

Of Minds and Language

Noam Chomsky

This article reviews, and rethinks, a few leading themes of the biolinguistic program since its inception in the early 1950s, at each stage influenced by developments in the biological sciences. The following also discusses how the questions now entering the research agenda develop in a natural way from some of the earliest concerns of these inquiries.

Keywords: biolinguistics; I-language; mind; minimalism

I have been thinking about various ways to approach this opportunity,¹ and on balance, it seemed that the most constructive tack would be to review, and rethink, a few leading themes of the biolinguistic program since its inception in the early 1950s, at each stage influenced by developments in the biological sciences. And to try to indicate how the questions now entering the research agenda develop in a natural way from some of the earliest concerns of these inquiries. Needless to say, this is from a personal perspective. The term “biolinguistics” itself was coined by Massimo Piattelli-Palmarini as the topic for an international conference in 1974 (Piattelli-Palmarini 1974) that brought together evolutionary biologists, neuroscientists, linguists, and others concerned with language and biology, one of many such initiatives, including the Royaumont conference that Massimo brought up (Piattelli-Palmarini 1980).

As you know, the 1950s was the heyday of the behavioral sciences. B.F. Skinner’s William James lectures, which later appeared as *Verbal Behavior* (Skinner 1957), were widely circulated by 1950, at least in Cambridge, Mass., and soon became close to orthodoxy, particularly as the ideas were taken up by W.V. Quine in his classes and work that appeared a decade later in his *Word and Object* (Quine 1960). Much the same was assumed for human capacity and cultural variety generally. Zellig Harris’s (1951) *Methods of Structural Linguistics* appeared

Editor’s note: We are grateful to Noam Chomsky for offering this contribution to the first issue of *Biolinguistics*. This would not have been possible without the support of the editors of the volume in which this article is also going to appear: Massimo Piattelli-Palmarini, Pello Salaburu, and Juan Uriagereka (see fn. 1). We also would like to express our gratitude to Oxford University Press, and to John Davey in particular, for granting us the permission to publish Chomsky’s contribution here.

¹ Editors’ note: The San Sebastian Meeting, June 2006; see Piattelli-Palmarini, Uriagereka & Salaburu (in press).



at the same time, outlining procedures for the analysis of a corpus of materials from sound to sentence, reducing data to organized form, and particularly within American linguistics, was generally assumed to have gone about as far as theoretical linguistics could or should reach. The fact that the study was called “methods” reflected the prevailing assumption that there could be nothing much in the way of a theory of language, because 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,” the formulation of Martin Joos, summarizing the reigning “Boasian tradition,” as he plausibly called it. The dominant picture in general biology was in some ways similar, captured in Gunther Stent’s (much later) observation 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.”

European structuralism was a little different, but not much: Trubetzkoy’s *Anleitung*, a classic introduction of phonological analysis (Trubetzkoy 1936, 2001), was similar in conception to the American procedural approaches, and in fact there was very little beyond phonology and morphology, the areas in which languages do appear to differ very widely and in complex ways, a matter of some more general interest, so recent work suggests.

Computers were on the horizon, and it was also commonly assumed that statistical analysis of vast corpora should reveal everything there is to learn about language and its acquisition, a severe misunderstanding of the fundamental issue that has been the primary concern of generative grammar from its origins at about the same time: To determine the structures that underlie semantic and phonetic interpretation of expressions and the principles that enter into growth and development of attainable languages. It was, of course, understood from the early 1950s that as computing power grows, it should ultimately be possible for analysis of vast corpora to produce material that would resemble the data analyzed. Similarly, it would be possible to do the same with videotapes of bees seeking nourishment. The latter might well give better approximations to what bees do than the work of bee scientists, a matter of zero interest to them; they want to discover how it works, resorting to elaborate and ingenious experiments. The former is even more absurd, since it ignores the core problems of the study of language.

A quite separate question is whether various characterizations of the entities and processes of language, and steps in acquisition, might involve statistical analysis and procedural algorithms. That they do was taken for granted in the earliest work in generative grammar, for example, in my *Logical Structure of Linguistic Theory* (*LSLT*, Chomsky 1955). I assumed that identification of chunked word-like elements in phonologically analyzed strings was based on analysis of transitional probabilities — which, surprisingly, turns out to be false, as Thomas Gambell and Charles Yang discovered, unless a simple UG prosodic principle is presupposed. *LSLT* also proposed methods to assign chunked elements to categories, some with an information-theoretic flavor; hand calculations in that pre-computer age had suggestive results in very simple cases, but to my knowledge, the topic has not been further pursued.

Information theory was taken to be a unifying concept for the behavioral

sciences, along the lines of Warren Weaver's essay in Shannon & Weaver's (1949/1998) famous monograph. Within the engineering professions, highly influential in these areas, it was a virtual dogma that the properties of language, maybe all human behavior, could be handled within the framework of Markov sources, in fact very elementary ones, not even utilizing the capacity of these simple automata to capture dependencies of arbitrary length. The restriction followed from the general commitment to associative learning, which excluded such dependencies. As an aside, my monograph *Syntactic Structures* (Chomsky 1957) begins with observations on the inadequacy in principle of finite automata, hence Markovian sources, but only because it was essentially notes for courses at MIT, where their adequacy was taken for granted. For similar reasons, the monograph opens by posing the task of distinguishing grammatical from ungrammatical sentences, on the analogy of well-formedness in formal systems, then assumed to be an appropriate model for language. In the much longer and more elaborate unpublished monograph *LSLT* two years earlier (Chomsky 1955), intended only for a few friends, there is no mention of finite automata, and a chapter is devoted to the reasons for rejecting any notion of well-formedness: The task of the theory of language is to generate sound-meaning relations fully, whatever the status of an expression, and in fact much important work then and since has had to do with expressions of intermediate status — the difference, say, between such deviant expressions as (1a) and (1b), that is, empty category principle vs. subadjacency violations, still not fully understood.

- (1) a. * Which book did they wonder why I wrote?
 b. * Which author did they wonder why wrote that book?

There were some prominent critics, like Karl Lashley, but his very important work on serial order in behavior (Lashley 1951), undermining prevailing associationist assumptions, was unknown, even at Harvard where he was a distinguished professor. Another sign of the tenor of the times.

This is a bit of a caricature, but not much. In fact it is understated, because the prevailing mood was also one of enormous self-confidence that the basic answers had been found, and what remained was to fill in the details in a generally accepted picture.

A few graduate students in the Harvard-MIT complex were skeptics. One was Eric Lenneberg, who went on to found the biology of language; another was Morris Halle. One change over the past 50 years is that we've graduated from sharing a cramped office to being in ample adjacent ones. From the early 1950s, we were reading and discussing work that was then well outside the canon: Lorenz, Tinbergen, Thorpe, and other work in ethology and comparative psychology. Also D'Arcy Thompson (1917/1992), though regrettably we had not come across Turing's work in biology (Turing 1952), and his thesis that "we must envisage a living organism as a special kind of system to which the general laws of physics and chemistry apply [...] and because of the prevalence of homologies, we may well suppose, as D'Arcy Thompson has done, that certain physical processes are of very general occurrence." The most recent evaluation of these aspects of Turing's work that I've seen, by Justin Leiber (2001), concludes that

Thompson and Turing “regard teleology, evolutionary phylogeny, natural selection, and history to be largely irrelevant and unfortunately effective distractions from fundamental ahistorical biological explanation,” the scientific core of biology. That broad perspective may sound less extreme today after the discovery of master genes, deep homologies, conservation, optimization of neural networks of the kind that Chris Cherniak has demonstrated, and much else, perhaps even restrictions of evolutionary/developmental processes so narrow that “replaying the protein tape of life might be surprisingly repetitive” (quoting a report on feasible mutational paths recently published in *Science*, Weinreich *et al.* 2006, reinterpreting a famous image of Steve Gould’s). Another major factor in the development of the biolinguistic perspective was work in recursive function theory and the general theory of computation and algorithms, then just becoming readily available, making it possible to undertake more seriously the inquiry into the formal mechanisms of generative grammars that were being explored from the late 1940s.

These various strands could, it seemed, be woven together to develop a very different approach to problems of language and mind, taking behavior and corpora to be not the object of inquiry, as in the behavioral sciences and structural linguistics, but merely data, and not necessarily the best data, for discovery of the properties of the real object of inquiry: The internal mechanisms that generate linguistic expressions and determine their sound and meaning. The whole system would then be regarded as one of the organs of the body, in this case a cognitive organ, like the systems of planning, interpretation, reflection, and whatever else falls among those aspects of the world loosely “termed mental”, which reduce somehow to “the organical structure of the brain”. I’m quoting chemist/philosopher Joseph Priestley in the late 18th century, articulating a standard conclusion after Newton had demonstrated, to his great dismay and disbelief, that the world is not a machine, contrary to the core assumptions of the 17th century scientific revolution. It follows that we have no choice but to adopt some non-theological version of what historians of philosophy call “Locke’s suggestion”: 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 action at a distance, a revival of occult properties, many leading scientists argued (with Newton’s partial agreement).

It is of some interest that all of this seems to have been forgotten. The American Academy of Arts and Sciences published a volume summarizing the results of the Decade of the Brain that ended the 20th century (Mountcastle 1998). The guiding theme, formulated by Vernon Mountcastle, is the thesis of the new biology that “[t]hings mental, indeed minds, are emergent properties of brains, [though] these emergences are [...] produced by principles that [...] we do not yet understand” (Mountcastle 1998: 1). The same thesis has been put forth in recent years by prominent scientists and philosophers as an “astonishing hypothesis” of the new biology, a “radical” new idea in the philosophy of mind, “the bold assertion that mental phenomena are entirely natural and caused by the neuro-physiological activities of the brain,” opening the door to novel and promising inquiries, a rejection of Cartesian mind-body dualism, and so on. All,

in fact, reiterate formulations of centuries ago, in virtually the same words, after mind-body dualism became unformulable with the disappearance of the only coherent notion of body (physical, material, etc.) — facts well understood in standard histories of materialism, like Friedrich Lange's (1892) 19th century classic.

It is also of some interest that although the traditional mind-body problem dissolved after Newton, the phrase "mind-body problem" has been resurrected for a problem that is only loosely related to the traditional one. The traditional mind-body problem developed in large part within normal science: Certain phenomena could not be explained by the principles of the mechanical philosophy, the presupposed scientific theory of nature, so a new principle was proposed, some kind of *res cogitans*, a thinking substance, alongside of material substance. The next task would be to discover its properties and to try to unify the two substances. That task was undertaken, but was effectively terminated when Newton undermined 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 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 developed in its most systematic form in scientific inquiry, which seeks to understand the world from outside any particular personal perspective.

The new version of the mind-body problem resurrects a thought experiment of Bertrand Russell's 80 years ago, though the basic observation traces back to the pre-Socratics. 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: "It is obvious that a man who can see knows things which a blind man cannot know; but a blind man can know the whole of physics. Thus the knowledge which other men have and he has not is not part of physics" (Russell 2003: 227). Russell's conclusion was that the natural sciences seek to discover "the causal skeleton of the world. Other aspects lie beyond their purview" (*ibid.*).

Recasting Russell's experiment in naturalistic terms, we might say that like all animals, our internal cognitive capacities reflexively provide us with a world of experience — the human *Umwelt*, in ethological lingo. But being reflective creatures, thanks to emergence of human intellectual capacities, we go on to seek a deeper understanding of the phenomena of experience. 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," though it was understood by Descartes and Hume, among others. It could be that these innate capacities do not lead us beyond some theoretical understanding of Russell's causal skeleton of the world. In principle these questions are subject to empirical inquiry into what we might call "the science-forming faculty," another "mental organ," now the topic of some investigation — Susan Carey's work, for example. But these issues are distinct from traditional dualism, which evaporated after Newton.

This is a rough sketch of the intellectual background of the biolinguistic perspective, in part with the benefit of some hindsight. Adopting this perspective, the term “language” means internal language, a state of the computational system of the mind/brain that generates structured expressions, each of which can be taken to be a set of instructions for the interface systems within which the faculty of language is embedded. There are at least two such interfaces: The systems of thought that use linguistic expressions for reasoning, interpretation, organizing action, and other mental acts. And the sensorimotor systems that externalize expressions in production and construct them from sensory data in perception. The theory of the genetic endowment for language is commonly called universal grammar (UG), adapting a traditional term to a different framework. Certain configurations are possible human languages, others are not, and a primary concern of the theory of human language is to establish the distinction between the two categories.

Within the biolinguistic framework, several tasks immediately arise. The first is to construct generative grammars for particular languages that yield the facts about sound and meaning. It was quickly learned that the task is formidable. Very little was known about languages, despite millennia of inquiry. The most extensive existing grammars and dictionaries were, basically, lists of examples and exceptions, with some weak generalizations. It was assumed that anything beyond could be determined by unspecified methods of “analogy” or “induction” or “habit.” But even the earliest efforts revealed that these notions concealed vast obscurity. Traditional grammars and dictionaries tacitly appeal to the understanding of the reader, either knowledge of the language in question or the shared innate linguistic capacity, or commonly both. But for the study of language as part of biology, it is precisely that presupposed understanding that is the topic of investigation, and as soon as the issue was faced, major problems were quickly unearthed.

The second task is to account for the acquisition of language, later called the problem of explanatory adequacy (when viewed abstractly). In biolinguistic terms, that means discovering the operations that map presented data to the internal language attained. With sufficient progress in approaching explanatory adequacy, a further and deeper task comes to the fore: To transcend explanatory adequacy, asking not just what the mapping principles are, but why language growth is determined by these principles, not innumerable others that can be easily imagined. The question was premature until quite recently, when it has been addressed in what has come to be called the minimalist program, the natural next stage of biolinguistic inquiry, to which I’ll briefly return.

Another question is how the faculty of language evolved. There are libraries of books and articles about evolution of language — in rather striking contrast to the literature, say, on the evolution of the communication system of bees. For human language, the problem is vastly more difficult for obvious reasons, and can be undertaken seriously, by definition, only to the extent that some relatively firm conception of UG is available, since that is what evolved.

Still another question is how the properties “termed mental” relate to “the organical structure of the brain,” in Priestley’s words (see also Chomsky 1998). And there are hard and important questions about how the internal language is

put to use, for example in acts of referring to the world, or in interchange with others, the topic of interesting work in neo-Gricean pragmatics in recent years.

Other cognitive organs can perhaps be studied along similar lines. In the early days of the biolinguistic program, George Miller and others sought to construct a generative theory of planning, modeled on early ideas about generative grammar (Miller & Johnson-Laird 1976). Other lines of inquiry trace back to David Hume, who recognized that knowledge and belief are grounded in a “species of natural instincts,” part of the “springs and origins” of our inherent mental nature, and that something similar must be true in the domain of moral judgment. 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,” those shared with animals. That should lead to efforts to develop something like a grammar of moral judgment. That task was undertaken by John Rawls, who adapted models of generative grammar that were being developed as he was writing his classic *Theory of Justice* (1971) in the 1960s. These ideas have recently been revived and developed and have become a lively field of theoretical and empirical inquiry (cf. Hauser 2006).

At the time of the 1974 biolinguistics conference, it seemed that the language faculty must be rich, highly structured, and substantially unique to this cognitive system. In particular, that conclusion followed from considerations of language acquisition. The only plausible idea seemed to be that language acquisition is rather like theory construction. Somehow, the child reflexively categorizes certain sensory data as linguistic experience, and then uses the experience as evidence to construct an internal language — a kind of theory of expressions that enter into the myriad varieties of language use.

To give a few of the early illustrations for concreteness, the internal language that we more or less share determines that sentence (2a) is three-ways ambiguous, though it may take a little reflection to reveal the fact; but the ambiguities are resolved if we ask (2b), understood approximately as (2c).

- (2) a. Mary saw the man leaving the store.
 b. Which store did Mary see the man leaving?
 c. Which store did Mary see the man leave?

The phrase *which store* is raised from the position in which its semantic role is determined as object of *leave*, and is then given an additional interpretation as an operator taking scope over a variable in its original position, so the sentence means, roughly, *for which x, x a store, Mary saw the man leav(ing) the store x* — and without going into it here, there is good reason to suppose that the semantic interface really does interpret the variable *x* as *the store x*, a well-studied phenomenon called “reconstruction”. The phrase that serves as the restricted variable is silent in the phonetic output, but must be there for interpretation. Only one of the underlying structures permits the operation, so the ambiguity is resolved in the interrogative, in the manner indicated. The constraints involved — so-called “island conditions” — have been studied intensively for about 45 years. Recent work indicates that they may reduce in large measure to minimal search

conditions of optimal computation, perhaps not coded in UG but more general laws of nature — which, if true, would carry us beyond explanatory adequacy.

Note that even such elementary examples as this illustrate the marginal interest of the notions “well-formed” or “grammatical” or “good approximation to a corpus”, however they are characterized.

To take a second example, illustrating the same principles less transparently, consider sentences (3a) and (3b).

- (3) a. John ate an apple.
b. John ate.

We can omit *an apple*, yielding (3b), which we understand to mean *John ate something unspecified*. Now consider

- (4) a. John is too angry to eat an apple.
b. John is too angry to eat.

We can omit *an apple*, yielding (4b), which, by analogy to (3b) should mean that *John is so angry that he wouldn't eat anything*. That's a natural interpretation, but there is also a different one in this case: namely, *John is so angry that someone or other won't eat him, John* — the natural interpretation for the structurally analogous expression

- (5) John is too angry to invite.

In this case, the explanation lies in the fact that the phrase *too angry to eat* does include the object of *eat*, but it is invisible. The invisible object is raised just as *which store* is raised in the previous example (2), again yielding an operator-variable structure. In this case, however, the operator has no content, so the construction is an open sentence with a free variable, hence a predicate. The semantic interpretation follows from general principles. The minimal search conditions that restrict raising of *which store* in example (2) also bar the raising of the empty object of *eat*, yielding standard island properties.

In both cases, the same general computational principles, operating efficiently, provide a specific range of interpretations as an operator-variable construction, with the variable unpronounced in both cases and the operator unpronounced in one. The surface forms in themselves tell us little about the interpretations.

Even the most elementary considerations yield the same conclusions. The simplest lexical items raise hard if not insuperable problems for analytic procedures of segmentation, classification, statistical analysis, and the like. A lexical item is identified by phonological elements that determine its sound along with morphological elements that determine its meaning. But neither the phonological nor morphological elements have the “beads-on-a-string” property required for computational analysis of a corpus. Furthermore, even the simplest words in many languages have phonological and morphological elements that are silent. The elements that constitute lexical items find their place in the

generative procedures that yield the expressions, but cannot be detected in the physical signal. For that reason, it seemed then — and still seems — that the language acquired must have the basic properties of an internalized explanatory theory. These are design properties that any account of evolution of language must deal with.

Quite generally, construction of theories must be guided by what Charles Sanders Peirce a century ago called an “abductive principle,” which he took to be a genetically determined instinct, like the pecking of a chicken. The abductive principle “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 what I was calling “the science-forming faculty,” but similar problems arise for language acquisition, though it is dramatically unlike scientific discovery. It is rapid, virtually reflexive, convergent among individuals, relying not on controlled experiment or instruction but only on the “blooming, buzzing confusion” that each infant confronts. The format that limits admissible hypotheses about structure, generation, sound and meaning must therefore be highly restrictive. The conclusions about the specificity and richness of the language faculty follow directly. Plainly such conclusions make it next to impossible to raise questions that go beyond explanatory adequacy — the “why” questions — and also pose serious barriers to inquiry into how the faculty might have evolved, matters discussed inconclusively at the 1974 conference (see Piattelli-Palmarini 1974).

A few years later, a new approach suggested ways in which these paradoxes might be overcome. This Principles-and-Parameters (P&P) approach (Chomsky 1981 *et seq.*) was based on the idea that the format consists of invariant principles and a “switch-box” of parameters — to adopt Jim Higginbotham’s image. The switches can be set to one or another value on the basis of fairly elementary experience. A choice of parameter settings determines a language. The approach largely emerged from intensive study of a range of languages, but as in the early days of generative grammar, it was also suggested by developments in biology — in this case, François Jacob’s ideas about how slight changes in the timing and hierarchy of regulatory mechanisms might yield great superficial differences (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. That’s discussed a bit in Kant lectures of mine at Stanford in 1978, which appeared a few years later in my book *Rules and Representations* (Chomsky 1980).

The approach crystallized in the early 1980s, and has been pursued with considerable success, with many revisions and improvements along the way. One illustration is Mark Baker’s demonstration, in his book *Atoms of Language* (Baker 2001), that languages that appear on the surface to be about as different as can be imagined (in his case Mohawk and English) turn out to be remarkably similar when we abstract from the effects of a few choices of values for parameters within a hierarchic organization that he argues to be universal, hence the outcome of evolution of language.

Looking with a broader sweep, the problem of reconciling unity and diversity has constantly arisen in biology and linguistics. The linguistics of the

early scientific revolution distinguished universal from particular grammar, though not in the biolinguistic sense. Universal grammar was taken to be the intellectual core of the discipline; particular grammars are accidental instantiations. With the flourishing of anthropological linguistics, the pendulum swung in the other direction, towards diversity, well captured in the Boasian formulation to which I referred. In general biology, a similar issue had been raised sharply in the Cuvier–Geoffroy debate in 1830 (Appel 1987). Cuvier’s position, emphasizing diversity, prevailed, particularly after the Darwinian revolution, leading to the conclusions about near infinitude of variety that have to be sorted out case by case, which I mentioned earlier. Perhaps the most quoted sentence in biology is Darwin’s final observation in *Origin of Species* about how “from so simple a beginning, endless forms most beautiful and most wonderful have been, and are being, evolved.” I don’t know if the irony was intended, but these words were taken by Sean Carroll (2005) as the title of his introduction to *The New Science of Evo Devo*, which seeks to show that the forms that have evolved are far from endless, in fact are remarkably uniform, presumably, in important respects, because of factors of the kind that Thompson and Turing thought should constitute the true science of biology. The uniformity had not passed unnoticed in Darwin’s day. Thomas Huxley’s naturalistic studies led him to observe that there appear to be “predetermined lines of modification” that lead natural selection to “produce varieties of a limited number and kind” for each species.²

Over the years, in both general biology and linguistics the pendulum has been swinging towards unity, in the evo–devo revolution in biology and in the somewhat parallel minimalist program.

The principles of traditional universal grammar had something of the status of Joseph Greenberg’s universals: They were descriptive generalizations. Within the framework of UG in the contemporary sense, they are observations to be explained by the principles that enter into generative theories, which can be investigated in many other ways. Diversity of language provides an upper bound on what may be attributed to UG: It cannot be so restricted as to exclude attested languages. Poverty of stimulus (POS) considerations provide a lower bound: UG must be at least rich enough to account for the fact that internal languages are attained. POS considerations were first studied seriously by Descartes to my knowledge, in the field of visual perception. Of course they are central to any inquiry into growth and development, though for curious reasons, these truisms are considered controversial only in the case of language and other higher human mental faculties (particular empirical assumptions about POS are of

² The passage quoted is, in its entirety:

The importance of natural selection will not be impaired even if further inquiries should prove that variability is definite, and is determined in certain directions rather than in others, by conditions inherent in that which varies. It is quite conceivable that every species tends to produce varieties of a limited number and kind, and that the effect of natural selection is to favour the development of some of these, while it opposes the development of others along their predetermined lines of modification. (Huxley 1893: 223)

See also Gates (1916: 128) and Chomsky (2004).

course not truisms, in any domain of growth and development).

For these and many other reasons, the inquiry has more stringent conditions to satisfy than generalization from observed diversity. That is one of many consequences of the shift to the biolinguistic perspective; another is that methodological questions about simplicity, redundancy, and so on, are transmuted into factual questions that can be investigated from comparative and other perspectives, and may reduce to natural law.

Apart from stimulating highly productive investigation of languages of great typological variety, at a depth never before even considered, the P&P approach also reinvigorated neighboring fields, particularly the study of language acquisition, reframed as inquiry into setting of parameters in the early years of life. The shift of perspective led to very fruitful results, enough to suggest that the basic contours of an answer to the problems of explanatory adequacy might be visible. On that tentative assumption, we can turn more seriously to the “why” questions that transcend explanatory adequacy. The minimalist program thus arose in a natural way from the successes of the P&P approach.

The P&P approach also removed the major conceptual barrier to the study of evolution of language. With the divorce of principles of language from acquisition, it no longer follows that the format that “limits admissible hypotheses” must be rich and highly structured to satisfy the empirical conditions of language acquisition, in which case inquiry into evolution would be virtually hopeless. That might turn out to be the case, but it is no longer an apparent conceptual necessity. It therefore became possible to entertain 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 also “principles of neural organization that may be even more deeply grounded in physical law” (quoting from my *Aspects of the Theory of Syntax*, Chomsky 1965 — a question then premature).

Assuming that language has general properties of other biological systems, we should be seeking three factors that enter into its growth in the individual: (i) genetic factors, the topic of UG, (ii) experience, which permits variation within a fairly narrow range, and (iii) principles not specific to language. The third factor includes principles of efficient computation, which would be expected to be of particular significance for systems such as language. UG is the residue when third factor effects are abstracted. The richer the residue, the harder it will be to account for the evolution of UG, evidently.

Throughout the modern history of generative grammar, the problem of determining the general nature of language has been approached “from top down,” so to speak: How much must be attributed to UG to account for language acquisition? The minimalist program seeks to approach the problem “from bottom up”: How little can be attributed to UG while still accounting for the variety of internal languages attained, relying on third factor principles? Let me end with a few words on this approach.

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. There are more complex modes of generation, such as the familiar phrase structure grammars explored in the early years of generative grammar. But a Merge-based system is the most elementary, so we assume it to be true of language unless empirical facts force greater UG complexity. 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 $\{X, Y\}$. That is sometimes called the No-Tampering condition, a natural principle of efficient computation, perhaps a special case of laws 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 the interfaces. The generated expressions provide the means to relate sound and meaning in traditional terms, a far more subtle process than had been assumed for millennia. UG must at least include the principle of unbounded Merge.

The conclusion holds whether recursive generation is unique to the language faculty or found elsewhere. If the latter, there still must be a genetic instruction to use unbounded Merge to form linguistic expressions. Nonetheless, it is interesting to ask whether this operation is language-specific. We know that it is not. The classic illustration is the system of natural numbers, raising problems for evolutionary theory noted by Alfred Russel Wallace. A possible solution is that the number system is derivative from language. If the lexicon is reduced to a single element, then unbounded Merge will easily yield arithmetic. Speculations about the origin of the mathematical capacity as an abstraction from language are familiar, as are criticisms, including apparent dissociation with lesions and diversity of localization. The significance of such phenomena, however, is far from clear. As Luigi Rizzi has pointed out (Rizzi 2003), they relate to use of the capacity, not its possession; for similar reasons, dissociations do not show that the capacity to read is not parasitic on the language faculty. The competence-performance distinction should not be obscured. To date, I am not aware of any real examples of unbounded Merge apart from language, or obvious derivatives from language, for example, taking visual arrays as lexical items.

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, 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 to some extent if it is to be usable at all. If that thesis were true, language would be something like a snowflake, taking the form it does by virtue of natural law, in which case UG would be very limited.

In addition to unbounded Merge, language requires atoms, or word-like elements, for computation. Whether these belong strictly to language or are appropriated from other cognitive systems, they pose extremely serious problems for the study of language and thought and also for the study of the evolution of human cognitive capacities. The basic problem is that even the simplest words and concepts of human language and thought lack the relation to mind-independent entities that has been reported for animal communication:

Representational systems based on a one-to-one relation between mind/brain processes and “an aspect of the environment to which these processes adapt the animal’s behavior,” to quote Randy Gallistel. The symbols of human language and thought are sharply different.

These matters were explored in interesting ways by 17th-18th century British philosophers, developing ideas that trace back to Aristotle. Carrying their work further, we find that human language appears to have no reference relation, in the sense stipulated in the study of formal systems, and presupposed — mistakenly, I think — in contemporary theories of reference for language in philosophy and psychology, which take for granted some kind of word-object relation, where the objects are extra-mental. What we understand to be a house, a river, a person, a tree, water, and so on, consistently turns out to be a creation of what 17th century investigators called the “cognoscitive powers,” which provide us with rich means to refer to the outside world from certain perspectives. The objects of thought they construct are individuated by mental operations that cannot be reduced to a “peculiar nature belonging” to the thing we are talking about, as David Hume summarized a century of inquiry. There need be no mind-independent entity to which these objects of thought bear some relation akin to reference, and apparently there is none in many simple cases (probably all). In this regard, internal conceptual symbols are like the phonetic units of mental representations, such as the syllable /ba/; every particular act externalizing this mental entity yields a mind-independent entity, but it is idle to seek a mind-independent construct that corresponds to the syllable. Communication is not a matter of producing some mind-external entity that the hearer picks out of the world, the way a physicist could. Rather, communication is a more-or-less affair, in which the speaker produces external events and hearers seek to match them as best they can to their own internal resources. Words and concepts appear to be similar in this regard, even the simplest of them. Communication relies on shared cognoscitive powers, and succeeds insofar as shared mental constructs, background, concerns, presuppositions, etc. allow for common perspectives to be (more or less) attained. These semantic properties of lexical items seem to be unique to human language and thought, and have to be accounted for somehow in the study of their evolution.

Returning to the computational system, 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 corresponds 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*, in accord with the No-Tampering condition. As in the examples I mentioned earlier, 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 of the two identical syntactic objects is pronounced, typically the structurally most salient occurrence. That illustrates the ubiquitous displacement property of language: Items are commonly pronounced in one position but interpreted somewhere else as well. Failure to pronounce all but one occurrence 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 heavy burden when we consider real cases. 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 (6),

(6) What did John eat?

it is necessary to locate and fill in the missing element, a severe burden on speech perception in more complex constructions. Here conditions of efficient computation conflict with facilitation of communication. Universally, languages prefer efficient computation. That appears to be true more generally. For example, island conditions are at least sometimes, and perhaps always, imposed by principles of efficient computation. They make certain thoughts inexpressible, except by circumlocution, thus impeding communication. The same is true of ambiguities, as in the examples I mentioned earlier. Structural ambiguities often fall out naturally from efficient computation, but evidently pose a communication burden.

Other considerations suggest the same conclusion. Mapping to the sensorimotor interface appears to be a secondary process, relating systems that are independent: the sensorimotor system, with its own properties, and the computational system that generates the semantic interface, optimally insofar as the strong minimalist thesis is accurate. That's basically what we find. Complexity, variety, effects of historical accident, and so on, are overwhelmingly restricted to morphology and phonology, the mapping to the sensorimotor interface. That's why these are virtually the only topics investigated in traditional linguistics, or that enter into language teaching. They are idiosyncrasies, so are noticed, and have to be learned. If so, then it appears that language evolved, and is designed, primarily as an instrument of thought. Emergence of unbounded Merge in human evolutionary history provides what has been called a "language of thought," an internal generative system that constructs thoughts of arbitrary richness and complexity, exploiting conceptual resources that are already available or may develop with the availability of structured expressions. If the relation to the interfaces is asymmetric, as seems to be the case, then unbounded Merge provides only a language of thought, and the basis for ancillary processes of externalization.

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. More general considerations suggest the same conclusion. The core principle of language, unbounded Merge, must have arisen from some rewiring of the brain, presumably the effect of some small mutation. 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. It is not easy to imagine an account of human evolution that does not assume at least this much. And empirical evidence is needed for any additional assumption about the evolution of language.

Such evidence is not easy to find. It is generally supposed that there are precursors to language proceeding from single words, to simple sentences, then more complex ones, and finally leading to unbounded generation. But there is no empirical evidence for the postulated precursors, and no persuasive conceptual argument for them either: Transition from 10-word sentences to unbounded Merge is no easier than transition from single words. A similar issue arises in language acquisition. The modern study of the topic began with the assumption that the child passes through a one and two-word stage, telegraphic speech, and so on. Again the assumption lacks a rationale, because at some point unbounded Merge must appear. Hence the capacity must have been there all along even if it only comes to function at some later stage. There does appear to be evidence about earlier stages: namely, what children produce. But that carries little weight. Children understand far more than what they produce, and understand normal language but not their own restricted speech, as was shown long ago by Lila Gleitman and her colleagues (Shipley *et al.* 1969). For both evolution and development, there seems little reason to postulate precursors to unbounded Merge.

In the 1974 biolinguistics conference, evolutionary biologist Salvador Luria was the most forceful advocate of the view that communicative needs would not have provided “any great selective pressure to produce a system such as language,” with its crucial relation to “development of abstract or productive thinking.” His fellow Nobel laureate François Jacob (1977) added later that “the role of language as a communication system between individuals would have come about only secondarily, as many linguists believe,” perhaps referring to discussions at the symposia (for an insightful reconstruction of those debates, see also Jenkins 2000). “The quality of language that makes it unique does not seem to be so much its role in communicating directives for action” or other common features of animal communication, Jacob continues, but rather “its role in symbolizing, in evoking cognitive images,” in “molding” our notion of reality and yielding our capacity for thought and planning, through its unique property of allowing “infinite combinations of symbols” and therefore “mental creation of possible worlds,” ideas that trace back to the 17th century cognitive revolution and have been considerably sharpened in recent years.

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. There are more radical proposals under which optimal satisfaction of semantic conditions becomes close to tautologous. That seems to me one way to understand the general drift of Jim Higginbotham’s work on the syntax-semantics border for many years. And from a different point of view, something

similar would follow from ideas developed by Wolfram Hinzen (2006, 2007), in line with Juan Uriagereka's suggestion that it is "as if syntax carved the path interpretation must blindly follow."

The general conclusions appear to fit reasonably well with evidence from other sources. It seems that brain size reached its current level about 100,000 years ago, which suggests to some specialists that "human language probably evolved, at least in part, as an automatic but adaptive consequence of increased absolute brain size," leading to dramatic changes of behavior (quoting George Striedter, in *Brain and Behavioral Sciences* February 2006, who adds qualifications about the structural and functional properties of primate brains; Striedter 2006: 9). This "great leap forward," as some call it, must have taken place before about 50,000 years ago, when the trek from Africa began. Even if further inquiry extends the boundaries, it remains a small window, in evolutionary time. The picture is consistent with the idea that some small rewiring of the brain gave rise to unbounded Merge, yielding a language of thought, later externalized and used in many ways. Aspects of the computational system that do not yield to principled explanation fall under UG, to be explained somehow in other terms, questions that may lie beyond the reach of contemporary inquiry, Richard Lewontin (1998) has argued. Also remaining to be accounted for are the apparently human-specific atoms of computation, the minimal word-like elements of thought and language, and the array and structure of parameters, rich topics that I barely mentioned.

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. The best guess about the nature of UG only a few years ago has been substantially improved by approaching the topic "from bottom up", by asking how far we can press the strong minimalist thesis. It seems now that much of the architecture that has been postulated can be eliminated without loss, often with empirical gain. That includes the last residues of phrase structure grammar, including the notion of projection or later "labeling," the latter perhaps eliminable in terms of minimal search. Also eliminable on principled grounds are underlying and surface structure, and also logical form, in its technical sense, leaving just the interface levels (and their existence too is not graven in stone, a separate topic). The several compositional cycles that have commonly been postulated can be reduced to one, with periodic transfer of generated structures to the interface at a few designated positions ("phases"), yielding further consequences. A very elementary form of transformational grammar essentially "comes free;" it would require stipulations to block it, so that there is a principled explanation, in these terms, for the curious but ubiquitous phenomenon of displacement in natural language, with interpretive options in positions that are phonetically silent. And by the same token, any other approach to the phenomenon carries an empirical burden. Some of the island conditions have principled explanations, as does the existence of categories for which there is no direct surface evidence, such as a functional category of inflection.

Without proceeding, it seems to me no longer absurd to speculate that there may be a single internal language, efficiently yielding the infinite array of

expressions that provide a language of thought. Variety and complexity of language would then be reduced to the lexicon, which is also the locus of parametric variation, and to the ancillary mappings involved in externalization, which might turn out to be best possible solutions to relating organs with independent origins and properties. There are huge promissory notes left to pay, and alternatives that merit careful consideration, but plausible reduction of the previously assumed richness of UG has been substantial.

With each step towards the goals of principled explanation we gain a clearer grasp of the essential nature of language, and of what remains to be explained in other terms. 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 beyond that lies the mysterious problem of the creative and coherent ordinary use of language, a central problem of Cartesian science, still scarcely even at the horizons of inquiry.

References

- Appel, Toby A. 1987. *The Cuvier–Geoffroy Debate: French Biology in the Decades before Darwin* (Monographs on the History and Philosophy of Biology). New York: Oxford University Press.
- Baker, Mark. 2001. *The Atoms of Language: The Mind’s Hidden Rules of Grammar*. New York: Basic Books.
- Carroll, Sean B. 2005. *Endless Forms Most Beautiful: The New Science of Evo Devo and the Making of the Animal Kingdom*. New York: W.W. Norton & Co.
- Chomsky, Noam. 1955. The logical structure of linguistic theory. Ms., Harvard University/Massachusetts Institute of Technology. [Published in part as *The Logical Structure of Linguistic Theory*, New York: Plenum, 1975.]
- Chomsky, Noam. 1957. *Syntactic Structures* (Januar Linguarum, Series Minor 4). The Hague: Mouton.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1980. *Rules and Representations* (Woodbridge Lectures 11). New York: Columbia University Press.
- Chomsky, Noam. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Dordrecht: Foris.
- Chomsky, Noam. 1998. Comments: Galen Strawson, “Mental Reality”. *Philosophy and Phenomenological Research* 58(2), 437-441.
- Chomsky, Noam. 2004. The biolinguistic perspective after 50 years. *Quaderni del Dipartimento di Linguistica, Università di Firenze* 14, 3-12. [Also published in *Sito Web dell’ Accademia della Crusca*, 2005. http://www.accademiadella-crusca.it/img_usr/Chomsky%20-%20ENG.pdf (19 November 2007).]
- Darwin, Charles. 1892. *The Origin of Species*. London: John Murray.

- Gates, R. Ruggles. 1916. Huxley as mutationist. *The American Naturalist* 50(592), 126-128.
- Harris, Zellig S. 1951. *Methods of Structural Linguistics*. Chicago: University of Chicago Press.
- Hauser, Marc D. 2006. *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York: Ecco.
- Hinzen, Wolfram. 2006. *Mind Design and Minimal Syntax*. Oxford: Oxford University Press.
- Hinzen, Wolfram. 2007. *An Essay on Names and Truth*. Oxford: Oxford University Press.
- Huxley, Thomas. 1893. *Darwiniana: Essays*. London: Macmillan.
- Jacob, François. 1977. Evolution and tinkering. *Science* 196(1), 161-171.
- Jenkins, Lyle. 2000. *Biolinguistics: Exploring the Biology of Language*. Cambridge: Cambridge University Press.
- Lange, Friedrich A. 1892. *History of Materialism and Criticism of Its Present Importance*, vol. 3 [transl. by Ernest Chester Thomas]. London: Kegan Paul, Trench, Trübner & Co. [3rd edn., with an introduction by Bertrand Russell, New York: The Humanities Press, reprinted 1950.]
- Lashley, Karl S. 1951. The problem of serial order in behavior. In Lloyd A. Jeffress (ed.), *Cerebral Mechanisms in Behavior*, 112-136. New York: John Wiley.
- Leiber, Justin. 2001. Turing and the fragility and insubstantiality of evolutionary explanations: A puzzle about the unity of Alan Turing's work with some larger implications. *Philosophical Philosophy* 14(1), 83-94.
- Lewontin, Richard C. 1998. The evolution of cognition: Questions we will never answer. In Don Scarborough & Saul Sternberg (eds.), *Methods, Models, and Conceptual Issues (An Invitation to Cognitive Science 4, 2nd edn.)*, 107-134. Cambridge, MA: MIT Press.
- Miller, George A. & Philip N. Johnson-Laird. 1976. *Language and Perception*. Cambridge: Cambridge University Press.
- Mountcastle, Vernon B. 1998. Brain science at the century's ebb. *Daedalus* 127(2), 1-36.
- Piattelli-Palmarini, Massimo. 1974. A debate on bio-linguistics. *Centre Royaumont pour une science de l'homme report*. [Conference held at Endicott House, Dedham, Massachusetts, 20-21 May, 1974.]
- Piattelli-Palmarini, Massimo (ed.). 1980. *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*. London: Routledge & Kegan Paul.
- Piattelli-Palmarini, Massimo, Juan Uriagereka & Pello Salaburu (eds.). In press. *Of Minds and Language: The Basque Country Encounter with Noam Chomsky*. New York: Oxford University Press.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Rizzi, Luigi. 2003. Some elements of the study of language as a cognitive capacity. In Nicola Dimitri (ed.), *Cognitive Processes and Economic Behavior (Routledge Siena Studies in Political Economy)*. London: Routledge.
- Russell, Bertrand. 2003. *Russell on Metaphysics: Selections from the Writings of Bertrand Russell* [ed. by Stephen Mumford]. London: Routledge.
- Quine, W.V.O. 1960. *Word and Object (Studies in Communication)*. Cambridge, MA: MIT Press.

- Shannon, Claude E. & Warren Weaver. 1949/1998. *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.
- Shipley, Elizabeth P., Carlota S. Smith & Lila R. Gleitman. 1969. A study in the acquisition of language: Free responses to commands. *Language* 45 (2), 322-42.
- Skinner, B.F. 1957. *Verbal Behavior*. New York: Appleton–Century–Crofts.
- Striedter, George F. 2006. Précis of *Principles of Brain Evolution*. *Behavioral and Brain Sciences* 29(1), 1-12.
- Thompson, D'Arcy Wentworth. 1917/1992. *On Growth and Form* [abridged edn., prepared by John Tyler Bonner]. Cambridge: Cambridge University Press.
- Trubetzkoy, Nicolaj S. 1936. *Anleitung zu phonologischen Beschreibungen*. Brno: Edition du Cercle Linguistique de Prague.
- Trubetzkoy, Nicolaj S. 2001. *Studies in General Linguistics and Language Structure* [ed. by Anatoly Liberman, transl. by Marvin Taylor & Anatoly Liberman]. Durham, NC: Duke University Press.
- Turing, Alan M. 1952. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London* 237(B 641), 37-72. [Reprinted in Turing, Alan M. 1992. *Morphogenesis* [ed. by Peter T. Saunders]. Amsterdam: Elsevier.]
- Weinreich, Daniel M., Nigel F. Delaney, Mark A. DePristo & Daniel L. Hartl. 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312(5770), 111-114.

Noam Chomsky
Massachusetts Institute of Technology
Department of Linguistics and Philosophy
Mass. Avenue 32–D808
Cambridge, MA 02139
USA
chomsky@mit.edu