

# Three Factors in Language Design

*Noam Chomsky*

The biolinguistic perspective regards the language faculty as an “organ of the body,” along with other cognitive systems. Adopting it, we expect to find three factors that interact to determine (I-) languages attained: genetic endowment (the topic of Universal Grammar), experience, and principles that are language- or even organism-independent. Research has naturally focused on I-languages and UG, the problems of descriptive and explanatory adequacy. The Principles-and-Parameters approach opened the possibility for serious investigation of the third factor, and the attempt to account for properties of language in terms of general considerations of computational efficiency, eliminating some of the technology postulated as specific to language and providing more principled explanation of linguistic phenomena.

*Keywords:* minimalism, principled explanation, Extended Standard Theory, Principles-and-Parameters, internal/external Merge, single-cycle derivation, phase

Thirty years ago, in 1974, an international meeting took place at MIT, in cooperation with the Royaumont Institute in Paris, on the topic of “biolinguistics,” a term suggested by the organizer, Massimo Piattelli-Palmarini, and the title of a recent book surveying the field and proposing new directions by Lyle Jenkins (2002).<sup>1</sup> This was only one of many such interactions in those years, including interdisciplinary seminars and international conferences.

The biolinguistic perspective began to take shape over 20 years before in discussions among a few graduate students who were much influenced by developments in biology and mathematics in the early postwar years, including work in ethology that was just coming to be known in the United States. One of them was Eric Lenneberg, whose seminal 1967 study *Biological Foundations of Language* remains a basic document of the field. Many of the leading questions discussed at the 1974 conference, and in the years leading up to it, remain very much alive today.

One of these questions, repeatedly brought up in the conference as “one of the basic questions to be asked from the biological point of view,” is the extent to which apparent principles of language, including some that had only recently come to light, are unique to this cognitive system or whether similar “formal arrangements” are found in other cognitive domains in humans or

This article is expanded from a talk presented at the annual meeting of the Linguistic Society of America, 9 January 2004. Thanks to Cedric Boeckx, Samuel David Epstein, Robert Freidin, Lyle Jenkins, Howard Lasnik, and Luigi Rizzi, among others, for comments on an earlier draft.

<sup>1</sup>The conference, titled “A Debate on Bio-Linguistics,” was held at Endicott House, Dedham, Massachusetts, 20–21 May 1974, and organized by the Centre Royaumont pour une science de l’homme, Paris.

other organisms. An even more basic question from the biological point of view is how much of language can be given a principled explanation, whether or not homologous elements can be found in other domains or organisms. The effort to sharpen these questions and to investigate them for language has come to be called the “Minimalist Program” in recent years, but the questions arise for any biological system and are independent of theoretical persuasion, in linguistics and elsewhere. Answers to these questions are fundamental not only to understanding the nature and functioning of organisms and their subsystems, but also to investigating their growth and evolution. For any biological system, language included, the only general question that arises about the program is whether it can be productively pursued or is premature.

In these remarks, I will try to identify what seem to me some of the significant themes in the past half-century of inquiry into problems of biolinguistics and to consider their current status. Several preliminary qualifications should be obvious. One is that the picture is personal; others would no doubt make different choices. A second is that things often seem clearer in retrospect than at the time, so there is some anachronism in this account, but not I think too much. A third is that I cannot even begin to mention the contributions of a great many people to the collective enterprise, particularly as the related fields have expanded enormously in the years since the 1974 conference.

The biolinguistic perspective views a person’s language as a state of some component of the mind, understanding “mind” in the sense of eighteenth-century scientists who recognized that after Newton’s demolition of the only coherent concept of body, we can only regard aspects of the world “termed mental” as the result of “such an organical structure as that of the brain” (Joseph Priestley). Among the vast array of phenomena that one might loosely consider language-related, the biolinguistic approach focuses attention on a component of human biology that enters into the use and acquisition of language, however one interprets the term “language.” Call it the “faculty of language,” adapting a traditional term to a new usage. This component is more or less on a par with the systems of mammalian vision, insect navigation, and others. In many of these cases, the best available explanatory theories attribute to the organism computational systems and what is called “rule-following” in informal usage—for example, when a recent text on vision presents the so-called rigidity principle as it was formulated 50 years ago: “if possible, and other rules permit, interpret image motions as projections of rigid motions in three dimensions” (Hoffman 1998:169). In this case, later work provided substantial insight into the mental computations that seem to be involved when the visual system follows these rules, but even for very simple organisms, that is typically no slight task, and relating mental computations to analysis at the cellular level is commonly a distant goal.

Adopting this conception, a language is a state of the faculty of language, an I-language, in technical usage.

The decision to study language as part of the world in this sense was regarded as highly controversial at the time, and still is. A more careful look will show, I think, that the arguments advanced against the legitimacy of the approach have little force (a weak thesis) and that its basic assumptions are tacitly adopted even by those who strenuously reject them, and indeed must be, even for coherence (a much stronger thesis). I will not enter into this interesting chapter of

contemporary intellectual history here, but will simply assume that crucial aspects of language can be studied as part of the natural world, adopting the biolinguistic approach that took shape half a century ago and that has been intensively pursued since, along different paths.

The language faculty is one component of what the cofounder of modern evolutionary theory, Alfred Russel Wallace, called “man’s intellectual and moral nature”: the human capacities for creative imagination, language and symbolism generally, mathematics, interpretation and recording of natural phenomena, intricate social practices, and the like, a complex of capacities that seem to have crystallized fairly recently, perhaps a little over 50,000 years ago, among a small breeding group of which we are all descendants—a complex that sets humans apart rather sharply from other animals, including other hominids, judging by traces they have left in the archaeological record. The nature of the “human capacity,” as some researchers now call it, remains a considerable mystery. It was one element of a famous disagreement between the two founders of the theory of evolution, with Wallace holding, contrary to Darwin, that evolution of these faculties cannot be accounted for in terms of variation and natural selection alone, but requires “some other influence, law, or agency,” some principle of nature alongside gravitation, cohesion, and other forces without which the material universe could not exist. Although the issues are framed differently today within the core biological sciences, they have not disappeared (see Wallace 1889: chap. 15, Marshack 1985).

It is commonly assumed that whatever the human intellectual capacity is, the faculty of language is essential to it. Many scientists agree with paleoanthropologist Ian Tattersall, who writes that he is “almost sure that it was the invention of language” that was the “sudden and emergent” event that was the “releasing stimulus” for the appearance of the human capacity in the evolutionary record—the “great leap forward” as Jared Diamond called it, the result of some genetic event that rewired the brain, allowing for the origin of modern language with the rich syntax that provides a multitude of modes of expression of thought, a prerequisite for social development and the sharp changes of behavior that are revealed in the archaeological record, also generally assumed to be the trigger for the rapid trek from Africa, where otherwise modern humans had apparently been present for hundreds of thousands of years (Tattersall 1998:24–25; see also Wells 2002). Tattersall takes language to be “virtually synonymous with symbolic thought.” Elaborating, one of the initiators of the Royaumont-MIT symposia, François Jacob, observed that “the role of language as a communication system between individuals would have come about only secondarily, as many linguists believe” (1982:59), perhaps referring to discussions at the symposia, where the issue repeatedly arose, among biologists as well. In the 1974 conference, his fellow Nobel laureate 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” (Luria 1974:195). “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 continued, 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 seventeenth-century cognitive revolution (1982:59). Jacob also stressed the common understanding that answers to questions about evolution “in most instances . . . can hardly be more than more or less reasonable guesses” (1982: 31).

We can add another insight of seventeenth- and eighteenth-century philosophy: that even the most elementary concepts of human language do not relate to mind-independent objects by means of some reference-like relation between symbols and identifiable physical features of the external world, as seems to be universal in animal communication systems. Rather, they are creations of the “cognoscitive powers” that provide us with rich means to refer to the outside world from certain perspectives, but are individuated by mental operations that cannot be reduced to a “peculiar nature belonging” to the thing we are talking about, as Hume summarized a century of inquiry. Those are critical observations about the elementary semantics of natural language, suggesting that its most primitive elements are related to the mind-independent world much as the internal elements of phonology are, not by a reference-like relation but as part of a considerably more intricate species of conception and action. It is for reasons such as these, though not clearly grasped at the time, that the early work in the 1950s adopted a kind of “use theory of meaning,” pretty much in the sense of John Austin and the later Wittgenstein: language was conceived as an instrument put to use for various human purposes, generating expressions including arrangements of the fundamental elements of the language, with no grammatical-ungrammatical divide, each basically a complex of instructions for use (see Chomsky 1955, hereafter *LSLT*).<sup>2</sup>

If this much is generally on the right track, then at least two basic problems arise when we consider the origins of the faculty of language and its role in the sudden emergence of the human intellectual capacity: first, the core semantics of minimal meaning-bearing elements, including the simplest of them; and second, the principles that allow infinite combinations of symbols, hierarchically organized, which provide the means for use of language in its many aspects. Accordingly, the core theory of language—Universal Grammar (UG)—must provide, first, a structured inventory of possible lexical items that are related to or perhaps identical with the concepts that are the elements of the “cognoscitive powers,” sometimes now regarded as a “language of thought” along lines developed by Jerry Fodor (1975); and second, means to construct from these lexical items the infinite variety of internal structures that enter into thought, interpretation, planning, and other human mental acts, and that are sometimes put to use in action, including the externalization that is a secondary process if the speculations just reviewed turn out to be correct. On the first problem, the apparently human-specific conceptual-lexical apparatus, there is important work on relational notions linked to syntactic structures and on the partially mind-internal objects that appear to play a critical role (events, propositions, etc.).<sup>3</sup> But there is little beyond descriptive remarks on the core referential apparatus that is used to talk about the world. The second problem has been central to linguistic research for half a century, with a long history before in different terms.

<sup>2</sup> For later discussion, see among others Chomsky 1966, 2001b, McGilvray 1999, Antony and Hornstein 2003.

<sup>3</sup> For insightful review and original analysis, see Borer 2004a,b.

The biolinguistic approach adopted from the outset the point of view that C. R. Gallistel (1997) calls “the norm these days in neuroscience” (p. 86), the “modular view of learning”: the conclusion that in all animals, learning is based on specialized mechanisms, “instincts to learn” (p. 82) in specific ways. We can think of these mechanisms as “organs within the brain” (p. 86), achieving states in which they perform specific kinds of computation. Apart from “extremely hostile environments” (p. 88), they change states under the triggering and shaping effect of external factors, more or less reflexively, and in accordance with internal design. That is the “process of learning” (Gallistel 1997, 1999), though “growth” might be a more appropriate term, avoiding misleading connotations of the term “learning.” The modular view of learning of course does not entail that the component elements of the module are unique to it: at some level, everyone assumes that they are not—the cellular level, for example—and the question of the level of organization at which unique properties emerge remains a basic one from a biological point of view, as it was at the 1974 conference.

Gallistel’s observations recall the concept of “canalization” introduced into evolutionary and developmental biology by C. H. Waddington over 60 years ago, referring to processes “adjusted so as to bring about one definite end result regardless of minor variations in conditions during the course of the reaction,” thus ensuring “the production of the normal, that is optimal type in the face of the unavoidable hazards of existence” (Waddington 1942). That seems to be a fair description of the growth of language in the individual. A core problem of the study of the faculty of language is to discover the mechanisms that limit outcomes to “optimal types.”

It has been recognized since the origins of modern biology that such constraints enter not only into the growth of organisms but also into their evolution, with roots in the earlier tradition that Stuart Kauffman calls “rational morphology” (1993:3–5).<sup>4</sup> In a classic contemporary paper, John Maynard Smith and associates trace the post-Darwinian reformulation back to Thomas Huxley, who was struck by the fact that there appear to be “predetermined lines of modification” that lead natural selection to “produce varieties of a limited number and kind” for every species (Maynard Smith et al. 1985:266).<sup>5</sup> They review a variety of such constraints in the organic world and describe how “limitations on phenotypic variability” are “caused by the structure, character, composition, or dynamics of the developmental system,” pointing out also that such “developmental constraints . . . undoubtedly play a significant role in evolution” though there is yet “little agreement on their importance as compared with selection, drift, and other such factors in shaping evolutionary history” (p. 265). At about the same time, Jacob wrote that “the rules controlling embryonic development,” almost entirely unknown, interact with other constraints imposed by general body plan, mechanical properties of building materials, and other factors in “restricting possible changes of structures and functions” in evolutionary development (1982:21), providing “architectural constraints” that “limit adaptive scope and channel evolutionary patterns” (Erwin 2003:1683). The best-known of the figures who devoted much of their work to these topics are

<sup>4</sup> For comment in a linguistic context, see Boeckx and Hornstein 2003. For more general discussion, see Jenkins 2000.

<sup>5</sup> For review of some of these topics, see Stewart 1998.

D'Arcy Thompson and Alan Turing, who took a very strong view on the central role of such factors in biology. In recent years, such considerations have been adduced for a wide range of problems of development and evolution, from cell division in bacteria to optimization of structure and function of cortical networks, even to proposals that organisms have “the best of all possible brains,” as argued by computational neuroscientist Christopher Cherniak (1995:522).<sup>6</sup> The problems are at the border of inquiry, but their significance is not controversial.

Assuming that the faculty of language has the general properties of other biological systems, we should, therefore, be seeking three factors that enter into the growth of language in the individual:

1. Genetic endowment, apparently nearly uniform for the species, which interprets part of the environment as linguistic experience, a nontrivial task that the infant carries out reflexively, and which determines the general course of the development of the language faculty. Among the genetic elements, some may impose computational limitations that disappear in a regular way through genetically timed maturation. Kenneth Wexler and his associates have provided compelling evidence of their existence in the growth of language, thus providing empirical evidence for what Wexler (to appear) calls “Lenneberg’s dream.”
2. Experience, which leads to variation, within a fairly narrow range, as in the case of other subsystems of the human capacity and the organism generally.
3. Principles not specific to the faculty of language.

The third factor falls into several subtypes: (a) principles of data analysis that might be used in language acquisition and other domains; (b) principles of structural architecture and developmental constraints that enter into canalization, organic form, and action over a wide range, including principles of efficient computation, which would be expected to be of particular significance for computational systems such as language. It is the second of these subcategories that should be of particular significance in determining the nature of attainable languages.

Those exploring these questions 50 years ago assumed that the primitive step of analysis of linguistic experience would be feature-based phonetic analysis, along lines described by Roman Jakobson and his associates (see Jakobson, Fant, and Halle 1953). We also tried to show that basic prosodic properties reflect syntactic structure that is determined by other principles, including crucially a principle of cyclic computation that was extended much more generally in later years (see Chomsky, Halle, and Lukoff 1956). The primitive principles must also provide what George Miller called “chunking,” identification of phonological words in the string of phonetic units. In *LSLT* (p. 165), I adopted Zellig Harris’s (1955) proposal, in a different framework, for identifying morphemes in terms of transitional probabilities, though morphemes do not have the required beads-on-a-string property. The basic problem, as noted in *LSLT*, is to show that such statistical

<sup>6</sup> See also Laughlin and Sejnowski 2003, Cherniak et al. 2004, and Physics News Update 2001 reporting Howard, Rutenberg, and de Vet 2001.

methods of chunking can work with a realistic corpus. That hope turns out to be illusory, as has recently been shown by Thomas Gambell and Charles Yang (2003), who go on to point out that the methods do, however, give reasonable results if applied to material that is preanalyzed in terms of the apparently language-specific principle that each word has a single primary stress. If so, then the early steps of compiling linguistic experience might be accounted for in terms of general principles of data analysis applied to representations preanalyzed in terms of principles specific to the language faculty, the kind of interaction one should expect among the three factors.

In *LSLT*, it was assumed that the next step would be assignment of chunked items to syntactic categories, again by general principles of data analysis. A proposal with an information-theoretic flavor was tried by hand calculations in that precomputer age, with suggestive results, but the matter has never been pursued, to my knowledge. Surely what are called “semantic properties” are also involved, but these involve nontrivial problems at the most elementary level, as mentioned earlier. The assumption of *LSLT* was that higher levels of linguistic description, including morphemes, are determined by a general format for rule systems provided by UG, with selection among them in terms of a computational procedure that seeks the optimal instantiation, a notion defined in terms of UG principles of significant generalization. Specific proposals were made then and in the years that followed. In principle, they provided a possible answer to what came to be called the “logical problem of language acquisition,” but they involved astronomical calculation and therefore did not seriously address the issues.

The main concerns in those years were quite different, as they still are. It may be hard to believe today, but it was commonly assumed 50 years ago that the basic technology of linguistic description was available and that language variation was so free that nothing of much generality was likely to be discovered. As soon as efforts were made to provide fairly explicit accounts of the properties of languages, however, it became obvious how little was known, in any domain. Every specific proposal yielded a treasure trove of counterevidence, requiring complex and varied rule-systems even to achieve a very limited approximation to descriptive adequacy. That was highly stimulating for inquiry into language, but it also left a serious quandary, since the most elementary considerations led to the conclusion that UG must impose narrow constraints on possible outcomes—sometimes called “poverty of stimulus” problems in the study of language, though the term is misleading because this is just a special case of basic issues that arise universally for organic growth.

A number of paths were pursued to try to resolve the tension. The most successful turned out to be efforts to formulate general principles, attributed to UG—that is, the genetic endowment—leaving a somewhat reduced residue of phenomena that would result, somehow, from experience. Early proposals were the A-over-A Principle, conditions on *wh*-extraction from *wh*-phrases (relatives and interrogatives), simplification of T-markers to base recursion (following observations by Charles Fillmore) and cyclicity (an intricate matter, as shown in an important paper of Robert Freidin’s (1978) and insightfully reviewed in a current paper of Howard Lasnik’s (to appear) which shows that many central questions remain unanswered), later John Robert Ross’s (1967) classic study of taxonomy of islands that still remains a rich store of ideas and observations to explore, then attempts to reduce islands to such properties as locality and structure

preservation, and so on. These approaches had some success, but the basic tensions remained unresolved at the time of the 1974 conference.

Within a few years, the landscape had changed considerably. In part this was because of great progress in areas that had hitherto been explored only in limited ways, including truth- and model-theoretic semantics and prosodic structures. In part it was the result of a vast array of new materials from studies of much greater depth than previously undertaken, and into a much wider variety of languages, much of it traceable to Richard Kayne's work and his lectures in Europe, which inspired far-reaching inquiry into Romance and Germanic languages, later other languages, also leading to many fruitful ideas about the principles of UG. About 25 years ago, much of this work crystallized in a radically different approach to UG, the Principles-and-Parameters (P&P) framework, which for the first time offered the hope of overcoming the tension between descriptive and explanatory adequacy. This approach sought to eliminate the format framework entirely, and with it, the traditional conception of rules and constructions that had been pretty much taken over into generative grammar. That much is familiar, as is the fact that the new P&P framework led to an explosion of inquiry into languages of the most varied typology, yielding new problems previously not envisioned, sometimes answers, and the reinvigoration of neighboring disciplines concerned with acquisition and processing, their guiding questions reframed in terms of parameter setting within a fixed system of principles of UG with at least visible contours. Alternative paths, variously interrelated, were leading in much the same direction, including Michael Brody's highly illuminating work (1995, 2003). No one familiar with the field has any illusion today that the horizons of inquiry are even visible, let alone at hand, in any domain.

Abandonment of the format framework also had a significant impact on the biolinguistic program. If, as had been assumed, acquisition is a matter of selection among options made available by the format provided by UG, then the format must be rich and highly articulated, allowing relatively few options; otherwise, explanatory adequacy is out of reach. The best theory of language must be a very unsatisfactory one from other points of view, with a complex array of conditions specific to human language, restricting possible instantiations. The only plausible theories had to impose intricate constraints on the permissible relations between sound and meaning, all apparently specific to the faculty of language. The fundamental biological issue of principled explanation could barely be contemplated, and correspondingly, the prospects for serious inquiry into evolution of language were dim; evidently, the more varied and intricate the conditions specific to language, the less hope there is for a reasonable account of the evolutionary origins of UG. These are among the questions that were raised at the 1974 symposium and others of the period, but they were left as apparently irresolvable problems.

The P&P framework offered prospects for resolution of these tensions as well. Insofar as this framework proves valid, acquisition is a matter of parameter setting and is therefore divorced entirely from the remaining format for grammar: the principles of UG. There is no longer a conceptual barrier to the hope that the UG might be reduced to a much simpler form, and that the basic properties of the computational systems of language might have a principled explanation instead of being stipulated in terms of a highly restrictive language-specific format for grammars. Within a P&P framework, what had previously been the worst theory—anything goes—might

be within the realm of conceivability at least, since it is no longer refuted instantly by the fact that it renders acquisition impossible. Returning to the three factors of language design, adoption of a P&P framework overcomes a difficult conceptual barrier to shifting the burden of explanation from the first factor, the genetic endowment, to the third factor, language-independent principles of data processing, structural architecture, and computational efficiency, thereby providing some answers to the fundamental questions of biology of language, its nature and use, and perhaps even its evolution.

As noted by Luigi Rizzi (pers. comm.), whose comments I paraphrase here, abandonment of the construction-based view of earlier (and traditional) work had further consequences for the biolinguistic program. In the earlier framework, not only rules but also UG principles were expressed in terms of grammatical constructions (islands, specified-subject and other constraints on operations, Emonds's structure-preserving hypothesis, filters, etc.), all inherently specific to language, without even remote counterparts in other biological systems. Within the P&P framework, the basic computational ingredients are considerably more abstract (locality, minimal search, basic recursion, etc.), and it becomes quite reasonable to seek principled explanation in terms that may apply well beyond language, as well as related properties in other systems.

As noted, the third factor subsumes two kinds of principles: (a) data processing, and (b) architectural/computational-developmental constraints. Consider the first category. Within the general P&P framework, language acquisition is interpreted in terms of parameter setting; an acquisition theory would seek to set forth the mechanisms used to fix parameter values. That requires some understanding of what the parameters are, and how they are organized, perhaps into a hierarchic structure with each choice of value setting limits on subsequent choices. The most far-reaching approach of this sort that I know of is Mark Baker's (2001). A somewhat different approach, also within the P&P framework, is suggested by Charles Yang (2002). Developing suggestions of Thomas Roeper's, Yang proposes that UG provides the neonate with the full array of possible languages, with all parameters valued, and that incoming experience shifts the probability distribution over languages in accord with a learning function that could be quite general. At every stage, all languages are in principle accessible, but only for a few are probabilities high enough so that they can actually be used. He provides interesting empirical evidence in support of this approach, which carries very early proposals about general data-processing principles to a new plane of sophistication.

Consider next language-independent principles of structural architecture and computational efficiency, the subcategory of the third factor in the design of any biological system that might be expected to be particularly informative about its nature. With the conceptual barriers imposed by the format framework overcome, we need no longer assume that the means of generating structured expressions are highly articulated and specific to language. We can seriously entertain the possibility that they might be reducible to language-independent principles, whether or not there are homologous elements in other domains and organisms. We can, in short, try to sharpen the question of what constitutes a principled explanation for properties of language, and turn to one of the most fundamental questions of the biology of language: to what extent does language approximate an optimal solution to conditions that it must satisfy to be usable at all, given

extralinguistic structural architecture? These conditions take us back to the traditional characterization of language, since Aristotle at least, as a system that links sound and meaning. In our terms, the expressions generated by a language must satisfy two interface conditions: those imposed by the sensorimotor system SM and by the conceptual-intentional system C-I that enters into the human intellectual capacity and the variety of speech acts.

Let us put aside the possibility that the linkage to sound is a secondary one, for reasons mentioned and in the light of recent discoveries on modality-independence of language. We can regard an explanation of properties of language as *principled* insofar as it can be reduced to properties of the interface systems and general considerations of computational efficiency and the like. Needless to say, these “external” conditions are only partially understood: we have to learn about the conditions that set the problem in the course of trying to solve it. The research task is interactive: to clarify the nature of the interfaces and optimal computational principles through investigation of how language partially satisfies the conditions they impose, not an unfamiliar feature of rational inquiry. Independently, the interface systems can be studied on their own, including comparative study that has been productively underway. And the same is true of principles of efficient computation, applied to language in recent work by many investigators with important results (see, e.g., Collins 1997, Epstein 1999, Epstein et al. 1998, Frampton and Gutmann 1999). In a variety of ways, then, it is possible to both clarify and address some of the basic problems of the biology of language.

It is perhaps worth recalling that the intuition that language is “well designed” relative to interface conditions has been a valuable heuristic for many years. The search for some reasonable notion of “simplicity of grammars” goes back to the earliest exploration of generative grammar in the late 1940s. It has also proven very useful over the years to ask whether apparent redundancy of principles is real, or indicates some error of analysis. A well-known example is passive forms of exceptional Case-marking constructions, which seemed at one time to be generated by both raising and passive transformations. Dissatisfaction with such overlapping conditions led to the realization that the transformations did not exist: rather, just a general rule of movement, which we can now see to be a virtual conceptual necessity. More recent work suggests that such apparently language-specific conditions as Jean-Roger Vergnaud’s Case Filter and its descendants may be reducible to virtual conceptual necessity as well, in this case, interpretability of constructions at the interface, topics reviewed and elaborated recently by Freidin and Vergnaud (2001). Throughout, the guiding intuition that redundancy in computational structure is a hint of error has proven to be productive and often verified, suggesting that something deeper is involved, a topic opened to more direct investigation as conceptual barriers were removed within the P&P framework.

To proceed further, for concreteness, consider the Extended Standard Theory (EST), now often called the “Y-model,” which was modified and improved extensively as it was reconstructed within the P&P framework—though I should perhaps stress again that the minimalist questions to which I am now turning arise within any theoretical orientation. In the early 1990s, Howard Lasnik and I published a sketch of our own best understanding of UG in these terms (Chomsky and Lasnik 1993). Taking that as a point of departure, let us ask to what extent its assumptions

can be revised or eliminated in favor of principled explanation in terms of interface conditions and general principles.

The EST/Y-model postulates three internal linguistic levels in addition to the interface levels: D-Structure, S-Structure, and LF. A linguistic level is postulated to host specific operations. Thus, D-Structure is the locus of lexical insertion and projection to X-bar constructions; S-Structure is the point of transfer of the computation to the sound interface, and the transition from overt to covert operations; LF is the output of overt and covert operations, and the input to the mapping to the meaning interface. Three internal levels require five operations, each assumed to be cyclic: the operations forming D-Structures by the cyclic operations of X-bar theory from selected items from the lexicon; the overt syntactic cycle from D- to S-Structure, the phonological/morphological cycle mapping S-Structure to the sound interface, the covert syntactic cycle mapping S-Structure to LF, and formal semantic operations mapping LF compositionally to the meaning interface. The similarity of the overt and covert cycles was recognized to be problematic as soon as the framework was developed 35 years ago, but the problem is in fact much broader: there are five cycles operating in rather similar fashion. One crucial question, then, is whether internal levels, not forced by interface conditions, can be eliminated, and the five cycles reduced to one. If possible, that would be a substantial step forward, with many consequences.

One natural property of efficient computation, with a claim to extralinguistic generality, is that operations forming complex expressions should consist of no more than a rearrangement of the objects to which they apply, not modifying them internally by deletion or insertion of new elements. If tenable, that sharply reduces computational load: what has once been constructed can be “forgotten” in later computations, in that it will no longer be changed. That is one of the basic intuitions behind the notion of cyclic computation. The EST/Y-model and other approaches violate this condition extensively, resorting to bar levels, traces, indices, and other devices, which both modify given objects and add new elements. A second question, then, is whether all of this technology is eliminable, and the empirical facts susceptible to principled explanation in accord with the “no-tampering” condition of efficient computation.

Other questions arise about the variety of operations—phrase structure, transformations, reconstruction, and so on; about the crucial reliance on such notions as government that seem to have no principled motivation; and about many principles that are hard even to formulate except in terms specific to the language faculty. The general question is, How far can we progress in showing that all such language-specific technology is reducible to principled explanation, thus isolating the core properties that are essential to the language faculty, a basic problem of biolinguistics?

An elementary fact about the language faculty is that it is a system of discrete infinity. Any such system is based on a primitive operation that takes  $n$  objects already constructed, and constructs from them a new object: in the simplest case, the set of these  $n$  objects. Call that operation Merge. Either Merge or some equivalent is a minimal requirement. With Merge available, we instantly have an unbounded system of hierarchically structured expressions. The simplest account of the “Great Leap Forward” in the evolution of humans would be that the brain was rewired,

perhaps by some slight mutation, to provide the operation Merge, at once laying a core part of the basis for what is found at that dramatic “moment” of human evolution, at least in principle; to connect the dots is no trivial problem. There are speculations about the evolution of language that postulate a far more complex process: first some mutation that permits two-unit expressions (yielding selectional advantage in overcoming memory restrictions on lexical explosion), then mutations permitting larger expressions, and finally the Great Leap that yields Merge. Perhaps the earlier steps really took place, but a more parsimonious speculation is that they did not, and that the Great Leap was effectively instantaneous, in a single individual, who was instantly endowed with intellectual capacities far superior to those of others, transmitted to offspring and coming to predominate, perhaps linked as a secondary process to the SM system for externalization and interaction, including communication as a special case. At best a reasonable guess, as are all speculations about such matters, but about the simplest one imaginable, and not inconsistent with anything known or plausibly surmised. In fact, it is hard to see what account of human evolution would not assume at least this much, in one or another form.<sup>7</sup>

Similar questions arise about growth of language in the individual. It is commonly assumed that there is a two-word stage, a three-word stage, and so on, with an ultimate Great Leap Forward to unbounded generation. That is observed in performance, but it is also observed that at the early stage the child understands much more complex expressions, and that random modification of longer ones—even such simple changes as placement of function words in a manner inconsistent with UG or the adult language—leads to confusion and misinterpretation. It could be that unbounded Merge, and whatever else is involved in UG, is present at once, but only manifested in limited ways for extraneous reasons (memory and attention limitations and the like)—matters discussed at the 1974 symposium and now possible to formulate more explicitly, for example, in the frameworks developed by Wexler and Yang, already mentioned.<sup>8</sup>

Suppose, then, that we adopt the simplest assumption: the Great Leap Forward yields Merge. The fundamental question of biology of language mentioned earlier then becomes, What else is specific to the faculty of language?

Unless some stipulation is added, there are two subcases of the operation Merge. Given A, we can merge B to it from outside A or from within A; these are external and internal Merge, the latter the operation called “Move,” which therefore also “comes free,” yielding the familiar displacement property of language. That property had long been regarded, by me in particular, as an “imperfection” of language that has to be somehow explained, but in fact it is a virtual conceptual necessity;<sup>9</sup> some version of transformational grammar seems to be the null hypothesis,

<sup>7</sup> For discussion in a broader framework, see Hauser, Chomsky, and Fitch 2002; and for intriguing extensions, see Piattelli-Palmarini and Uriagereka, to appear.

<sup>8</sup> Rizzi (to appear) argues that performance restrictions interact with UG principles: inconsistency with the target language in early production is grammatically based, with parameters fixed and principles observed, but performance-driven until the production system matures. The logic is similar to that of Gambell and Yang (2003) and Wexler (to appear).

<sup>9</sup> Options that “come free” may, of course, not be used, for example, in invented symbolic systems that keep to external Merge—possibly because they do not have to make use of the edge properties that are typically discourse-related.

and any other mechanisms, beyond internal Merge, carry a burden of proof. Assuming the no-tampering condition that minimizes computational load, both kinds of Merge to A will leave A intact. That entails merging to the edge, the “extension condition,” which can be understood in different ways, including the “tucking-in” theory of Norvin Richards (2001), which is natural within the probe-goal framework of recent work, and which can also be interpreted to accommodate head adjunction.

The no-tampering condition also entails the so-called copy theory of movement, which leaves unmodified the objects to which it applies, forming an extended object. As movement operations began to emerge with improvements of the more intricate *LSLT* framework, it was assumed that they must compound movement and deletion, because of the general architecture of the system. In the EST/Y-model architecture, that was unnecessary, but since there did appear to be clear effects of the position left by Move, it was assumed that the operation leaves a trace, a new object coindexed with the moved element; thus, the objects subject to the operation are modified, and new entities are introduced, violating elementary conditions of optimal computation. That was an error—mine in this case. Neither indexing nor the notion of traces is required if we adopt the simplest assumptions: the copy theory. An important additional gain is that rules of reconstruction can be eliminated, and the phenomena can be accounted for more effectively, as has been shown in a great deal of recent work.

Move leaves a copy, perhaps several, all transmitted to the phonological component. If language is optimized for communicative efficiency, we would expect all to be spelled out: that would overcome many of the filler-gap problems faced by processing models. If language is optimized for satisfaction of interface conditions, with minimal computation, then only one will be spelled out, sharply reducing phonological computation. The latter is the case, one of many reasons to suspect that speculations of the kind mentioned earlier about language design and evolution are on the right track. Which one is spelled out? There are several ideas about that. The most convincing seems to me Jon Nissenbaum’s (2000), in phase-theoretic terms to which I will return. At the phase level, two basic operations apply: transfer to the interface, and Merge, either external or internal. If internal Merge precedes transfer, movement is overt; otherwise, it is covert. If movement is covert, transfer has already spelled out the lower copy; if overt, the choice is delayed to the next phase. Željko Bošković (2001) has pointed out that there should be an exception to such minimal computation if the phonetic interface requires partial spell-out of a copy, and he has produced interesting examples. To take a different one discovered by Idan Landau (to appear), VP-fronting in Hebrew sometimes leaves the V behind, namely, when it has to satisfy what Lasnik called the “Stranded Affix Filter,” a phonetic condition on affixes.<sup>10</sup> If these are the only exceptions, then syntactic and phonological computation are optimally efficient in this regard, at least.

The two kinds of Merge appear to correlate closely with interface properties, capturing much of the dual aspect of semantics that was becoming clearer in the 1960s in work of Ray Jackendoff

<sup>10</sup> Landau’s formulation is somewhat different.

(1969), Peter Culicover (1970), Lyle Jenkins (1972), and others, sometimes expressed in terms of deep and surface interpretation. External Merge correlates with argument structure, internal Merge with edge properties, scopal or discourse-related (new and old information, topic, etc.). The correlation is not perfect, at least on the surface, but close enough to suggest that apparent violations should fall under broader principles. For external Merge, these go beyond argument structure associated with substantive categories, and they presumably include functional categories of the kind coming to light in a variety of inquiries, notably the very fruitful cartography projects (Cinque 2002, Rizzi 2004, Belletti 2004). Possibly such apparent exceptions as morphological identification of typical edge properties involve covert Move, also matters studied in current work.

External Merge yields the ‘‘base structures’’ of the EST/Y-model and earlier proposals. Since structures are constructed in parallel, there is no locus of lexical insertion. Accordingly, the notion of D-Structure is not only superfluous, but unformulable, a more far-reaching and desirable result. If base structures require nothing more than Merge, bar levels can also be eliminated, extending ideas of Naoki Fukui (1986) and Margaret Speas (1986) to full ‘‘bare phrase structure.’’ Proceeding further, along lines to which I will briefly return, S-Structure and LF seem to be unformulable as well, leaving us with just the interface levels, the optimal outcome.

The specifier-complement distinction loses independent significance, except in that the complement of a head H should be the only domain accessible to operations driven by H, by conditions of minimal search, the core property of c-command, but barring m-command and specifier-head relations (unless the specifier is a head, hence a possible probe), a strong and controversial empirical consequence of minimal search. ‘‘Complement’’ and ‘‘specifier’’ are just notations for first-Merge and later-Merge. Banning of multiple specifiers requires stipulation, hence empirical justification. The empirical evidence seems to me to leave the null hypothesis viable, maybe supported, though the issue is very much alive.

Each syntactic object generated contains information relevant to further computation. Optimally, that will be captured entirely in a single designated element, which should furthermore be identifiable with minimal search: its label, the element taken to be ‘‘projected’’ in X-bar-theoretic systems. The label, which will invariably be a lexical item introduced by external Merge, should be the sole probe for operations internal to the syntactic object, and the only element visible for further computations. We would hope that labels are determined by a natural algorithm, perhaps even by methods as far-reaching as those that Chris Collins (2002) has proposed in his label-free analysis. Again, these optimal possibilities are explored extensively in recent work by many researchers, along different though related paths.

Note that labels, or some counterpart, are the minimum of what is required, on the very weak assumption that at least some information about a syntactic object is needed for further computation, both for search within it and for its external role. Labeling underlies a variety of asymmetries: for example, in a head-XP construction, the label will always be the head, and the XP a ‘‘dependency’’; for substantive heads, an argument with a semantic role that it assigns. Therefore, such asymmetries need not be specified in the syntactic object itself, and must not be, because they are redundant. A head-complement structure is a set, not an ordered pair. There are many similar cases.

One asymmetry imposed by the phonetic interface is that the syntactic object derived must be linearized. Optimally, linearization should be restricted to the mapping of the object to the SM interface, where it is required for language-external reasons. If so, then no order is introduced in the rest of the syntactic computation: the narrow syntax and the mapping to the C-I interface. That has been a research objective for the past 25 years at least, ever since Tanya Reinhart's (1979) work on c-command, and it seems at least a plausible conclusion, though the consequences are quite restrictive and inconsistent with some of the most interesting work: papers by Danny Fox and Jon Nissenbaum on quantifier raising, antecedent-contained deletion, and late Merge of adjuncts, to mention some current examples (Fox 2002, Fox and Nissenbaum 1999). But I think there is reason to suspect that the simplest possibility can be maintained. How syntactic objects are linearized is also a lively topic. An early proposal is the head parameter, perhaps general, perhaps category-dependent. A different approach is based on Kayne's Linear Correspondence Axiom, which has inspired a great deal of productive work, including Kayne's (to appear) discovery of surprising gaps in language typology, and accounts for them. Other ideas are also being explored. A desideratum, throughout, is that the conclusion introduce few assumptions, optimally none, that cannot be reduced to the interface requirement and conditions of computational efficiency.

If linear order is restricted to the mapping to the phonetic interface, then it gives no reason to require the basic operation Merge to depart from the simplest form. As noted, many asymmetries can be derived from configurations, from properties of lexical items, and from the independently needed principles of labeling, so need not—hence should not—be specified by the elementary operation that forms the configuration. If that is universal, then the basic operation will be the simplest possible one: unstructured Merge, forming a set.

There may be—I think is—some reason to introduce an asymmetrical operation, for adjuncts, basically serving the function of composition of predicates. The most elementary choice is an ordered pair, which, at some point, must be simplified to a set for the purposes of linearization. Simplification optimally should be at the point of Spell-Out, that is, transfer to the two interfaces. Adopting this assumption, we can derive at least some, maybe all, of the intricate properties of reconstruction of adjuncts explored by Robert Freidin, Henk van Riemsdijk and Edwin Williams, David Lebeaux, and others since (Chomsky 2004).

It might be useful to allay some misgivings that appear in the technical literature. With minimal technical apparatus, a syntactic object  $X$  with the label  $A$  can be taken to be the set  $\{A, X\}$ , where  $X$  itself is a set  $\{Z, W\}$ . If  $A$  happens to be one of  $Z, W$ , then the set  $\{A, X\}$  has the formal properties of an ordered pair, as Norbert Wiener pointed out 90 years ago; that is, it will have formal properties of the pairs constructed for adjuncts.<sup>11</sup> That has been wrongly assumed to be problematic. It isn't. The possibility of giving a set-theoretic interpretation of ordered pairs, along Wiener's lines, has no bearing on whether ordered pairs are formal objects with their own

<sup>11</sup> The standard Wiener-Kuratowski definition of ordered pair is slightly different, Howard Lasnik points out, taking  $\langle a, b \rangle = \{\{a\}, \{a, b\}\}$ . Wiener's original proposal was more complex.

properties, “primitive objects” for mental computation that are distinct from sets, including those indicating label.

Without further stipulations, external Merge yields  $n$ -ary constituents. Particularly since Kayne’s (1981) work on unambiguous paths, it has been generally assumed that these units are mostly, maybe always, binary. If so, we would like to find a principled explanation. For internal Merge, it follows from elementary conditions of efficient computation, within the probe-goal framework, though precisely how this works raises interesting questions, with considerable empirical bite. For external Merge, one source could be conditions of argument structure imposed at the C-I interface. Another possibility has been suggested by Luigi Rizzi (pers. comm.), in terms of minimization of search in working memory.

The most restrictive case of Merge applies to a single object, forming a singleton set. Restriction to this case yields the successor function, from which the rest of the theory of natural numbers can be developed in familiar ways. That suggests a possible answer to a problem that troubled Wallace over a century ago: in his words, that the “gigantic development of the mathematical capacity is wholly unexplained by the theory of natural selection, and must be due to some altogether distinct cause” (1889:467), if only because it remained unused. One possibility is that the natural numbers result from a simple constraint on the language faculty, hence are not given by God, in accord with Kronecker’s famous aphorism, though the rest is created by man, as he continued. Speculations about the origin of the mathematical capacity as an abstraction from linguistic operations are not unfamiliar. There are apparent problems, including apparent dissociation with lesions and diversity of localization, but the significance of such phenomena is unclear for many reasons (including the issue of possession vs. use of the capacity). There may be something to these speculations, perhaps along the lines just indicated, which have the appropriate minimalist flavor.

On very weak assumptions about efficient computation, there should be no linguistic levels beyond those that are imposed by interface conditions: that is, the interface levels themselves. In fact, it is not even obvious that these exist. Thus, one can imagine, say, a computational process that sends parts of generated objects to the SM systems in the course of derivation (say, certain but not all phonetic properties), not just at its final stage; and the same on the meaning side. Though such proposals have never been explored, to my knowledge, they could turn out to be correct. Beyond this, however, we would like to determine whether all internal levels are dispensable, with the five cycles of the EST/Y-model reduced to one—better still unformulable, as D-Structure is. That will follow if computation relies solely on Merge, perhaps pair- as well as set-Merge, yielding syntactic objects that, at some point in the derivation, are transferred to the two interfaces: transfer to the sound interface is often called “Spell-Out.” Let us call the syntactic objects that are transferred “phases.” Optimally, once a phase is transferred, it should be mapped directly to the interface and then “forgotten”; later operations should not have to refer back to what has already been mapped to the interface—again, a basic intuition behind cyclic operations. We therefore hope to be able to establish a “Phase Impenetrability Condition,” which guarantees that mappings to the two interfaces can forget about what they have already done, a substantial

saving in memory. Whether that is feasible is a question only recently formulated, and barely explored. It raises many serious issues, but so far at least, no problems that seem insuperable.

If these general ideas are on the right track, then all internal levels are unformulable, hence dispensable in the strongest sense. We are left with only the interface levels, and the five cycles of the EST/Y-model are reduced to a single one, based on Merge. The cyclic properties of the mappings to the interface follow without comment. Pure cyclic computation is also required for the simplest account of uninterpretable features. Their values are redundant, determined by an Agree relation with interpretable features. They should therefore be unvalued in the lexicon, and when assigned a value, must be treated as what Juan Uriagereka (1998) calls a ‘‘virus’’: eliminated as soon as possible, hence at the phase level, where they are transferred to the phonological component for possible phonetic realization but eliminated from the syntactic derivation and from transfer to the semantic interface to avoid crash. That is again an immediate consequence of phase-based computation.<sup>12</sup>

The topic of uninterpretable features has moved to center stage since Vergnaud’s proposals concerning Case theory 25 years ago, and particularly in the past few years. Vergnaud’s original thesis that such elements as structural Case may be present even when not spelled out has received considerable empirical support. A very strong version of the thesis holds that inflectional features are in fact universal, ideas developed particularly by Halldór Sigurðsson (2003) as part of an extensive analysis of the role of these features, including Case, in speech events, and their fine-grained syntactic analysis.

Since the copy of a moved element remains, it is in principle accessible for higher syntactic computations (much as it can be accessed more globally at the meaning interface, under reconstruction). How deeply is such access allowed? Evidence so far is meager, but as far as I know, it never goes beyond one phase. We find examples of this, for example, with Icelandic quirky Case: a phase-internal nominative object can agree with an external inflection. The example is interesting and merits further inquiry, because it appears to undermine a possible phase-theoretic explanation for some of Vergnaud’s account of subjects that cannot appear where Case cannot be assigned, namely, that they are ‘‘invisible’’ because they are in the interior of a phase that has already been passed. In the Icelandic case, the object is not invisible, though the Phase Impenetrability Condition, which prevents any tampering with an earlier phase, blocks extraction. A different account, based on a richer variety of linguistic materials and some different assumptions, has been proposed by Andrew Nevins (2004).

What objects constitute phases? They should be as small as possible, to minimize computational load. For reasons mentioned earlier, they should at least include the domains in which uninterpretable features are valued; agreement features apart from nouns and structural Case for nouns. There is evidence, which I will not review here, that these domains are CP and vP (where v is ‘‘small v’’ that introduces full argument structure, transitive, or experiencer verbal structures) and that these are the only phases in the core clausal structures. Perhaps phases also include DP,

<sup>12</sup> See Chomsky 2001a. A consequence is the Earliness Principle of David Pesetsky and Esther Torrego (2001).

as is worked out by Peter Svenonius (2003), developing further the parallels between DP and CP. Keeping to CP and vP, the objects derived have a skeletal structure in which C and v are the labels driving internal operations and relevant to external Merge, and are also the points of feature valuation (as is a virtual necessity) and transfer. If so, a number of properties have a unified account. A possible speculation is that the more elaborate structures revealed by the cartographic inquiries are based on linearization of features in these labels, and possibly labels closely linked to them (as in the C-T connection).

The phases have correlates at the interfaces: argument structure or full propositional structure at the meaning side, relative independence at the sound side. But the correlation is less than perfect, which raises questions.

This computational architecture, if sustainable, seems to be about as good as possible. S-Structure and LF are no longer formulable as levels, hence disappear along with D-Structure, and computations are reduced to a single cycle. It should be that all operations are driven by the designated element of the phase, the phase head C, v. That should include internal Merge. But a problem arises, discovered by Julie Legate (2003): there are reconstruction sites at smaller categories, though movement cannot stop there but must move on to the phase head that drives the operation. A solution to the conundrum is developed by Cedric Boeckx (2003) in a slightly different framework, adapting ideas of Daiko Takahashi's (1994). In brief, the phase head drives the Move operation, but it proceeds category by category, in a manner occasionally suggested for other reasons, until it reaches a point where it must stop. That is usually the phase head, though there can be an exception, namely, if the phase head assigns its uninterpretable features to the head it selects. Thus, the phase head C may be the locus of agreement, selecting T and assigning it (unvalued)  $\phi$ -features, so that when raising of DP driven by C-agreement reaches the TP level, its uninterpretable features are valued and it is frozen, unable to proceed further. That yields many of the properties of the Extended Projection Principle (EPP) and, if generalized, also yields interesting empirical consequences that would carry us too far afield. But partial movement is impossible phase-internally in other cases; in particular, we can overcome the annoying lack of structures of the form *\*There seems a man to be in the room*, without the apparatus that has been proposed to account for the gap, which generalizes to the specifier of participial phrases and other such constructions. In a somewhat different framework, Jan Koster (2003) has developed the interesting idea that V-movement to T is "partial movement" to C, the true locus of Tense, analogous to the partial *wh*-movement in German discussed by Henk van Riemsdijk (1983) and Dana McDaniel (1989).

Internal Merge seems to be driven in part at least by uninterpretable features of the phase head, as a reflex of probe-goal matching. Insofar as that is true, it carries us part of the way toward a principled explanation for why languages should have uninterpretable features at all. But it is not clear how far. For example, there is considerable evidence for successive-cyclic movement through phase edges, but for  $\bar{A}$ -movement it is unclear whether there is independent motivation for the features that would be required to formulate the operations within this framework. It may be that phase heads have an edge feature, sometimes called an "EPP-feature" (by extension of the notion EPP) or an "occurrence feature" OCC (because the object formed is an

occurrence of the moved element in the technical sense). This edge feature permits raising to the phase edge without feature matching. Similar questions arise for A-movement, particularly in languages where EPP is satisfied by a nonnominative element, as in locative inversion, which may agree with T, as in Bantu, though not in English, though there is reason to doubt that locative inversion involves A-movement to subject in English. Mark Baker (2003) suggests a parameter distinguishing Bantu from Indo-European in this respect, with interesting empirical consequences, and it is possible that his conclusions extend to apparent attraction to T by a D-feature in Slavic and Japanese-type languages, along lines developed by James E. Lavine and Robert Freidin (2002) and by Shigeru Miyagawa (2004). It seems possible that the variety of cases might fall under a probe-goal framework with T parameterized somewhat along the lines that Baker suggests, but keeping to  $\phi$ -feature probe-goal matching with or without the requirement that the goal be active, that is, with unvalued Case; or, along lines that Miyagawa develops extensively, with Focus playing the role of agreement in some class of languages. If the phase heads have both agreement features (or Focus features) and edge features, then we would expect both to apply simultaneously. That assumption directly yields some curious phenomena about intervention effects discussed by Ken Hiraiwa (2002) and by Anders Holmberg and Thorbjörg Hróarsdóttir (2003), and also an account of some apparent exceptions to the Subject Island Condition. We move here well into the terrain of open problems—a vast terrain.

I have barely touched on the rich range of inquiries with these general motivations in the past few years, and now in progress. One broad-ranging summary, with many new ideas and new empirical materials, is Cedric Boeckx's recent book *Islands and Chains* (2003). It seems reasonably clear that there has been considerable progress in moving toward principled explanation that addresses fundamental questions of the biology of language. It is even more clear that these efforts have met one primary requirement for a sensible research program: stimulating inquiry that has been able to overcome some old problems while even more rapidly bringing to light new ones, previously unrecognized and scarcely even formulable, and enriching greatly the empirical challenges of descriptive and explanatory adequacy that have to be faced.

## References

- Antony, Louise M., and Norbert Hornstein, eds. 2003. *Chomsky and his critics*. Malden, Mass.: Blackwell.
- Baker, Mark. 2001. *The atoms of language*. New York: Basic Books.
- Baker, Mark. 2003. Agreement, dislocation, and partial configurationality. In *Formal approaches to function in grammar*, ed. by Andrew Carnie, Heidi Harley, and MaryAnn Willie, 107–132. Amsterdam: John Benjamins.
- Belletti, Adriana, ed. 2004. *The cartography of syntactic structures*. Vol. 3, *Structures and beyond*. Oxford: Oxford University Press.
- Boeckx, Cedric. 2003. *Islands and chains: Resumption as stranding*. Amsterdam: John Benjamins.
- Boeckx, Cedric, and Norbert Hornstein. 2003. The varying aims of linguistic theory. Ms., University of Maryland, College Park.
- Borer, Hagit. 2004a. *Structuring sense*. Vol. 1, *In name only*. Oxford: Oxford University Press.
- Borer, Hagit. 2004b. *Structuring sense*. Vol. 2, *The normal course of events*. Oxford: Oxford University Press.

- Bošković, Željko. 2001. *On the nature of the syntax-phonology interface*. Amsterdam: Elsevier.
- Brody, Michael. 1995. *Lexico-Logical Form: A radically minimalist theory*. Cambridge, Mass.: MIT Press.
- Brody, Michael. 2003. *Towards an elegant syntax*. London: Routledge.
- Cherniak, Christopher. 1995. Neural component placement. *Trends in Neuroscience* 18:522–527.
- Cherniak, Christopher, Zekeria Mokhtarzada, Raul Rodriguez-Esteban, and Kelly Changizi. 2004. Global optimization of cerebral cortex layout. *Proceedings of the National Academy of Science* online, 13 January 2004 (print version 27 January 2004, 101(4):1081–1086).
- Chomsky, Noam. 1955. *Logical structure of linguistic theory (LSLT)*. Ms. Parts of 1956 revision published in New York: Plenum, 1975, and Chicago: University of Chicago Press, 1985.
- Chomsky, Noam. 1966. *Cartesian linguistics*. New York: Harper & Row.
- Chomsky, Noam. 2001a. Derivation by phase. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 1–52. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 2001b. *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.
- Chomsky, Noam. 2004. Beyond explanatory adequacy. In *The cartography of syntactic structures*. Vol. 3, *Structures and beyond*, ed. by Adriana Belletti. Oxford: Oxford University Press.
- Chomsky, Noam, Morris Halle, and Fred Lukoff. 1956. On accent and juncture in English. In *For Roman Jakobson*, compiled by Morris Halle and others, 65–80. The Hague: Mouton.
- Chomsky, Noam, and Howard Lasnik. 1993. The theory of principles and parameters. In *Syntax: An international handbook of contemporary research*, ed. by Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 506–569. Berlin: de Gruyter. Reprinted in Noam Chomsky, *The Minimalist Program*, Cambridge, Mass.: MIT Press, 1995.
- Cinque, Guglielmo, ed. 2002. *The cartography of syntactic structures*. Vol. 1, *Functional structure in DP and IP*. Oxford: Oxford University Press.
- Collins, Chris. 1997. *Local economy*. Cambridge, Mass.: MIT Press.
- Collins, Chris. 2002. Eliminating labels. In *Derivation and explanation in the Minimalist Program*, ed. by Samuel David Epstein and T. Daniel Seely, 42–64. Malden, Mass.: Blackwell.
- Culicover, Peter. 1970. Syntactic and semantic investigation. Doctoral dissertation, MIT, Cambridge, Mass.
- Epstein, Samuel David. 1999. Un-principled syntax: The derivation of syntactic relations. In *Working minimalism*, ed. by Samuel David Epstein and Norbert Hornstein, 317–345. Cambridge, Mass.: MIT Press.
- Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima, and Hisatsugu Kitahara. 1998. *A derivational approach to syntactic relations*. Oxford: Oxford University Press.
- Erwin, Douglas. 2003. The Goldilocks hypothesis. *Science* 302 (5 December 2003):1682–1683.
- Fodor, Jerry. 1975. *Language of thought*. New York: Crowell.
- Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33: 63–96.
- Fox, Danny, and Jon Nissenbaum. 1999. Extraposition and scope: A case for overt QR. In *Proceedings of the 18th West Coast Conference on Formal Linguistics*, ed. by Sonya Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest, 132–144. Somerville, Mass.: Cascadilla Press.
- Frampton, John, and Sam Gutmann. 1999. Cyclic computation, a computationally efficient minimalist syntax. *Syntax* 2:1–27.
- Freidin, Robert. 1978. Cyclicity and the theory of grammar. *Linguistic Inquiry* 9:519–549.
- Freidin, Robert, and Jean-Roger Vergnaud. 2001. Exquisite connections: Some remarks on the evolution of linguistic theory. *Lingua* 111:639–666.
- Fukui, Naoki. 1986. A theory of category projection and its applications. Doctoral dissertation, MIT, Cambridge, Mass. Revised version published as *Theory of projection in syntax*. Cambridge: Cambridge University Press, and Stanford, Calif.: CSLI Publications, 1995.

- Gallistel, C. R. 1997. Neurons and memory. In *Conversations in the cognitive neurosciences*, ed. by Michael S. Gazzaniga. Cambridge, Mass.: MIT Press.
- Gallistel, C. R. 1999. The replacement of general-purpose learning models with adaptively specialized learning modules. In *The cognitive neurosciences*, ed. by Michael S. Gazzaniga. 2nd ed. Cambridge, Mass.: MIT Press.
- Gambell, Thomas, and Charles Yang. 2003. Scope and limits of statistical learning in word segmentation. Ms., Yale University, New Haven, Conn.
- Harris, Zellig S. 1955. From phoneme to morpheme. *Language* 31:190–222.
- Hauser, Marc D., Noam Chomsky, and W. Tecumseh Fitch. 2002. The faculty of language. *Science* 198: 1569–1579.
- Hiraiwa, Ken. 2002. Multiple Agree. Paper presented at the 25th GLOW Colloquium, Utrecht.
- Hoffman, Donald. 1998. *Visual intelligence*. New York: Norton.
- Holmberg, Anders, and Thorbjörg Hróarsdóttir. 2003. Agreement and movement in Icelandic raising constructions. *Lingua* 113:997–1019.
- Jackendoff, Ray. 1969. Some rules of semantic interpretation in English. Doctoral dissertation, MIT, Cambridge, Mass.
- Jacob, François. 1982. *The possible and the actual*. New York: Pantheon.
- Jakobson, Roman, Gunnar Fant, and Morris Halle. 1953. *Preliminaries to speech analysis*. Cambridge, Mass.: MIT Press.
- Jenkins, Lyle. 1972. Modality in English syntax. Doctoral dissertation, MIT, Cambridge, Mass.
- Jenkins, Lyle. 2000. *Biolinguistics*. Cambridge: Cambridge University Press.
- Kauffman, Stuart. 1993. *The origins of order*. Oxford: Oxford University Press.
- Kayne, Richard. 1981. Unambiguous paths. In *Levels of syntactic representation*, ed. by Robert May and Jan Koster, 143–183. Dordrecht: Reidel.
- Kayne, Richard. To appear. Anti-symmetry in Japanese. In *Variation and universals in biolinguistics*, ed. by Lyle Jenkins.
- Koster, Jan. 2003. All languages are tense second. Ms., University of Groningen.
- Landau, Idan. To appear. Chain resolution in Hebrew V(P) fronting. *Syntax*.
- Lasnik, Howard. To appear. Conceptions of the cycle. In *Wh-movement on the move*, ed. by Lisa Lai-Shen Cheng and Norbert Corver. Cambridge, Mass.: MIT Press.
- Laughlin, Simon, and Terrence J. Sejnowski. 2003. Communication in neuronal networks. *Science* 301 (26 September 2003): 1870–1874.
- Lavine, James E., and Robert Freidin. 2002. The subject of defective T(ense) in Slavic. *Journal of Slavic Linguistics* 10:251–287.
- Legate, Julie Anne. 2003. Some interface properties of the phase. *Linguistic Inquiry* 34:506–516.
- Lenneberg, Eric. 1967. *Biological foundations of language*. New York: Wiley.
- Luria, Salvador. 1974. Transcript of remarks at “A Debate on Bio-Linguistics,” a conference organized by the Centre Royaumont pour une science de l’homme, Paris, held at Endicott House, Dedham, Mass., 20–21 May 1974.
- Marshall, Alexander. 1985. *Hierarchical evolution of the human capacity*. New York: American Museum of Natural History.
- Maynard Smith, J., R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, D. Raup, and L. Wolpert. 1985. Developmental constraints and evolution. *Quarterly Review of Biology* 60: 265–287.
- McDaniel, Dana. 1989. Partial and multiple *wh*-movement. *Natural Language & Linguistic Theory* 7: 565–604.
- McGilvray, James. 1999. *Chomsky: Language, mind, and politics*. Cambridge: Polity.
- Miyagawa, Shigeru. 2004. On the EPP. In *Proceedings of the EPP/Phase Workshop*, ed. by Martha McGinnis

- and Norvin Richards. MIT Working Papers in Linguistics. Cambridge, Mass.: MIT, Department of Linguistics and Philosophy, MITWPL.
- Nevins, Andrew. 2004. Derivations without the Activity Condition. In *Proceedings of the EPP/Phase Workshop*, ed. by Martha McGinnis and Norvin Richards. MIT Working Papers in Linguistics. Cambridge, Mass.: MIT, Department of Linguistics and Philosophy, MITWPL.
- Nissenbaum, Jon. 2000. Investigations of covert phrase movement. Doctoral dissertation, MIT, Cambridge, Mass.
- Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 355–426. Cambridge, Mass.: MIT Press.
- Physics News Update*. 2001. American Institute of Physics Bulletin of Physics, News Number 570, December 21, 2001, reporting study by M. Howard, A. D. Rutenberg, and S. de Vet, *Physical Review Letters*, 31 December 2001.
- Piattelli-Palmarini, Massimo, and Juan Uriagereka. To appear. In *Variation and universals in biolinguistics*, ed. by Lyle Jenkins.
- Reinhart, Tanya. 1979. Syntactic domains for semantic rules. In *Formal semantics and pragmatics for natural language*, ed. by Franz Guenther and Siegfried J. Schmidt, 107–130. Dordrecht: Reidel.
- Richards, Norvin. 2001. *Movement in language*. Oxford: Oxford University Press.
- Riemsdijk, Henk van. 1983. Correspondence effects and the Empty Category Principle. In *Studies in generative grammar and language acquisition*, ed. by Y. Otsu et al., 5–16. Tokyo: International Christian University.
- Rizzi, Luigi, ed. 2004. *The cartography of syntactic structures*. Vol. 2, *The structure of CP and IP*. Oxford: Oxford University Press.
- Rizzi, Luigi. To appear. On the grammatical basis of language development. In *Handbook of comparative syntax*, ed. by Guglielmo Cinque and Richard Kayne. Oxford: Oxford University Press.
- Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT, Cambridge, Mass. Published as *Infinite syntax!*. Norwood, N.J.: Ablex, 1986.
- Sigurðsson, Halldór. 2003. Meaningful silence, meaningless sounds. Ms., Lund University.
- Speas, Margaret. 1986. Adjunctions and projections in syntax. Doctoral dissertation, MIT, Cambridge, Mass.
- Stewart, Ian. 1998. *Life's other secret*. New York: Wiley.
- Svenonius, Peter. 2003. On the edge. Ms., University of Tromsø.
- Takahashi, Daiko. 1994. Minimality of movement. Doctoral dissertation, University of Connecticut, Storrs.
- Tattersall, Ian. 1998. *The origin of the human capacity*. New York: American Museum of Natural History.
- Uriagereka, Juan. 1998. *Rhyme and reason*. Cambridge, Mass.: MIT Press.
- Waddington, C. H. 1942. Canalization of development and the inheritance of acquired characters. *Nature* 150:563–565.
- Wallace, Alfred Russel. 1889. *Darwinism*. London: Macmillan.
- Wells, Spencer. 2002. *The journey of man: A genetic odyssey*. Princeton, N.J.: Princeton University Press.
- Wexler, Kenneth. To appear. Lenneberg's dream: Learning, normal language development and specific language impairment. In *Variation and universals in biolinguistics*, ed. by Lyle Jenkins.
- Yang, Charles. 2002. *Knowledge and learning in natural language*. Oxford: Oxford University Press.

*Department of Linguistics and Philosophy*

32-D840

MIT

Cambridge, Massachusetts 02139

chomsky@mit.edu