

Part I

Evolution

OUP UNCORRECTED PROOF – FIRST PROOF, 10/5/2010, SPi

2

The Bilingual Program: The Current State of its Development

ROBERT C. BERWICK AND NOAM CHOMSKY

Before discussing language, particularly in a biological context, we should be clear about what we mean by the term, which has engendered much confusion. Sometimes the term “language” is used to refer to human language; sometimes it is used to refer to any symbolic system or mode of communication or representation, as when one speaks of the language of the bees, or programming languages, or the language of the stars, and so on. Here we will keep to the first sense: human language, a particular object of the biological world. The study of language, so understood, has come to be called the *biolinguistic* perspective.

Among the many puzzling questions about language, two are salient: First, why are there any languages at all, evidently unique to the human lineage, what evolutionary biologists call an “autapomorphy”? Second, why are there so many languages? These are in fact the basic questions of origin and variation that so occupied Darwin and other evolutionary thinkers and comprise modern biology’s explanatory core: why do we observe *this* particular array of living forms in the world and not others? From this standpoint, linguistic science stands squarely within the modern biological tradition, despite its seemingly abstract details, as has often been observed.

According to a fairly general consensus among paleoanthropologists and archaeologists, these questions are very recent ones in evolutionary time. Roughly 100,000 years ago, the first question did not arise, because there were no languages. About 50,000 years ago, the answers to both questions were settled: our ancestors began their trek from Africa, spreading over the entire world, and as far as is known, the language faculty has remained essentially unchanged—which is not surprising in such a brief period. An infant from a Stone Age tribe in the Amazon, if brought to Boston, will be indistinguishable in linguistic and other cognitive functions from children born in Boston who

trace their ancestry to the first English colonists; and conversely. The actual dates are uncertain, and do not matter much for our purposes. The general picture appears to be roughly accurate.

We are therefore concerned with a curious biological object, *language*, which has appeared on earth quite recently. It is a species property of humans, a common endowment with no significant variation apart from serious pathology, unlike anything else known in the organic world in its essentials, and surely central to human life since its emergence. It is a central component of what the co-founder of modern evolutionary theory, Alfred Russell Wallace, called “man’s intellectual and moral nature:” the human capacities for creative imagination, language and symbolism generally, recording and interpretation of natural phenomena, intricate social practices and the like, a complex that is sometimes simply called “the human capacity.” This complex seems to have crystallized fairly recently among a small group in East Africa of whom we are all descendants, distinguishing contemporary humans sharply from other animals, with enormous consequences for the whole of the biological world. It is commonly and plausibly assumed that the emergence of language was a core element in this sudden and dramatic transformation. Furthermore, language is one component of the human capacity that is accessible to study in some depth. That is another reason why even research that is purely linguistic in character actually falls under the heading of biolinguistics despite its superficial remove from biology, as exemplified in the chapters by Lasnik and Larson in this volume.

From the biolinguistic perspective, we can think of language as, in essence, an “organ of the body,” more or less on a par with the visual or digestive or immune systems. Like others, it is a subcomponent of a complex organism that has sufficient internal integrity so that it makes sense to study it in abstraction from its complex interactions with other systems in the life of the organism. In this case it is 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,” in the words of the eighteenth-century scientist and philosopher Joseph Priestley. He was articulating the natural conclusion after Newton had demonstrated, to Newton’s own great dismay and disbelief, that the world is not a machine, contrary to the core assumptions of the seventeenth-century scientific revolution—a conclusion that in effect eliminated the traditional mind–body problem, because there is no longer a coherent concept of *body* (*matter, physical*), a matter well understood in the eighteenth and nineteenth centuries. We can think of language as a *mental organ*, where the term “mental” simply refers to certain aspects of the world, to be studied in the same way

as chemical, optical, electrical, and other aspects, with the hope for eventual unification—noting that such unification in these other domains in the past was often achieved in completely unexpected ways, not necessarily by reduction.

As mentioned at the outset with regard to the curious mental organ *language*, two obvious questions arise. One is: Why does it exist at all, evidently unique to our species? Second: Why is there more than one language? In fact, why is there such a multitude and variety that languages appear to “differ from each other without limit and in unpredictable ways” and therefore the study of each language must be approached “without any preexistent scheme of what a language must be,” here quoting the formulation of the prominent theoretical linguist Martin Joos 50 years ago, summarizing the reigning “Boasian tradition,” as he plausibly called it, tracing it to the work of one of the founders of modern anthropology and anthropological linguistics, Franz Boas. The publication that was the foundation of American structural linguistics in the 1950s, Zellig Harris’s *Methods in Structural Linguistics* (1951), was called “methods” because there seemed to be little to say about language beyond the methods for reducing the data from limitlessly varying languages to organized form. European structuralism was much the same. Nikolai Trubetzkoy’s classic introduction to phonological analysis was similar in conception. More generally, structuralist inquiries focused almost entirely on phonology and morphology, the areas in which languages do appear to differ widely and in complex ways, a matter of broader interest, to which we will return.¹

The dominant picture in general biology at about the same time was rather similar, captured in molecular biologist Gunther Stent’s 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” (as quoted in (123) 2005a: 24).

In fact the problem of reconciling unity and diversity has constantly arisen in general biology as well as in linguistics. The study of language that developed within the seventeenth century scientific revolution distinguished universal from particular grammar, though not quite in the sense of the contemporary biolinguistic approach. Universal grammar was taken to be the intellectual core of the discipline; particular grammars were regarded as accidental instantiations of the universal system. With the flourishing of anthropological linguistics, the pendulum swung in the other direction, towards diversity, well articulated in the Boasian formulation we quoted. In general

¹ See Joos (1957); Trubetzkoy (1939, trans. 1969).

biology, the issue had been raised sharply in a famous debate between the naturalists Georges Cuvier and Geoffroy St. Hilaire in 1830. Cuvier's position, emphasizing diversity, prevailed, particularly after the Darwinian revolution, leading to the conclusions about the near infinitude of variety that have to be sorted out case by case. 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." It is unclear if the irony was intended, but these words were taken by evolutionary biologist Sean Carroll 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.

Reconciliation of the apparent diversity of organic forms with their evident underlying uniformity—why do we see *this* array of living things in the world and not others, just as why do we see *this* array of languages/grammars and not others?—comes about through the interplay of three factors, famously articulated by the biologist Monod in his book *Le Hazard et la Nécessité: (1970; Chance and Necessity, 1972)*. First, there is the historically contingent fact that we are all common descendants from a single tree of life, and so share common ancestry with all other living things, which apparently have explored only a minute fraction of a space that includes a much larger set of possible biological outcomes. It should by now be no surprise that we therefore possess common genes, biochemical pathways, and much else.

Second, there are the physio-chemical constraints of the world, necessities that delimit biological possibilities, like the near-impossibility of wheels for locomotion due to the physical difficulty of providing a nerve control and a blood supply to a rotating object. Third, there is the sieving effect of natural selection, which winnows out from a pre-existing menu of possibilities—offered by historical contingency and physio-chemical constraints—the actual array of organisms that we observe in the world around us. Note that the effect of the constrained menu of options is of utmost importance; if the options are extremely constrained, then selection would have very little to choose from: it should be no surprise that when one goes to a fast-food restaurant one is usually seen leaving with a hamburger and french fries. Thus, just as Darwin would have it, natural selection is by no means the "exclusive" means that has shaped the natural world: "Furthermore, I am convinced that Natural Selection has been the main but not exclusive means of modification" (Darwin 1859: 7).

Recent discoveries have reinvigorated the general approach of D'Arcy Thompson (1992) and Alan Turing on principles that constrain the variety of organisms. In Turing and Wardlaw's words, the true science of biology should

regard each “living organism as a special kind of system to which the general laws of physics and chemistry apply,” sharply constraining their possible variety and fixing their fundamental properties (Turing and Wardlaw 1953). That perspective may sound less extreme today after the discovery of master genes, deep homologies and conservation, 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 by Weinrich et al. (2006) on feasible mutational paths, reinterpreting a famous image of Steven Gould’s, who had suggested that the tape of life, if replayed, might follow a variety of paths. As Michael Lynch further notes (2007: 367), “we have known for decades that all eukaryotes share most of the same genes for transcription, translation, replication, nutrient uptake, core metabolism, cytoskeletal structure, and so forth. Why would we expect anything different for development?”

In a recent review of the evo–devo approach, Gerd Müller (2007) notes how much more concrete our understanding of the Turing-type patterning models have become, observing that several

Generic forms . . . result from the interaction of basic cell properties with different pattern-forming mechanisms. Differential adhesion and cell polarity when modulated by different kinds of physical and chemical patterning mechanisms . . . lead to standard organizational motifs . . . differential adhesion properties and their polar distribution on cell surfaces lead to hollow spheres when combined with a diffusion gradient, and to invaginated spheres when combined with a sedimentation gradient. . . . The combination of differential adhesion with a reaction-diffusion mechanism generates radially periodic structures, whereas a combination with chemical oscillation results in serially periodic structures. Early metazoan body plans represent an exploitation of such generic patterning repertoires. (Müller 2007: 947)

For example, the contingent fact that we have five fingers and five toes may be better explained by an appeal to how toes and fingers develop than that five is optimal for their function.²

Biochemist Michael Sherman (2007) argues, perhaps more controversially, that a “Universal Genome that encodes all major developmental programs essential for various phyla of Metazoa emerged in a unicellular or a primitive multicellular organism shortly before the Cambrian period” about 500 million years ago, when there was a sudden explosion of complex animal forms;

² As Ahouse and Berwick (1998) note, “Five fingers and toes were not the original number of digits in tetrapods (see the discussion by M. I. Coates and J. A. Clark in *Nature* 347, 1990, 66–9) and amphibians probably never had more than four digits (and generally have three) on their front and back feet. There is a clever explanation from molecular developmental genetics that rationalizes why there are at most five different types of digits even if some are duplicated.”

and, further, that the many “Metazoan phyla, all having similar genomes, are nonetheless so distinct because they utilize specific combinations of developmental programs.” On this view, there is but one multicellular animal from a sufficiently abstract point of view—the point of view that might be taken by a Martian scientist from a much more advanced civilization viewing events on earth. Superficial variety would result in part from various arrangements of an evolutionarily conserved “developmental–genetic toolkit,” as it is sometimes called. If ideas of this kind prove to be on the right track, the problem of unity and diversity will be reformulated in ways that would have surprised some recent generations of scientists. The degree to which the conserved toolkit is the sole explanation for the observed uniformity deserves some care. As mentioned, observed uniformity arises in part because there has simply not been enough time, and contingent ancestry-by-descent bars the possibility of exploring “too much” of the genetic–protein–morphological space—particularly given the virtual impossibility of “going backwards” and starting the search over again for greater success. Given these inherent constraints, it becomes much less of a surprise that organisms are all built according to a certain set of *Baupläne*, as Steven Gould has emphasized, among others. It is in this sense that if sophisticated Martian scientists came to earth, they would probably see in effect just one organism, though with many apparent superficial variations.

The uniformity had not passed unnoticed in Darwin’s day. The naturalistic studies of Darwin’s close associate and expositor Thomas Huxley led him to observe, with some puzzlement, 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 (Maynard-Smith *et al.* 1985: 266). Indeed, the study of the sources and nature of possible variation constituted a large portion of Darwin’s own research program after *Origin*, as summarized in his *Variation of Plants and Animals under Domestication* (1868). Huxley’s conclusion is reminiscent of earlier ideas of “rational morphology”, a famous example being Goethe’s theories of archetypal forms of plants, which have been partially revived in the evo–devo revolution. Indeed, as indicated earlier, Darwin himself was sensitive to this issue, and, grand synthesizer that he was, dealt more carefully with such “laws of growth and form”: the constraints and opportunities to change are due to the details of development, chance associations with other features that may be strongly selected for or against, and finally selection on the trait itself. Darwin noted that such laws of “correlation and balance” would be of considerable importance to his theory, remarking, for example, that “white cats if they have blue eyes are almost invariably deaf” (Darwin, 1856 letter to W. D. Fox).

When the evolutionary Modern Synthesis, pioneered by Fisher, Haldane, and Wright, held sway through most of the last half of the previous century, emphasis in evolution was focused on micro-mutational events and gradualism, singling out the power of natural selection operating via very small incremental steps. More recently, however, in general biology the pendulum has been swinging towards a combination of Monod's three factors, yielding new ways of understanding traditional ideas.

Let us return to the first of the two basic questions: Why should there be any languages at all, apparently an autapomorphy? As mentioned, very recently in evolutionary time the question would not have arisen: there were no languages. There were, of course, plenty of animal communication systems. But they are all radically different from human language in structure and function. Human language does not even fit within the standard typologies of animal communication systems—Marc Hauser's, for example, in his comprehensive review of the evolution of communication (1997). It has been conventional to regard language as a system whose function is communication. This is indeed the widespread view invoked in most selectionist accounts of language, which almost invariably start from this interpretation. However, to the extent that the characterization has any meaning, this appears to be incorrect, for a variety of reasons to which we turn below.

The inference of a biological trait's "purpose" or "function" from its surface form is always rife with difficulties. Lewontin's remarks in *The Triple Helix* (2001: 79) illustrate how difficult it can be to assign a unique function to an organ or trait even in the case of what at first seems like a far simpler situation: bones do not have a single, unambiguous function. While it is true that bones support the body, allowing us to stand up and walk, they are also a storehouse for calcium and bone marrow for producing new red blood cells, so they are in a sense part of the circulatory system.

What is true for bones is also true in spades for human language. Moreover, there has always been an alternative tradition, expressed by Burling (1993), among others, that humans may well possess a secondary communication system like those of other primates, namely a nonverbal system of gestures or even calls, but that this is not language, since, as Burling notes, "our surviving primate communication system remains sharply distinct from language."³

Language can of course be used for communication, as can any aspect of what we do: Style of dress, gesture, and so on. And it can be and commonly is used for much else. Statistically speaking, for whatever that is worth, the

³ Laura Petitto's work on the acquisition of sign language (1987) demonstrates Burling's point rather dramatically—the same gesture is used for pointing and pronominal reference, but in the latter case the gesture is counter-iconic at the age when infants typically reverse "I" and "you."

overwhelming use of language is internal—for thought. It takes an enormous act of will to keep from talking to oneself in every waking moment—and asleep as well, often a considerable annoyance. The distinguished neurologist Harry Jerison (1973: 55) among others expressed a stronger view, holding that “language did not evolve as a communication system . . . the initial evolution of language is more likely to have been . . . for the construction of a real world,” as a “tool for thought.” Not only in the functional dimension, but also in all other respects—semantic, syntactic, morphological, and phonological—the core properties of human language appear to differ sharply from animal communication systems, and to be largely unique in the organic world.

How, then, did this strange object appear in the biological record, apparently within a very narrow evolutionary window, perhaps about 50–100,000 years ago? There are of course no definite answers, but it is possible to sketch what seem to be some reasonable speculations, which relate closely to work of recent years in the biolinguistic framework.

Anatomically modern humans are found in the fossil record several hundred thousand years ago, but evidence of the human capacity is much more recent, not long before the trek from Africa. Paleoanthropologist Ian Tattersall reports that “a vocal tract capable of producing the sounds of articulate speech” existed over half a million years before there is any evidence that our ancestors were using language. “We have to conclude,” he writes, “that the appearance of language and its anatomical correlates was not driven by natural selection, however beneficial these innovations may appear in hindsight”—a conclusion which raises no problems for standard evolutionary biology, contrary to illusions in popular literature (Tattersall 1998). It appears that human brain size reached its current level recently, perhaps 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” (neuroscientist Georg Striedter 2004). With regard to language, Tattersall concludes that “after a long—and poorly understood—period of erratic brain expansion and reorganization in the human lineage, something occurred that set the stage for language acquisition. This innovation would have depended on the phenomenon of emergence, whereby a chance combination of preexisting elements results in something totally unexpected,” presumably “a neural change . . . in some population of the human lineage . . . rather minor in genetic terms, [which] probably had nothing whatever to do with adaptation” though it conferred advantages, and then proliferated. Perhaps it was an automatic consequence of absolute brain size, as Striedter suggests, or perhaps some minor chance mutation. Some time later—not very long in evolutionary time—came further innovations,

perhaps culturally driven, that led to behaviorally modern humans, the crystallization of the human capacity, and the trek from Africa (Tattersall 1998, 2002, 2005).

What was that neural change in some small group that was rather minor in genetic terms? To answer that, we have to consider the special properties of language. The most elementary property of our shared language capacity is that it enables us to construct and interpret a discrete infinity of hierarchically structured expressions: discrete because there are five-word sentences and six-word sentences, but no five-and-a-half-word sentences; infinite because there is no longest sentence. Language is therefore based on a recursive generative procedure that takes elementary word-like elements from some store, call it the lexicon, and applies repeatedly to yield structured expressions, without bound. To account for the emergence of the language faculty—hence for the existence of at least one language—we have to face two basic tasks. One task is to account for the “atoms of computation,” the lexical items—commonly in the range of 30–50,000. The second is to discover the computational properties of the language faculty. This task in turn has several facets: we must seek to discover the generative procedure that constructs infinitely many expressions in the mind, and the methods by which these internal mental objects are related to two *interfaces* with language-external (but organism-internal) systems: the system of thought, on the one hand, and also to the sensory-motor system, thus *externalizing* internal computations and thought. This is one way of reformulating the traditional conception, at least back to Aristotle, that language is sound with a meaning. All of these tasks pose very serious problems, far more so than was believed in the recent past, or often today.

Let us turn then to the basic elements of language, beginning with the generative procedure, which, it seems, emerged some time in the 50,000–100,000 year range, barely a flick of an eye in evolutionary time, presumably involving some slight rewiring of the brain. At this point the evo–devo revolution in biology becomes relevant. It has provided compelling evidence for two relevant conclusions. One is that genetic endowment even for regulatory systems is deeply conserved. A second is that very slight changes can yield great differences in observed outcome—though phenotypic variation is nonetheless limited, in virtue of the deep conservation of genetic systems, and laws of nature of the kind that interested Thompson and Turing. To cite a simple illustration, there are two kinds of stickleback fish, with or without spiky spines on the pelvis. About 10,000 years ago, a mutation in a genetic “switch” near a gene involved in spine production differentiated the two varieties, one with spines and one without, one adapted to oceans and the other to lakes (Colosimo et al. 2004, 2005; Orr 2005).

Much more far-reaching results have to do with the evolution of eyes, an intensively studied topic. It turns out that there are very few types of eye, in part because of constraints imposed by the physics of light, in part because only one category of proteins, opsin molecules, can perform the necessary functions. The genes encoding opsin had very early origins, and are repeatedly recruited, but only in limited ways, again because of physical constraints. The same is true of eye lens proteins. The evolution of eyes illustrates the complex interactions of physical law, stochastic processes, and the role of selection in choosing within a narrow physical channel of possibilities (Gehring 2005).

Jacob and Monod's work from 1961 on the discovery of the "operon" in *E. coli* for which they won the Nobel Prize, led to Monod's famous quote (cited in Jacob 1988): "what is true for the colon bacillus [*E. coli*] is true for the elephant" (Jacob 1988: 290). While this has sometimes been interpreted as anticipating the modern evo–devo account, it seems that what Monod actually meant was that his and François Jacob's generalized negative regulation theory should be sufficient to account for all cases of gene regulation. But this conjecture turned out to be incorrect. Eukaryotes (non-bacteria) do not (cannot) use the operon regulatory machinery that bacteria (Prokaryotes) use, because they do not have their genes lined up neatly in a linear fashion, strung out like beads along a string, without breaks and without intervening non-protein coding regions (introns). Roughly, it is this arrangement in Prokaryotes that permits the negative feedback operon system to work. Indeed, much of the modern evo-devo revolution is about the discovery of the rather more sophisticated methods for gene regulation and development employed by Eukaryotes. Nonetheless, Monod's basic notion that slight differences in timing and arrangement of regulatory mechanisms that activate genes could result in enormous differences did turn out to be correct, though the machinery was unanticipated. It was left to Jacob (1977) to provide a suggestive model for the development of other organisms based on the notion that "thanks to complex regulatory circuits" what "accounts for the difference between a butterfly and a lion, a chicken and a fly... are the result of mutations which altered the organism's regulatory circuits more than its chemical structure" (1977: 26). Jacob's model in turn provided part of the inspiration for the Principles and Parameters (P&P) approach to language, a matter discussed in lectures shortly after (Chomsky 1980: 67).

The P&P approach is based on the assumption that languages consist of fixed and invariant principles connected to a kind of switchbox of parameters, questions that the child has to answer on the basis of presented data in order to fix a language from the limited variety available in principle—or perhaps,

as Charles Yang has argued (2002), to determine a probability distribution over languages resulting from a learning procedure for parameter setting. For example, the child has to determine whether the language to which it is exposed is head-initial, like English, a language in which substantive elements precede their objects, as in *read the book* or *in the room*; or whether it is head-final, like Japanese, where the counterparts would be *book read* and *room in*. As in the somewhat analogous case of rearrangement of regulatory mechanisms, the approach suggests a framework for understanding how essential unity might yield the appearance of the limitless diversity that was assumed not long ago for language (as for biological organisms generally).

The P&P research program has been very fruitful, yielding rich new understanding of a very broad typological range of languages, opening new questions that had never been considered, sometimes providing answers. It is no exaggeration to say that more has been learned about languages in the past twenty-five years than in the earlier millennia of serious inquiry into language. With regard to the two salient questions with which we began, the approach suggests that what emerged, fairly suddenly in evolutionary terms, was the generative procedure that provides the principles, and that diversity of language results from the fact that the principles do not determine the answers to all questions about language, but leave some questions as open parameters. Notice that the single illustration above has to do with ordering. Though the matter is contested, it seems that there is by now substantial linguistic evidence that ordering is restricted to externalization of internal computation to the sensory-motor system, and plays no role in core syntax and semantics, a conclusion for which there is also accumulating biological evidence of a sort familiar to mainstream biologists, to which we return below.

The simplest assumption, hence the one we adopt unless counter-evidence appears, is that the generative procedure emerged suddenly as the result of a minor mutation. In that case we would expect the generative procedure to be very simple. Various kinds of generative procedures have been explored in the past 50 years. One approach familiar to linguists and computer scientists is phrase-structure grammar, developed in the 1950s and since extensively employed. The approach made sense at the time. It fit very naturally into one of the several equivalent formulations of the mathematical theory of recursive procedures—Emil Post’s rewriting systems—and it captured at least some basic properties of language, such as hierarchic structure and embedding. Nevertheless, it was quickly recognized that phrase-structure grammar is not only inadequate for language but is also quite a complex procedure with many arbitrary stipulations, not the kind of system we would hope to find, and unlikely to have emerged suddenly.

Over the years, research has found ways to reduce the complexities of these systems, and finally to eliminate them entirely in favor of the simplest possible mode of recursive generation: an operation that takes two objects already constructed, call them X and Y , and forms from them a new object that consists of the two unchanged, hence simply the set with X and Y as members. Call this operation Merge. Provided with conceptual atoms of the lexicon, the operation Merge, iterated without bound, yields an infinity of hierarchically constructed expressions. If these can be interpreted by conceptual systems, the operation provides an internal language of thought.

A very strong thesis, called the “strong Minimalist thesis,” is that the generative process is optimal: the principles of language are determined by efficient computation and language keeps to the simplest recursive operation, Merge, designed to satisfy interface conditions in accord with independent principles of efficient computation. Language is something like a snowflake, assuming its particular form by virtue of laws of nature—in this case principles of computational efficiency—once the basic mode of construction is available, and satisfying whatever conditions are imposed at the interfaces. The basic thesis is expressed in the title of a recent collection of technical essays: “Interfaces + Recursion = Language?” (Sauerl and Gärtner 2007). Optimally, recursion can be reduced to Merge. The question mark in the title is of course highly appropriate: the questions arise at the border of current research. We will suggest below that there is a significant asymmetry between the two interfaces, with the semantic–pragmatic interface—the link to systems of thought and action—having primacy. Just how rich these external conditions may be is also a serious research question, and a hard one, given the lack of much evidence about these thought–action systems that is independent of language. A very strong thesis, suggested by Wolfram Hinzen (2006) is that central components of thought, such as propositions, are basically derived from the optimally constructed generative procedure. If such ideas can be sharpened and validated, then the effect of the semantic–pragmatic interface on language design would be correspondingly reduced.

The strong minimalist thesis is very far from established, but it looks much more plausible than it did only a few years ago. Insofar as it is correct, the evolution of language will reduce to the emergence of the operation Merge, the evolution of conceptual atoms of the lexicon, the linkage to conceptual systems, and the mode of externalization. Any residue of principles of language not reducible to Merge and optimal computation will have to be accounted for by some other evolutionary process—one that we are unlikely to learn much about, at least by presently understood methods, as Lewontin (1998) notes.

Notice that there is no room in this picture for any precursors to language—say a language-like system with only short sentences. There is no rationale for postulation of such a system: to go from seven-word sentences to the discrete infinity of human language requires emergence of the same recursive procedure as to go from zero to infinity, and there is of course no direct evidence for such protolanguages. Similar observations hold for language acquisition, despite appearances, a matter that we put to the side here.

Crucially, the operation Merge yields the familiar *displacement* property of language: the fact that we pronounce phrases in one position, but interpret them somewhere else as well. Thus in the sentence *Guess what John is eating*, we understand *what* to be the object of *eat*, as in *John is eating an apple*, even though it is pronounced somewhere else. This property has always seemed paradoxical, a kind of *imperfection* of language. It is by no means necessary in order to capture semantic facts, but it is ubiquitous. It surpasses the capacity of phrase structure grammars, requiring that they be still further complicated with additional devices. But it falls within Merge, automatically. To see how, suppose that the operation Merge has constructed the mental expression corresponding to *John is eating what*. A larger expression can be constructed by Merge in two ways: Internal Merge can add something from within the expression, so as to form *what John is eating what*; and External Merge can add something new, yielding *Guess what John is eating what*.

That carries us part of the way towards displacement. In *what John is eating what*, the phrase *what* appears in two positions, and in fact those two positions are required for semantic interpretation: the original position provides the information that *what* is understood to be the direct object of *eat*, and the new position, at the edge, is interpreted as a quantifier ranging over a variable, so that the expression means something like “for which thing *x*, John is eating the thing *x*.”

These observations generalize to a wide range of constructions. The results are just what is needed for semantic interpretation, but they do not yield the objects that are pronounced in English. We do not pronounce *guess what John is eating what*, but rather *guess what John is eating*, with the original position suppressed. That is a universal property of displacement, with minor (and interesting) qualifications that we can ignore here. The property follows from elementary principles of computational efficiency. In fact, it has often been noted that serial motor activity is computationally costly, a matter attested by the sheer quantity of motor cortex devoted to both motor control of the hands and for oro-facial articulatory gestures.

To externalize the internally generated expression *what John is eating what*, it would be necessary to pronounce *what* twice, and that turns out to place

a very considerable burden on computation, when we consider expressions of normal complexity and the actual nature of displacement by Internal Merge. With all but one of the occurrences of *what* suppressed, the computational burden is greatly eased. The one occurrence that must be pronounced is the most prominent one, the last one created by Internal Merge: otherwise there will be no indication that the operation has applied to yield the correct interpretation. It appears, then, that the language faculty recruits a general principle of computational efficiency for the process of externalization.

The suppression of all but one of the occurrences of the displaced element is computationally efficient, but imposes a significant burden on interpretation, hence on communication. The person hearing the sentence has to discover the position of the gap where the displaced element is to be interpreted. That is a highly non-trivial problem in general, familiar from parsing programs. There is, then, a conflict between computational efficiency and interpretive-communicative efficiency. Universally, languages resolve the conflict in favor of computational efficiency. These facts at once suggest that language evolved as an instrument of internal thought, with externalization a secondary process. There is a great deal of evidence from language design that yields similar conclusions; so called “island properties,” for example.

There are independent reasons for the conclusion that externalization is a secondary process. One is that externalization appears to be modality-independent, as has been learned from studies of sign language in recent years. The structural properties of sign and spoken language are remarkably similar. Additionally, acquisition follows the same course in both, and neural localization seems to be similar as well. That tends to reinforce the conclusion that language is optimized for the system of thought, with mode of externalization secondary.

Note further that the constraints on externalization holding for the auditory modality also appear to hold in the case of the visual modality in signed languages. Even though there is no physical constraint barring one from “saying” with one hand that *John likes ice-cream* and *Mary likes beer* with the other hand, nevertheless it appears that one hand is dominant throughout and delivers sentences (via gestures) in a left-to-right order in time, linearized as in vocal-tract externalization, while the non-dominant hand adds markings for emphasis, morphology, and the like.

Indeed, it seems possible to make a far stronger statement: all recent relevant biological and evolutionary research leads to the conclusion that the process of externalization is secondary. This includes the recent and highly publicized discoveries of genetic elements putatively involved in language,

specifically, the FOXP2 regulatory (transcription factor) gene. FOXP2 is implicated in a highly heritable language defect, so-called verbal dyspraxia. Since this discovery it has been intensely analyzed from an evolutionary and comparative standpoint, with small amino-acid differences between the human variant and other primates and non-human mammals posited as the target of recent positive natural selection, perhaps concomitant with language emergence (Fisher et al. 1998; Enard et al. 2005); and with similarities between those same two amino acids in humans and Neandertals also suggested as possibly significant with respect to language (Krause, Lalueza et al. 2007; *Science Daily*, 21 Oct. 2007).

However, we might ask whether this gene is centrally involved in language or, as now seems more plausible, is part of the secondary externalization process. Recent discoveries in birds and mice over the past few years point to an “emerging consensus” that this transcription factor gene is not so much part of a blueprint for internal syntax, the narrow faculty of language, and most certainly not some hypothetical “language gene” (just as there are no single genes for eye color or autism) but rather part of regulatory machinery related to externalization (Vargha-Khadem et al. 2005; Groszer et al. 2008). FOXP2 aids in the development of serial fine motor control, orofacial or otherwise: the ability to literally put one sound or gesture down in place, one point after another in time.

In this respect it is worth noting that members of the KE family in which this genetic defect was originally isolated exhibit a quite general motor dyspraxia, not localized to simply their oro-facial movements. Recent studies where a mutated FOXP2 gene built to replicate the defects found in the KE family was inserted in mice confirm this view: “We find that *Foxp2*-R552H heterozygous mice display subtle but highly significant deficits in learning of rapid motor skills . . . These data are consistent with proposals that human speech faculties recruit evolutionarily ancient neural circuits involved in motor learning” (Groszer et al. 2008:359).

If this view is on the right track, then FOXP2 is more akin to the blueprint that aids in the construction of a properly functioning input–output system for a computer, like its printer, rather than the construction of the computer’s central processor itself. From this point of view, what has gone wrong in the affected KE family members is thus something awry with the externalization system, the “printer,” not the central language faculty itself. If this is so, then the evolutionary analyses suggesting that this transcription factor was under positive selection approximately 100,000 years ago (in itself arguable) could in fact be quite inconclusive about the evolution of the core components of the faculty of language, syntax, and the mapping to the

“semantic” (conceptual–intensional) interface. It is difficult to determine the causal sequence: the link between FOXP2 and high-grade serial motor coordination could be regarded as either an opportunistic prerequisite substrate for externalization, no matter what the modality, as is common in evolutionary scenarios, or the result of selection pressure for efficient externalization solutions after Merge arose. In either case, FOXP2 becomes part of a system extrinsic to core syntax/semantics.

There is further recent evidence from Michael Coen (p.c.) regarding serial coordination in vocalization suggesting that discretized serial motor control might simply be a substrate common to all mammals, and possibly all vertebrates. If so, then the entire FOXP2 story, and motor externalization generally, is even further removed from the picture of core syntax/semantics evolution. The evidence comes from the finding that all mammals tested (people, dogs, cats, seals, whales, baboons, tamarin monkeys, mice) and unrelated vertebrates (crows, finches, frogs, etc.) possess what was formerly attributed just to the human externalization system: each of the vocal repertoires of these various species is drawn from a *finite* set of distinctive phonemes (or, more accurately, songemes in the case of birds, barkemes in the case of dogs, etc.). Coen’s hypothesis is that each species has some finite number of articulatory productions, for example, phonemes, that are genetically constrained by its physiology, according to principles such as minimization of energy during vocalization, physical constraints, and the like. This is similar to Kenneth Stevens’s picture of the quantal nature of speech production (Stevens 1972, 1989).

On this view, any given species uses a subset of species-specific primitive sounds to generate the vocalizations common to that species. (It would not be expected that each animal uses all of them, in the same way that no human employs all phonemes.) If so, then our hypothetical Martian would conclude that even at the level of peripheral externalization, there is one human language, one dog language, one frog language, and the like.

Summarizing, FOXP2 does not speak to the question of the core faculty of human language because it really has nothing to do with the core language *phenotype*, Merge and syntax. From an explanatory point of view, this makes it quite unlike the case of, say, sickle-cell anemia where a genetic defect directly leads to the aberrant trait, the formation of an abnormal haemoglobin protein and resulting red blood cell distortion. To be sure, FOXP2 remains a possibly necessary component of the language system in the same way that a printer is part of a computer system. But it is not human language *tout court*. If all this is so, then the explanation “for” the core language phenotype

may be even more indirect and difficult than Richard Lewontin (1998) has sketched.⁴

In fact, in many respects this focus on FOXP2 and dyspraxia is quite similar to the near-universal focus on language as communication.⁵ Both efforts examine properties apparently particular only to the externalization process, which, we conjecture, is not part of the core faculty of human language. In this sense both efforts are misdirected, unrevealing of the internal computations of the mind/brain, the cognitive revolution notwithstanding. By expressly stating the distinction between internal syntax and externalization, many new research directions may be opened up, and new concrete, testable predictions posed particularly from a biological perspective, as the example of animal vocal productions illustrates.

Returning to the core principles of language, unbounded Merge (hence displacement) must have arisen from some rewiring of the brain, hence 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 partially transmitted to offspring, and because of the selective advantages it confers, it might come to dominate a small breeding group, though as with all such novel mutations, there is an issue about how an initially small number of copies of such an allele might survive, despite a large selective advantage. As first noted by Haldane (1927), the probability of even a highly advantageous heterozygous mutation with a selective advantage of 1 percent—about an order of magnitude greater than any selective advantage found in field measurement—would nevertheless have approximately a $1/e$ or about a 30 percent chance of going extinct within one generation. Gillespie (1991) notes that a 99 percent certainty of fixing such an advantageous mutation is attained only after reaching approximately 4000 copies of such an allele. Assuming an effective population size of about

⁴ Note that the argument still goes through if we suppose that there's another possibility: that FOXP2 builds part of the input–output system for vocal learning where one must externalize and then re-internalize song/language—sing or talk to oneself. This would remain a way to “pipe” items in and out of the internal system, and serialize them, possibly a critical component to be sure, in the same sense that one might require a way to print output from a computer.

⁵ This is much like attending solely to the different means by an LCD television and the old cathode-ray tube TVs display moving images without paying any attention to what image is being displayed. The old TVs “painted” a picture by sweeping an electron beam over a set of chemical dots that would glow or not. Liquid crystal displays operate by an entirely different means: roughly, they pass light or not through a liquid crystal array of dots depending on an electric charge applied to each dot, but there is no single sweeping beam. One generates the same flat image by an entirely different means. Similarly, whether the externalized, linear timing slots are being set out by motor commands to the vocal tract or by moving fingers is irrelevant to the more crucial “inner” representations.

this number—not unreasonable for what we understand about early demography in Africa at this time, albeit poorly—this would suggest that such a beneficial allele would have to spread to the entire breeding group in order that natural selection could operate unimpeded to sweep the mutation to fixation. This is not paradoxical, but simply reflects the stochastic character of evolution by natural selection itself; the same principle applies to all beneficial mutations. What it implies is that the emergence of language in this sense could indeed have been a unique event, accounting for its species-specific character. Such founder effects in population bottleneck situations are not uncommon.

When the beneficial mutation has spread through the group, 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 as a special case. It is not easy to imagine an account of human evolution that does not assume at least this much, in one or another form. Any additional assumption requires both evidence and rationale, not easy to come by.

Most alternatives do in fact posit additional assumptions, grounded on the language-as-communication viewpoint, presumably related to externalization as we have seen. In a recent survey Számado and Szathmáry (2006) list what they consider to be the major alternative theories explaining the emergence of human language. These include: (1) language as gossip; (2) language as social grooming; (3) language as outgrowth of hunting cooperation; (4) language as outcome of motherese; (5) sexual selection; (6) language as requirement of exchanging status information; (7) language as song; (8) language as requirement for tool making or the outcome of tool making; (9) language as an outgrowth of gestural systems; (10) language as Machiavellian device for deception; and finally, (11) language as internal mental tool.

Note that it is only this last theory, language as internal mental tool, that does not assume, explicitly or implicitly, that the primary function of language is for external communication. But this leads to a kind of adaptive paradox, since animal signaling ought to then suffice. Számado and Szathmáry note:

Most of the theories do not consider the kind of selective forces that could encourage the use of conventional communication in a given context instead of the use of ‘traditional’ animal signals . . . thus, there is no theory that convincingly demonstrates a situation that would require a complex means of symbolic communication rather than the existing simpler communication systems. (2006: 679)

They further note that the language-as-mental-tool theory does not suffer from this defect. However, they, like most researchers in this area, do not seem to draw the obvious inference but instead maintain a focus on externalization and communication.

Proposals as to the primacy of internal language—similar to Harry Jerison’s observation, already noted, about language as an “inner tool”—have also been made by eminent evolutionary biologists. At an international conference on biolinguistics in 1974, 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.” The same idea was taken up by François Jacob, who suggested that “the role of language as a communication system between individuals would have come about only secondarily . . . 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, 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.” These ideas trace back to the cognitive revolution of the seventeenth century, which in many ways foreshadows developments from the 1950s (Jacob 1982; Luria 1974).

We can, however, go beyond speculation. Investigation of language design can yield evidence on the relation of language to the sensory-motor system and thought systems. As noted, we think there is mounting evidence to support the natural conclusion that the relation is asymmetrical in the manner illustrated in the critical case of displacement.

Externalization is not a simple task. It has to relate two quite distinct systems: one is a sensorimotor system that appears to have been basically intact for hundreds of thousands of years; the second is a newly emerged computational system for thought, which is perfect insofar as the strong minimalist thesis is correct. We would expect, then, that morphology and phonology—the linguistic processes that convert internal syntactic objects to the entities accessible to the sensory-motor system—might turn out to be quite intricate, varied, and subject to accidental historical events. Parameterization and diversity, then, would be mostly—possibly entirely—restricted to externalization. That is pretty much what we seem to find: a computational system efficiently generating expressions interpretable at the semantic–pragmatic interface, with diversity resulting from complex and highly varied modes

of externalization which, furthermore, are readily susceptible to historical change.⁶

If this picture is more or less accurate, we may have an answer to the second of the two basic questions: Why are there so many languages? The reason might be that the problem of externalization can be solved in many different and independent ways, either before or after the dispersal of the original population.

We have no reason to suppose that solving the externalization problem involved an evolutionary change—that is, genomic change. It might simply be a problem addressed by existing cognitive processes, in different ways, and at different times. There is sometimes an unfortunate tendency to confuse literal evolutionary (genomic) change with historical change, two entirely distinct phenomena. As already noted, there is very strong evidence that there has been no relevant evolution of the language faculty since the trek from Africa some 50,000 years ago, though undoubtedly there has been a great deal of change, even invention of modes of externalization (as in sign language). Confusion about these matters could be overcome by replacing the metaphoric notions “evolution of language” and “language change” by their more exact counterparts: evolution of the organisms that use language, and change in the ways they do so. In these more accurate terms, emergence of the language faculty involved evolution, while historical change (which continues constantly) does not.

Again, these seem to be the simplest assumptions, and there is no known reason to reject them. If they are generally on the right track, it follows that externalization may not have evolved at all; rather, it might have been a process of problem solving using existing cognitive capacities. Evolution in the biological sense of the term would then be restricted to the mutation that yielded the operation Merge, along with whatever residue resists explanation in terms of the strong minimalist thesis and any language-specific constraints that might exist on the solution to the cognitive problem of externalization. Accordingly, any approach to evolution of language that focuses on communication, or the sensory-motor system, or statistical properties of spoken language, and the like, may well be seriously misguided. That judgment covers quite a broad range, as those familiar with the literature will be aware.

Returning to the two initial salient questions, we have at least some suggestions—reasonable ones we think—about how it came about that there is even one language, and why languages appear to vary so widely—the latter

⁶ Positing an independent, recursive, “language of thought” as a means to account for recursion in syntax leads to an explanatory regress as well as being unnecessary and quite obscure.

partly an illusion, much like the apparent limitless variety of organisms, all of them based on deeply conserved elements with phenomenal outcomes restricted by laws of nature (for language, computational efficiency).

There are other factors that may strongly influence language design— notably properties of the brain, now unknown—and there is plainly a lot more to say even about the topics to which we have briefly alluded here. But instead of pursuing these questions, let us turn briefly to lexical items, the conceptual atoms of thought and its ultimate externalization in varied ways.

Conceptual structures are found in other primates: probably actor–action–goal schemata, categorization, possibly the singular–plural distinction, and others. These were presumably recruited for language, though the conceptual resources of humans that enter into language use appear to be far richer. Specifically, even the “atoms” of computation, lexical items/concepts, appear to be uniquely human.

Crucially, even the simplest words and concepts of human language and thought lack the relation to mind-independent entities that appears to be characteristic of animal communication. The latter is held to be based on a one–one relation between mind/brain processes and “an aspect of the environment to which these processes adapt the animal’s behavior,” to quote cognitive neuroscientist Randy Gallistel, introducing a major collection of papers on animal communication (Gallistel 1990). According to Jane Goodall, the closest observer of chimpanzees in the wild, for them “the production of a sound in the *absence* of the appropriate emotional state seems to be an almost impossible task” (Goodall, cited in Tattersall 2002).

The symbols of human language and thought are sharply different. Their use is not automatically keyed to emotional states, and they do not pick out mind-independent objects or events in the external world. For human language and thought, it seems, there is no *reference* relation in the sense of Frege, Peirce, Tarski, Quine, and contemporary philosophy of language and mind. What we understand to be a river, a person, a tree, water, and so on, consistently turns out to be a creation of what seventeenth-century investigators called the human “cognoscitive powers,” which provide us with rich means to refer to the outside world from intricate perspectives. As the influential neo-Platonist Ralph Cudworth put the matter, it is only by means of the “inward ideas” produced by its “innate cognoscitive power” that the mind is able to “know and understand all external individual things,” articulating ideas that influenced Kant. The objects of thought constructed by the cognoscitive powers cannot be reduced to a “peculiar nature belonging” to the thing we are talking about, as David Hume summarized a century of inquiry. 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 object 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, succeeding insofar as shared mental constructs, background, concerns, presuppositions, and so on, allow for common perspectives to be (more or less) attained. These 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. How, no one has any idea. The fact that there even is a problem has barely been recognized, as a result of the powerful grip of the doctrines of referentialism.

Human cognoscitive powers provide us with a world of experience, different from the world of experience of other animals. Being reflective creatures, thanks to the emergence of the human capacity, humans try to make some sense of experience. These efforts are called myth, or religion, or magic, or philosophy, or, in modern English usage, science. For science, the concept of reference in the technical sense is a normative ideal: we hope that the invented concepts *photon* or *verb phrase* pick out some real thing in the world. And of course the concept of reference is just fine for the context for which it was invented in modern logic: formal systems, in which the relation of reference is stipulated, holding for example between numerals and numbers. But human language and thought do not seem to work that way, and endless confusion has resulted from failure to recognize that fact.

We enter here into large and extremely interesting topics that we will have to put aside. Let us just summarize briefly what seems to be the current best guess about unity and diversity of language and thought. In some completely unknown way, our ancestors developed human concepts. At some time in the very recent past, maybe about 75,000 years ago, an individual in a small group of hominids in East Africa underwent a minor mutation that provided the operation Merge—an operation that takes human concepts as computational atoms, and yields structured expressions that provide a rich language of thought. These processes might be computationally perfect, or close to it, hence the result of physical laws independent of humans. The innovation had obvious advantages, and took over the small group. At some later stage, the internal language of thought was connected to the sensory-motor system, a complex task that can be solved in many different ways and at different times,

The Bilingual Program: The Current State of its Evolution 41

and quite possibly a task that involves no evolution at all. In the course of these events, the human capacity took shape, yielding a good part of our “moral and intellectual nature,” using Wallace’s phrase (1871). The outcomes appear to be highly diverse, but they have an essential unity, reflecting the fact that humans are in fundamental respects identical, just as the hypothetical extraterrestrial scientist we conjured up earlier might conclude that there is only one language with minor dialectal variations, primarily—perhaps entirely—in mode of externalization.

To conclude, recall that even if this general story turns out to be more or less valid, and the huge gaps can be filled in, it will still leave unresolved problems that have been raised for hundreds of years. Among these is the question of how properties “termed mental” relate to “the organical structure of the brain,” in the eighteenth-century formulation; and the more mysterious problems of the creative and coherent ordinary use of language, a central concern of Cartesian science, still scarcely even at the horizons of inquiry.