distinct persons if two distinct psychic continuities were connected to it, and a person could persist even if its substance became completely different, properties of our concepts exploited by science fiction. These are all properties of the common notion or innate idea of person, which appears to be part of our intrinsic nature. Locke argued further that person, unlike tree or cat, is a term that involves "actions and their merit," a property that belongs only to "intelligent agents, capable of a law, and happiness, and misery," hence even more remote from mind-independent physical investigation, and at the heart of our moral faculty and intuitions.

These properties of innate ideas are available to very young children; a lot of children's literature is based on them, extending to animals as well. My grandchildren are entranced by the story of a baby donkey, named Sylvester, who is turned into a rock by some evil force, and spends the rest of the story trying to convince his parents that he is really Sylvester. Since children's stories have happy endings, at the end he is turned back to the baby donkey, and everyone lives happily ever after. But every child understands that the rock was really Sylvester all along, no matter what its material constitution was.

In all such cases, there is no mind-independent object, which could in principle be identified by a physicist, related to the name. As we proceed, we find much more intricate properties, no matter how simple the terms of language we investigate. As Hume and others recognized, for natural language and thought there is no meaningful word-object relation because we do not think or talk about the world in terms of mind-independent objects; rather, we focus attention on intricate aspects of the world by resort to our cognoscitive powers. Accordingly, for natural language and thought there is no notion of reference in the sense of the modern philosophical tradition, developed in the work of Frege, Peirce, Russell, Tarski, Carnap, Quine, and others, or contemporary theorists of reference: "externalists," in contemporary terminology. These technical concepts are fine for the purpose for which they were originally invented: formal systems where the symbols, objects, and relations are stipulated. Arguably they also provide a norm for science: its goal is to construct systems in which terms really do pick out an identifiable mind-independent element of the world, like "neutron," or "noun phrase." But human language and thought do not work that way.

If that is true – and the evidence seems overwhelming – then natural language diverges sharply in these elementary respects from animal communication, which has been held to rely on a one-one relation between mind/brain processes and "an aspect of the environment to which these processes adapt the animal's behavior" (quoting Randy Gallistel, a leading neuroscientist and specialist on animal communication): some particular type of event in the world elicits a warning call, for example. If so, then one basic problem facing the study of evolution of human language and thought is to account for the emergence of these universal and innate human properties, not a trivial matter, as we quickly discover when we take words and ideas seriously.

When we move beyond the simplest elements of language, problems of this kind escalate very sharply. That was discovered as soon as the long-forgotten tradition was revived in the 1950s. There is no time to discuss concrete examples, but what we find, quite generally, is that the surface forms that are produced and heard do not yield their interpretations in any straightforward way. The general conclusion is that even the simplest elements of language, including the simplest elements of sound and meaning, find their place in the internal computational processes that yield the expressions, but cannot in general be detected in the physical signal. For that reason, the language acquired must have the basic properties of an explanatory theory. These are design properties that must be dealt with by an account of the growth of language in the individual – language acquisition – and its evolution in the species, alongside of the elementary semantic properties of the simplest words.

Quite generally, construction of theories must be guided by what the philosopher and logician Charles Sanders Peirce a century ago called an "abductive principle," genetically determined, which "puts a limit upon admissible hypotheses," so that the mind is capable of "imagining correct theories of some kind" and discarding infinitely many others consistent with the evidence. Peirce was concerned with the development of scientific theories, but his observation holds as well for the growth of language -- language acquisition, as it is called. In this case, the principles that limit admissible hypotheses must be highly restrictive, given the empirical facts of rapidity of acquisition and convergence among individuals. That much was recognized from the earliest modern work. But the observation confronted what seemed to be impossible barriers. To see why, we have to look more closely at the problems.

In the 1950s, when this work began to take shape, the prevailing ideas in both linguistics and general biology adopted a particular interpretation of what may be the most quoted passage in biology, Darwin's concluding words of the first edition of Origin of Species: that "from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved." A widely-held view among evolutionary biologists was that the variability of organisms is so free as to constitute "a near infinitude of particulars which have to be sorted out case by case" (Gunther Stent). A similar view was held in the study of language, the belief that languages can "differ from each other without limit and in unpredictable ways," so that the study of each language must be approached "without any preexistent scheme of what a language must be" (linguist Martin Joos). Although they cannot literally be correct, the statements did capture the spirit of prevailing conceptions, which had substantial motivation: languages and organisms do appear on the surface to exhibit virtually endless variety. In both biology and linguistics, these basically incoherent assumptions have been gradually shown to be untenable over the years, in somewhat parallel ways, and with some interaction. There is reason to hope that the analogies may become more substantive in coming years. I will return briefly to that.

The prevailing conceptions posed a kind of paradox, in both domains. For the study of language, considerations of acquisition entailed that the innate preexisting scheme must be highly restricted so as to limit admissible hypotheses. But that runs counter to the apparent limitless variety of attainable languages. There is a similar paradox in general biology. In the study of language, about 25 years ago efforts to overcome the paradox converged in an approach that seemed to offer a kind of conceptual breakthrough. This approach proposed a sharp distinction between process of acquisition and the format of the internal theory of language. The basic idea is that the genetic endowment provides a fixed system of principles that interact in the generation of the expressions of a language, and a fixed set of parameters that are assigned a value by exposure to external data. A language is determined by setting the values and selecting lexical items from a fixed and highly structured store, with properties of the kind I've mentioned. The promissory note is to be paid by showing that the parameters are simple enough so that the values can be set by the data available to children, and that the fixed principles, with parameters set, function automatically to yield the rich variety of expressions for every possible human language, with their specific interpretations. That is a huge empirical task, of course, but one that has been pursued with considerable success.

This revised approach largely emerged from intensive study of a range of languages, but it was also suggested by an analogy to new developments in biology, specifically the discovery of regulatory mechanisms, and

Nobel laureate François Jacob's speculations about how slight changes in these mechanisms might yield great superficial differences – between a butterfly or an elephant, and so on. The model seemed natural for language as well: slight changes in parameter settings might yield superficial variety, through interaction of invariant principles with parameter choices.

The approach also removed a crucial conceptual barrier to the study of evolution of language. In earlier approaches, it appeared that the genetically-determined format for language had to be highly articulated, richly structured, and specific to language, sharply restricting admissible hypotheses so as to account for the rapidity and convergence of acquisition. Plainly, the greater the richness and specificity of the format, the harder it will be to account for its evolution. But by divorcing acquisition from the format for language, that conceptual problem is overcome. It may still turn out that the innate mechanisms are rich, highly structured, and specific to language, but that is no longer a conceptual necessity, within the new framework.
It therefore became possible to pursue more seriously the recognition, from the earliest days of generative grammar, that acquisition of language involves not just a few years of experience and millions of years of evolution, yielding the genetic endowment, but it also involves "principles of neural organization that may be even more deeply grounded in physical law" (quoting from work 40 years ago). Meanwhile what some biologists now call the evo-devo revolution was increasingly showing that "the rules controlling embryonic development" interact with other physical conditions "to restrict possible changes of structures and functions" in evolutionary development, providing "architectural constraints" that "limit adaptive scope and channel evolutionary patterns" (quoting Jacob and others). By the 1980s, evo-devo discoveries were reviving interest in ideas of rational morphology that pre-date Darwin, and observations in the early days of modern evolutionary biology that there appear to be "predetermined lines of modification" that lead natural selection to "produce varieties of a limited number and kind" for every species (Darwin's colleague Thomas Huxley. By now such considerations have been adduced for a wide range of problems of development and evolution, from cell division to optimization of structure and function of cortical networks, and just recently proposals as to how principles of physics determine that vertebrate brains have a certain distribution of gray and white matter, with specific characteristics. Such discoveries, once again, have suggestive similarities to ongoing linguistic investigation, in this case efforts to show that general principles of computational efficiency determine specific properties of the language faculty, raising the question to what extent the linguistic organ is also optimally designed.

Assuming that language has general properties of other biological systems, we should be seeking three factors that enter into the growth of language in the individual: (1) Genetic factors, which interpret part of the environment as linguistic experience, and determine the general course of development to the languages attained; (2) Experience, which permits variation within a fairly narrow range; (3) Principles not specific to the faculty of language.

The third factor includes principles of efficient computation, which would be expected to be of particular significance for systems such as language, determining the general character of attainable languages.

An elementary fact about the language faculty is that it is a system of discrete infinity. In the simplest case, such a system is based on a primitive operation that takes objects already constructed, and constructs from them a new object. Call that operation Merge. If computation is efficient, then when X and Y are merged, neither will change, so that the outcome can be taken to be simply the set {X, Y}: sometimes called the No-Tampering Condition, an obvious principle of efficient computation, plausibly understood as a law of nature. With Merge available, we instantly have an unbounded system of hierarchically structured expressions. For language to be usable, these expressions have to link to language-external but organism-internal systems, the systems of thought and the sensorimotor system. We can think of each generated expression as a collection of instructions for such systems: instructions for producing and perceiving external signals, and for internal thought, planning, interpretation, and organizing such actions as referring to the world from certain perspectives. The generated expressions provide the means to relate sound and meaning, in traditional terms, a far more subtle and intricate process than had been assumed for millennia.

Rephrasing the observation, we can say that the expressions generated by the language have to satisfy the interface conditions determined by the systems with which it interacts. Insofar as third factor properties function, language will satisfy these conditions in an optimal way, meeting conditions of efficient computation. We can regard an account of some linguistic phenomena as principled insofar as it derives them by efficient computation satisfying interface conditions. A very strong proposal, sometimes called "the strong minimalist thesis," is that all phenomena of language have a principled account in this sense, that language is a perfect solution to interface conditions, the conditions it must satisfy if it is to be usable. If that thesis were true, language would be something like a snowflake, taking the form it does by virtue of natural law. Genetic endowment is the residue when this thesis is not satisfied. An account of the evolution of language will have to deal with the property of unbounded Merge, and whatever else remains in the genetic endowment.

Emergence of unbounded Merge at once provides a kind of "language of thought," an internal system to allow pre-existent conceptual resources to construct expressions of arbitrary richness and complexity. It is often argued that another independent language of thought must be postulated, but the arguments for that do not seem to me compelling, a matter that would carry us too far afield.

As a simple matter of logic, there are two kinds of Merge, external and internal. External Merge takes two objects, say "eat" and "apples," and forms the new object that correspond to "eat apples." Internal Merge – often called Move -- is the same, except that one of the objects is internal to the other. So applying Internal Merge to "John ate what," we form the new object corresponding to "what John ate what." At the semantic interface, both occurrences of "what" are interpreted: the first occurrence as an operator and the second as the variable over which it ranges, so that the expression means something like: "for which thing x, John ate the thing x." At the sensorimotor side, only one is pronounced, namely, the first occurrence. That illustrates the ubiquitous displacement property of language: items are commonly pronounced in one position but interpreted somewhere else as well: in the expression we hear as "what did John eat," we understand "what" to have the semantic properties of the object of "eat," as in "John ate something." The structurally highest copy must be pronounced, or there would be no indication that Internal Merge has applied. Failure to pronounce all other copies follows from third factor considerations of efficient computation, since it reduces the burden of repeated application of the rules that transform internal structures to phonetic form – a very considerable burden when we consider more complex cases. The questions do not arise on the semantic side. As always, there is more to say, but this seems the heart of the matter.

This simple example suggests that the relation of the internal language to the interfaces is asymmetrical. Optimal design yields the right properties at the semantic side, but causes processing problems at the sound side. To understand the perceived sentence "what did John eat," and far more complex cases, it is necessary to locate and fill in the missing element, a burden on speech perception (and a well-known problem in language processing systems). In this case, conditions of efficient computation conflict with facilitation of communication. Universally, languages prefer efficient computation. That appears to be true quite generally. If so, it appears that language evolved, and is designed, primarily as an instrument of thought, with externalization a secondary process.

There are other reasons to believe that something like that is true. One is that externalization appears to be independent of sensory modality, as has been learned from studies of sign language in recent years. The structural properties of sign and spoken language appear to be remarkably similar. Acquisition follows the same course, and neural localization seems to be similar as well. Children raised in bimodal environments – one parent speaking and the other signing – seem to exhibit no preference, treating the two languages just as if they were exposed to Spanish and English. That tends to reinforce the conclusion that language is optimized for the system of thought, with mode of externalization secondary.

There are more general reasons that suggest the same conclusion. The core principle of language, unbounded Merge, must have arisen from some rewiring of the brain, presumably not too long before the "great leap forward," hence very recently in evolutionary time. Such changes take place in an individual, not a group. The individual so endowed would have had many advantages: capacities for complex thought, planning, interpretation, and so on. The capacity would be transmitted to offspring, coming to dominate a small breeding group. At that stage, there would be an advantage to externalization, so the capacity would be linked as a secondary process to the sensorimotor system for externalization and interaction, including communication – a special case, at least if we invest the term "communication" with substantive content, not just using it for any form of interaction. It is not easy to imagine an account of human evolution that does not assume at least this much, in one or another form. And empirical evidence is needed for any additional assumption about the evolution of language.

We can, however, go beyond speculation. Investigation of language design can yield evidence on the relation of language to the interfaces. There is, I think, mounting evidence that the relation is asymmetrical in the manner indicated. If so, principled explanation will reduce to efficient computation and satisfaction of semantic interface conditions, with satisfaction of sensorimotor conditions a secondary process.

At this point we have to move on to more technical discussion than is possible here, but I think it is fair to say that there has been considerable progress in moving towards principled explanation in terms of third factor considerations, considerably sharpening the question of the specific properties that determine the nature of language – the core problem of theoretical study of language since its origins millennia ago, now taking quite new forms.

The quest for principled explanation faces daunting tasks. We can formulate the goals with reasonable clarity. With each step towards the goal we gain a clearer grasp of the essential nature of language. It should be kept in mind, however, that any such progress still leaves unresolved problems that have been raised for hundreds of years. Among these is the question how properties "termed mental" relate to "the organical structure of the brain," in the 18th century formulation, and also the mysterious problems of the creative and coherent ordinary use of language, a central problem of Cartesian science, still scarcely even at the horizons of inquiry. And perhaps, if understanding someday reaches that far, we may reach the point where understanding of language and of our moral nature may converge, lending some substance to tantalizing speculations about these topics centuries ago.