

〈特別寄稿／Special Contribution〉

# Minimal Computation and the Architecture of Language

Noam CHOMSKY

Massachusetts Institute of Technology

From the early days of the modern scientific revolution, there has been intense interest in human language, recognized to be a core feature of human nature and the primary capacity distinguishing modern humans from other creatures. In a contemporary interpretation, Ian Tattersall, one of the leading students of human evolution, writes that “the acquisition of the uniquely modern [human] sensibility was instead an abrupt and recent event.... And the expression of this new sensibility was almost certainly crucially abetted by the invention of what is perhaps the single most remarkable thing about our modern selves: language.”<sup>1</sup>

Centuries earlier, Galileo and the seventeenth century Port Royal logicians and grammarians were awed by the “marvelous invention” of a means to construct “from 25 or 30 sounds that infinity of expressions, which bear no resemblance to what takes place in our minds, yet enable us to reveal [to others] everything that we think, and all the various movements of our soul.” Descartes took this capacity to be a primary difference between humans and any beast-machine, providing a basic argument for his mind-body dualism. The great humanist Wilhelm von Humboldt characterized language as “a generative activity [*eine Erzeugung*]” rather than “a lifeless product” [*ein todes Erzeugtes*], *Energieia* rather than *Ergon*, and pondered the fact that somehow this activity “makes infinite use of finite means.”<sup>2</sup> For the last great representative of this tradition, Otto Jespersen, the central question of the study of language is how its structures “come into existence in the mind of a speaker” on the basis of finite experience, yielding a “notion of structure” that is “definite enough to guide him in framing sentences of his own,” crucially “free expressions” that are typically new to speaker and hearer. And more deeply, to go beyond to unearth “the great principles underlying the grammars of all languages” and by so doing to gain “a deeper insight into the innermost nature of human language and of human thought” – ideas that sound much less strange today than they did during the structuralist/behavioral science era that came to dominate much of the field through the first half of the 20<sup>th</sup> century, marginalizing the leading ideas and concerns of the tradition.<sup>3</sup>

Throughout this rich tradition of reflection and inquiry there were efforts to comprehend how humans can freely and creatively employ “an infinity of expressions” to express their thoughts in ways that are appropriate to circumstances though not determined by them, a

crucial distinction. However, tools were not available to make much progress in carrying these ideas forward. That difficulty was partially overcome by mid-20<sup>th</sup> century, thanks to the work of Gödel, Turing, and other great mathematicians that laid the basis for the modern theory of computability. These accomplishments provided a very clear understanding of how “finite means” can generate an “infinity of expressions,” thereby opening the way to formulating and investigating what we may consider to be the Basic Property of the human language faculty: a finitely-specified generative procedure, represented in the brain, that yields a discrete infinity of hierarchically structured expressions, each with a determinate interpretation at two interfaces: the sensorymotor interface SM for externalization in one or another sensory modality (usually, though not necessarily, sound); and the conceptual-intentional interface CI for reflection, interpretation, inference, planning, and other mental acts. Nothing analogous, even remotely similar, has been discovered in any other organism, thus lending substance to the judgments of the rich tradition.

It is important to recognize that the unbounded *use* of these finite means -- the actual production of speech in the free and creative ways that intrigued the great figures of the past -- still remains a mystery, not just in this domain, but for voluntary action generally. The mystery is graphically described by two of the most prominent scientists who study voluntary motion, Emilio Bizzi and Robert Ajemian, reviewing the state of the art today: “we have some idea as to the intricate design of the puppet and the puppet strings,” they write, “but we lack insight into the mind of the puppeteer.”<sup>4</sup>

That is not a slight problem. It lies at the borders of feasible scientific inquiry if not beyond, in a domain which human intelligence cannot penetrate. And if we if we are willing to accept the fact that we are organic creatures, not angels, we will join leading thinkers of the past – Descartes, Newton, Locke, Hume and others – in recognizing that some problems may be permanent mysteries for us.

The study of the finite means that are used in linguistic behavior – the puppet and the strings – has been pursued very successfully since the mid-twentieth century in what has come to be called the “generative enterprise” and “biolinguistic framework,” drawing from and contributing to the “cognitive revolution” that has been underway during this period. The kinds of questions that students are investigating today could not even have been formulated not many years ago, and there has been a vast explosion in the languages of the widest typological variety that have come under investigation, at a level of depth never before contemplated in the long and rich history of investigation of language since classical Greece and ancient India. There have been many discoveries along the way, regularly raising new problems and opening new directions of inquiry. In these respects the enterprise has had considerable success.

Departing from the assumptions of the structuralist/behaviorist era and returning to the

spirit of the tradition in new forms, the generative/biolinguistic enterprise takes a language to be an internal system, a “module” of the system of human cognitive capacities. In technical terms, a language is taken to be an “I-language” – where “I” stands for internal, individual, and intensional (meaning that we are concerned with the actual nature of the biological object itself rather than with some set of objects that it generates, such as a corpus of expressions or set of behaviors). Each I-language satisfies the Basic Property of human language, formulated above. Jespersen’s “great principles underlying the grammars of all languages” are the topic of Universal Grammar (UG), adapting a traditional term to a new framework, interpreted now as the theory of the genetic endowment for the faculty of language, the innate factors that determine the class of possible I-languages.

There is by now substantial evidence that UG is a species property, uniform among humans apart from severe pathology, and with no close analogue, let alone anything truly homologous, in the rest of the animal world. It seems to have emerged quite recently in evolutionary time, as Tattersall concluded, probably within the last 100,000 years. And we can be fairly confident that it has not evolved at least since our ancestors began to leave Africa some 50-60 thousand years ago. If so, then the emergence of the language faculty – of UG – was quite sudden in evolutionary time, which leads us to suspect that the Basic Property, and whatever else constitutes UG, should be very simple. Furthermore, since Eric Lenneberg’s pioneering work in the 1950s,<sup>5</sup> evidence has been accumulating that the human language faculty is dissociated from other cognitive capacities – though of course the use of language in perception (parsing) and production integrates the internal I-language with other capacities. That too suggests that whatever emerged quite suddenly (in evolutionary time) should be quite simple.

As the structuralist and behavioral science approaches took shape through the first half of the 20<sup>th</sup> century, it came to be generally assumed that the field faced no fundamental problems. Methods of analysis were available, notably Zellig Harris’s *Methods in Structural Linguistics*, which provided the means to reduce a corpus of materials to an organized form, the primary task of the discipline. The problems of phonology, the major focus of inquiry, seemed to be largely understood. As a student in the late 1940s, I remember well the feeling that “this is really interesting work, but what happens to the field when we have structural grammars for all languages?” These beliefs made sense within the prevailing framework, as did the widely-held “Boasian” conception articulated by theoretical linguist Martin Joos that languages can “differ from each other without limit and in unpredictable ways,” so that the study of each language must be approached “without any preexistent scheme of what a language must be.”<sup>6</sup>

These beliefs collapsed as soon as the first efforts to construct generative grammars were undertaken by mid-20<sup>th</sup> century. It quickly became clear that very little was known about human language, even the languages that had been well studied. It also became clear that many of the

fundamental properties of language that were unearthed must derive in substantial part from the innate language faculty, since they are acquired with little or no evidence. Hence there must be sharp and determinate limits to what a language can be. Furthermore, many of the properties that were revealed with the first efforts to construct rules satisfying the Basic Principle posed serious puzzles, some still alive today, along with many new ones that continue to be unearthed.

In this framework, the study of a specific language need not rely just on the behavior and products of speakers of this language. It can also draw from conclusions about other languages, from neuroscience and psychology, from genetics, in fact from any source of evidence, much like science generally, liberating the inquiry from the narrow constraints imposed by strict structuralist/behavioral science approaches.

In the early days of the generative enterprise, it seemed necessary to attribute great complexity to UG in order to capture the empirical phenomena of languages. It was always understood, however, that this cannot be correct. UG must meet the condition of evolvability, and the more complex its assumed character, the greater the burden on some future account of how it might have evolved – a very heavy burden in the light of the few available facts about evolution of the faculty of language, as just indicated.

From the earliest days, there were efforts to reduce the assumed complexity of UG while maintaining, and often extending, its empirical coverage. And over the years there have been significant steps in this direction. By the early 1990s it seemed to a number of researchers that it might be possible to approach the problems in a new way: by constructing an “ideal solution” and asking how closely it can be approximated by careful analysis of apparently recalcitrant data, an approach that has been called “the minimalist program.” The notion “ideal solution” is not precisely determined a priori, but we have a grasp of enough of its properties for the program to be pursued constructively.<sup>7</sup>

I-languages are computational systems, and ideally should meet conditions of Minimal Computation MC, which are to a significant extent well understood. I-languages should furthermore be based on operations that are minimally complex. The challenges facing this program are naturally very demanding ones, but there has been encouraging progress in meeting them, though vast empirical domains remain to be explored.

The natural starting point in this endeavor is to ask what is the simplest computational operation that would satisfy the Basic Property. The answer is quite clear. Every unbounded computational system includes, in some form, an operation that selects two objects X and Y already constructed, and forms a new object Z. In the simplest and hence optimal case, X and Y are not modified in this operation, and no new properties are introduced (in particular, order). Accordingly, the operation is simple set-formation:  $Z = \{X, Y\}$ . The operation is called *Merge*

in recent literature.

Every computational procedure must have a set of *atoms* that initiate the computation – but like the atoms of chemistry, may be analyzed by other systems of language. The atoms are the minimal meaning-bearing elements of the lexicon, mostly word-like but of course not words. Merge must have access to these, and since it is a recursive operation, it must also apply to syntactic objects SO constructed from these, to the new SOs formed by this application, etc., without limit. Furthermore, to satisfy the Basic Property some of the SOs created by Merge must be mapped by fixed procedures to the SM and CI interfaces.

By simple logic, there are two cases of Merge(X,Y). Either Y is distinct from X (*External Merge EM*) or one of the two (say Y) is a part of the other that has already been generated (*Internal Merge IM*). In both cases,  $\text{Merge}(X,Y) = \{X,Y\}$ , by definition. In the case of IM, with Y a part of X,  $\text{Merge}(X,Y) = \{X,Y\}$  contains two *copies* of Y, one the SO that is merged and the other the one that remains in X. For example, EM takes the SOs *read and books* (actually, the SOs underlying them, but let us skip this refinement for simplicity of exposition) and forms the new SO  $\{\textit{read, books}\}$  (unordered). IM takes the SOs *John will read which book* and *which book* and forms  $\{\textit{which book, John will read which book}\}$ .

In both cases, other rules convert the SOs to the SM and CI forms. Mapping to CI is straightforward in both cases. The IM example has (roughly) the form “for which x, x a book, John will read the book x.” Mapping to SM adds linear order, prosody, and detailed phonetic properties, and in the IM example deletes the lower copy of *which book*, yielding *which book John will read*. This SO can appear either unchanged, as in *guess [which book John will read]*, or with a raising rule of a type familiar in many languages, yielding *which book will John read*.

It is important to note that throughout, the operations described satisfy MC. That includes the deletion operation in the mapping to SM, which sharply reduces the computational and articulatory load in externalizing the Merge-generated SO. To put it loosely, what reaches the mind has the right semantic form, but what reaches the ear has gaps that have to be filled by the hearer. These “filler-gap” problems pose significant complications for parsing/perception. In such cases, I-language is “well-designed” for thought but poses difficulties for language use, an important observation that in fact generalizes quite widely and might turn out to be exceptionless, when the question arises.

Note that what reaches the mind lacks order, while what reaches the ear is ordered. Linear order, then, should not enter into the syntactic-semantic computation. Rather, it is imposed by externalization, presumably as a reflex of properties of the SM system, which requires linearization: we cannot speak in parallel or articulate structures. For many simple cases, this seems accurate: thus there is no difference in the interpretation of verb-object constructions in

head-initial or head-final constructions.

The same is true in more complex cases, including “exotic” structures that are particularly interesting because they rarely occur but are understood in a determinate way, for example, parasitic gap constructions. The “real gap” RG (which cannot be filled) may either precede or follow the “parasitic gap” PG (which can be filled), but cannot be in a dominant (c-command) structural relation to the PG, as illustrated in the following:

- (1) Guess who [[your interest in PG] clearly appeals to RG
- (2) Who did you [talk to RG [without recognizing PG].
- (3) \*Guess who [GAP [admires [NP your interest in GAP]]

Crucially, grammatical status and semantic interpretation are determined by structural hierarchy while linear order is irrelevant, much as in the case of verb-initial versus verb-final. And all of this is known by the language user even though evidence for language acquisition is minuscule or entirely non-existent.

The general property of language illustrated by these cases is that linguistic rules are invariably *structure-dependent*. The principle is so strong that when there is a conflict between the computationally simple property of minimal linear distance and the far more complex computational property of minimal structural distance, the latter is always selected. That is an important and puzzling fact, which was observed when early efforts to construct generative grammars were undertaken. On the surface, it seems to conflict with the quite natural and generally operative principles of MC.

To illustrate, consider the following sentences:

- (4) Birds that fly instinctively swim
- (5) The desire to fly instinctively appeals to children
- (6) Instinctively, birds that fly swim
- (7) Instinctively, the desire to fly appeals to children.

The structures of (6) and (7) are, roughly, as indicated by bracketing in (6') and (7') respectively:

- (6') Instinctively, [birds that fly] [swim]]
- (7') Instinctively, [[the desire to fly] [appeals [to children]]]

In both cases, “fly” is the closest verb to “instinctively” in linear distance, but the more remote in structural distance.

Examples (4) and (5) are ambiguous (“fly instinctively”, “instinctively swim/appeal”), but in (6’) and (7’) the adverb is construed only with the remote verb. The immediate question is why the ambiguity disappears; and more puzzling, why is it resolved in terms of the computationally complex operation of locating the *structurally* closest verb rather than the much simpler operation of locating the *linearly* closest verb? The property holds of all relevant constructions in all languages, in apparent conflict with MC. Furthermore, the knowledge is once again acquired without relevant evidence.

There have been many attempts by linguists and other cognitive scientists to show that these outcomes can be determined by some kind of learning mechanism from available data. All fail, irremediably, which is not surprising, as the simple example just given indicates.<sup>8</sup>

There is a very simple and quite natural solution to the puzzle, the only one known: languages are optimally designed, based on the simplest computational operation, Merge, which is order-free. Equipped with just this information, the child acquiring language never considers linear order in determining how to account for the data of experience; this is the only option open to the I-language satisfying the Basic Property, given that the general architecture of language established by UG satisfies MC.

The strongest thesis we can formulate about human language is that MC holds quite generally: the Strong Minimalist Thesis SMT. Not many years ago it would have appeared to be so absurd that it was never even contemplated. In recent years, evidence has been mounting that suggests otherwise.

Assuming SMT we at once have an explanation for a number of quite puzzling phenomena. One is the ubiquity of displacement - interpretation of a phrase where it appears as well as in another position in which its basic semantic role is determined, a phenomenon that appeared to require mechanisms that are an “imperfection” of language design. Under SMT, as we have seen, displacement should be the norm (namely, by IM) and would have to be blocked by some arbitrary stipulation. Hence a barrier against displacement would be an “imperfection.” Correspondingly, any approach to language that bars IM has a double burden of justification: for the stipulation blocking the simplest case, and for any new mechanisms introduced to account for the phenomena. Furthermore, SMT yields at once the “copy theory” of displacement, which provides an appropriate structure for interpretation at CI, dispensing with complex operations of “reconstruction.” And as we have just seen, it provides a solution to the puzzles of structure-dependence, an overarching principle of language design. These are, I think, quite significant results, of a kind hitherto not attained or even contemplated in the rich tradition of linguistic inquiry.

Further inquiries, which I cannot review here, carry the account further.

Note again that while MC yields appropriate structures at CI, it poses difficulties at SM. Looking further, there is substantial evidence that externalization to SM is the primary locus of the complexity, variability, and mutability of language, and that, correspondingly, mastering the specific mode of externalization is the main task of language acquisition: mastering the phonetics and phonology of the language, its morphology, and its lexical idiosyncracies (including what is called “Saussurean arbitrariness,” the specific choice of sound-meaning correspondences for minimal word-like elements).

We might proceed to entertain another bold but not implausible thesis: that generation of CI – narrow syntax and construal/interpretive rules – is uniform among languages, or nearly so. In fact, realistic alternatives are not easy to imagine, in the light of the fact that the systems are acquired on the basis of little or no evidence, as even the few simple examples given earlier illustrate. The conclusion also comports well with the very few known facts about origin of language. These appear to place the emergence of language within a time frame that is very brief in evolutionary time, hardly more than an instant, and with no evolutionary change since. Hence we would expect what evolved – UG – to be quite simple.

Note that to be seriously considered, any speculation about origin of language must account for the emergence of the Basic Property, which cannot be approached in small steps, just as evolution of the arithmetical capacity is not facilitated by ability to deal with small numbers. The leap from 4 or 1 million to an unbounded number system is no easier than the leap from 1 to that goal.

The considerations just reviewed are a sample of those that suggest that the Basic Property is not quite as formulated so far in this discussion. Rather, the Basic Property of I-language, determined by UG, is a finitely-specified generative procedure, represented in the brain, that yields a discrete infinity of hierarchically structured expressions, each with a determinate interpretation at the CI interface. Ancillary principles may externalize the internally-generated expression in one or another sensory modality.

There is neurological and psycholinguistic evidence to support these conclusions about the architecture of language and the ancillary character of externalization. Research conducted in Milan a decade ago, initiated by Andrea Moro, showed that nonsense systems keeping to UG principles of structure-dependence elicit normal activation in the language areas of the brain, but much simpler systems using linear order in violation of UG yield diffuse activation, implying that subjects are treating them as a puzzle, not a language. There is confirming evidence by Neil Smith and Ianthi-Maria Tsimpli in their investigation of a cognitively deficient but linguistically gifted subject; he was able to master the nonsense language satisfying structure-dependence, but not the one using the simpler computation involving linear order. Smith and Tsimpli also made the interesting observation that normals can solve the problem of the UG-



violating language if it is presented to them as a puzzle, but not if it is presented as a language, presumably activating the language faculty. These studies suggest very intriguing paths that can be pursued in neuroscience and experimental psycholinguistics.<sup>9</sup>

Note that these conclusions about language architecture undermine a conventional contemporary doctrine that language is primarily a system of communication, and presumably evolved from simpler communication systems. If, as the evidence strongly indicates, even externalization is an ancillary property of language, then specific uses of externalized language, as in communication, are an even more peripheral phenomenon – a conclusion also supported by other evidence, I think. Language appears to be primarily an instrument of thought, much in accord with the spirit of the tradition. There is no reason to suppose that it evolved as a system of communication.

A reasonable surmise today, I think, is that within the narrow time frame suggested by the available facts, some small rewiring of the brain yielded the modified Basic Property – of course in an individual, who was then uniquely capable of thought: reflection, planning, inference, and so on, in principle without bounds. It is possible that other capacities, notably arithmetical competence, are by-products. In the absence of external pressures, the Basic Property should be optimal, as determined by laws of nature, notably MC, satisfying SMT, rather as a snowflake takes its intricate shape. The mutation might proliferate to further generations, possibly coming to dominate a small breeding group. At that point, externalization would be valuable. The task would be to map the products of the internal system satisfying SMT to sensorimotor systems that had been present for hundreds of thousands of years, in some cases far more; thus there is evidence that the auditory systems of apes are very much like those of humans. That mapping poses a hard cognitive problem, which can be solved in many ways, each of them complex, all of them mutable – very much what we observe, as noted earlier. Carrying out these tasks might involve little or no evolutionary change.

If these reflections are on the right track, then the primary task for linguistic research is to fill in the huge gaps in this picture, that is, to show that the vast array of phenomena in the humanly accessible languages can be explained properly in something like these terms. And remaining in the domain of mystery for the moment – perhaps forever – is the origin of the atoms of computation and the nature of the “puppeteer,” the creative aspect of language use that was the prime concern of the long and rich tradition that has been revived in a new form in the generative/biolinguistic enterprise.

I remarked earlier that in my student days in mid-20<sup>th</sup> century, it seemed as though the major problems of the study of language had been pretty much solved and that the enterprise, though challenging, was approaching a terminal point that could be fairly clearly perceived. The picture today could not be more different.

## Notes

- 1 Tattersall (2012).
- 2 For sources, see Chomsky (1996, 2009).
- 3 Jespersen (1922).
- 4 Bizzi and Ajemian (2015).
- 5 Lenneberg (1967), Curtiss (2012).
- 6 Jos (1957).
- 7 Chomsky (1995)
- 8 For review, see Berwick et al. (2011).
- 9 Musso et al. (2003). Smith and Tsimpli (1995). Smith (1999).

## References

Berwick, R., P. Pietroski, B. Yankama, and N. Chomsky (2011). "Poverty of the Stimulus Revisited," *Cognitive Science* 35: 1–36.

Bizzi, E., and R. Ajemian (2015). "A Hard Scientific Quest: Understanding Voluntary Movements," *Daedalus* 144.1: 83-95.

Chomsky, N. (1966). *Cartesian Linguistics* (Harper & Row). Third edition (2009, revised, edited with an introduction by James McGilvray; ebooks.) Cambridge: Cambridge University.

Chomsky, N. (1995). *The Minimalist Program*, Cambridge, Mass: MIT Press.

Curtiss, S. (2012). "Revisiting Modularity: Using Language as a Window to the Mind," in R. Berwick and M. Piattelli-Palmarini, eds. *Rich Languages from Poor Inputs*. Oxford: Oxford University.

Jespersen, O. (1922). *Philosophy of Grammar*. Chicago: Chicago University Press.

Joos, M., ed. (1957). *Readings in Linguistics*. Washington: American Council of Learned Societies.

Lenneberg, E. (1967). *Biological Foundations of Language*. New York: John Wiley and Sons.

Musso, M., A. Moro, V. Gluache, M. Rijntjes, J. C. Bücheli, and C. Weiller (2003). "Broca's Area and the Language Instinct," *Nature Neuroscience* 6. 774–781.

Smith, N. (1999). *Chomsky: Ideas and Ideals*. Cambridge: Cambridge University Press.

Smith, N., and I-M. Tsimpli (1995). *The Mind of a Savant: Language Learning and Modularity*. Oxford: Blackwell.

Tattersall, I. (2012). *Masters of the Planet*. London: Palgrave Macmillan.

