15

UNIVERSAL GRAMMAR

Ian G. Roberts, Jeffrey Watumull, and Noam Chomsky

Introduction

Is language universal? In particular, is the grammar – the computational system in the mind/brain – that powers human language universal? Could it be universal in the way laws of nature are universal, such that any sufficiently intelligent system would all but inevitably converge on it (in either its evolution or its science)? Here we will expound on the conjecture that the answers to these questions are affirmative: grammar – particularly human grammar – is not specific to our species, but universal in the deepest of senses. The implications for xenolinguistics are obviously profound: we should predict that any extraterrestrial intelligence (ETI) – indeed any sufficiently intelligent system (e.g., artificial intelligence) – we encounter would likely be endowed with a cognitive computational system that runs human-style linguistic "software", thus eliminating any principled limit to effective communication. However, there could exist be material differences in the "hardware" used to physically externalize linguistic information, but these would pose mere engineering challenges rather than insoluble conceptual problems. It is not unreasonable to suppose that the combined intelligence of humans and ETIs could construct the necessary interface(s). In that case, effective communication between these "universal minds" would be guaranteed.

The human mind, Descartes argued, is undoubtedly in some sense a "universal instrument" (Descartes 1637 [1995]). We cannot know with certainty what he intended by this provocative comment, but we do know that the Cartesians would have understood language as fundamental to any nontrivial notion of "universality" and "intelligence" because it is language that empowers humans to generate an unbounded set of hierarchically structured expressions that can enter into (or in fact *constitute*) effectively infinitely many thoughts and actions – that is, the

DOI: 10.4324/9781003352174-15

competence of every human, but no beast or machine, to use language in creative ways appropriate to situations but not caused by them, and to formulate and express these thoughts coherently and without bound, perhaps "incited or inclined" to speak in particular ways by internal and external circumstances but not "compelled" to do so. This linguistic competence (and especially its creative use), in concert with other mental faculties, establishes the general intelligence necessary for the evolutionary "great leap forward" of our species (see Chomsky 2016):

> there might have been a crucial mutation in human evolution which led, in almost no time from an evolutionary perspective, from [humans living in] caves to [their creating knowledge of such sophistication as to enable us to imagine and construct things as complex as, say,] spaceships. It's a plausible speculation that the mutation in question was whatever it is that makes our brains capable of computing recursive syntax, since it's the recursive syntax that really gives language – and thought – their unlimited expressive power. It's one small step from syntax to spaceships, but a great leap for humans.
>
> *(Roberts 2017: 182)*

A great leap for humans – and, on Earth, *only* humans, evidently (see Berwick and Chomsky 2016).

Moreover, it has been assumed that the essential properties of human language are not only unique, but *logically contingent*:

> Let us define "universal grammar" (UG) as the system of principles, conditions, and rules that are elements or properties of all human languages not merely by accident but by necessity – of course, I mean biological, not logical necessity. Thus UG can be taken as expressing "the essence of human language."
>
> *(Chomsky 1975: 29)*

> There is no *a priori* reason to expect that human language will have such properties; Martian could be different.
>
> *(Chomsky 2000: 16)*

This assumption, we submit, merits rethinking in light of progress in the Minimalist Program (Chomsky 1995). Recent work demonstrating the *simplicity* (Watumull et al. 2017) and *optimality* (Chomsky et al. 2017) of language increases the cogency of the following: "the basic principles of language are formulated in terms of notions drawn from the domain of (virtual) conceptual necessity", the domain defined by "general considerations of conceptual naturalness that have some independent plausibility, namely, simplicity, economy, symmetry, nonredundancy, and the like" (Chomsky 1995: 171, 1) that render linguistic computation optimal. To the extent that this *strong minimalist thesis* (SMT) is true, the essential – computational –

properties of language would derive from laws of nature – language- and even biology-independent principles that, once realized in the mind/brain, *do* entail particular properties as logically necessary. For instance, it is simply a fact of logic that the simplest (optimal) form of the recursive procedure generative of syntactic structures, *Merge*, has two and only two forms of application (i.e., "external" and "internal" in the sense of combining two separate objects or two where one is inside the other). Relatedly, *given* the nature of the structures Merge generates, minimal structural distance is *necessarily* the simplest computation for the structure dependence of rules. And so on and so forth (see Berwick et al. 2011; Chomsky 2013; Watumull 2015 for additional examples).

Research in the Minimalist Program starts with the optimality conjecture and proceeds to inquire whether and to what extent it can be sustained given the observed complexities and variety of natural languages. If a gap is discovered, the task is to inquire whether the data can be reinterpreted, or whether principles of simplicity and optimal computation can be reformulated, so as to solve the puzzles within the framework of SMT, thus generating some support, in an interesting and unexpected domain, for Galileo's precept that nature is simple and it is the task of the scientist to prove it.

As we discover more and more of "the essence of human language" to be defined by (virtual) conceptual necessity, the less and less absurd it is to question just how contingent a phenomenon human language really is. It may well be with language as with other phenomena studied in the natural sciences that "[b]ehind it all is surely an idea so simple, so beautiful, that when we grasp it – in a decade, a century, or a millennium – we will all say to each other, how could it have been otherwise?" (Wheeler 1986: 304). In other words, there may well be some a priori reasons to expect human language to have the properties it does; so the ETI's language might *not* be so different from human language, after all.

Simplicity, Universality, and Merge

Our conjecture is based on the notion of *simplicity* as originally conceived in generative linguistics. "[S]implicity, economy, compactness, etc." were proffered in the first work on generative grammar as criteria the grammar of a language must satisfy:

> Such considerations are in general not trivial or "merely esthetic". It has been recognized of philosophical systems, and it is, I think, no less true of grammatical systems, that the motives behind the demand for economy are in many ways the same as those behind the demand that there be a system at all.
>
> *(Chomsky 1951: 1, 67)*

The idea is elementary but profound: if the theory is no more simple, economical, compact, etc., than the data it is proffered to explain, it is not a theory at all; hence, the more compressed the theory, the more successful – i.e., the more explanatory – it is. Incidentally, this idea is appreciated surprisingly seldom today: many computational cognitive scientists and machine learning theorists (and hence virtually all "artificial intelligence" [AI] labs in academia and industry) have perversely redefined a successful theory or computer program to be one that merely approximates or classifies unanalyzed data.[1] This contrasts dramatically with the Enlightenment definition in which data are selectively analyzed as evidence for/against conjectured *explanations*.

In generative grammar since approximately 1980 (principles-and-parameters [P&P] theory), language acquisition has been explained as the process of setting the values for the finitely many universal parameters of the initial state of the language faculty (universal grammar, UG). The apparent complexity and diversity of linguistic phenomena are illusory and epiphenomenal, emerging from the interaction of invariant principles under varying conditions. This was a radical shift from the early work in generative linguistics, which sought only an evaluation measure that would select among alternative grammars – the simplest congruent with the format encoded in UG and consistent with the primary linguistic data. But with the P&P shift in perspective, simplicity can be reformulated. As discussed in the earliest work in generative linguistics, notions of simplicity assume two distinct forms: the imprecise but profound notion of simplicity that enters into rational inquiry generally, and the theory-internal measure of simplicity that selects among grammars (i-languages). The former notion of simplicity is language-independent, but the theory-internal notion is a component of UG, a subcomponent of the procedure for determining the relation between experience and i-language. In early work, the internal notion was implemented in the form of the evaluation procedure to select among proposed grammars/i-languages consistent with the UG format for rule systems, but the P&P approach transcends that limited, parochial conception of simplicity: with no evaluation procedure, there is no internal notion of simplicity in the earlier sense. There remains only the universal notion.

In P&P, grammars – i-languages – are simple, but they are so by virtue of objective principles of computational efficiency (Chomsky 2005), not by stipulation in UG. In fact, rather than "simple", we propose to define P&P-style acquisition as "economical", which, in the Leibnizian spirit, we understand to subsume simplicity:

> The most economical idea, like the most economical engine, is the one that accomplishes most by using least. Simplicity – or fuel consumption – is a different factor from power [i.e., generative capacity, empirical coverage, etc.] but has to be taken equally into consideration [. . .]. The economy of a basis may be said to be the ratio of its *strength* to its simplicity. But superfluous power is also a waste. Adequacy for a given system is the only relevant factor in the power of a basis; and where we are comparing several alternative bases for some one system, as is normally the case, that factor is a constant. Thus in practice the simplest basis is the most economical.
>
> *(Goodman 1943: 111)*

Economy, in other words, is a *minimax* property. In Leibniz's words (see Roberts and Watumull 2015): "the simplicity of the means counterbalances the richness of the effects" so that in nature "the maximum effect [is] produced by the simplest means". This is the Galilean ideal (see Chomsky 2002).

The maximally economical form of P&P-style learning explicable in terms of objective factors is the traversal of a parameter hierarchy for parameter specification (see Roberts 2012, 2019). In such a system, the child is not enumerating and evaluating grammars.[2] Instead, the i-language matures to a steady state in a relatively deterministic process of "answering questions" that *emerge naturally and necessarily* in the sense that there exist "choices" in acquisition that logically must be "made" for the system to function at all; none of the parameters need be encoded in the genetic endowment (see Obata et al. 2015 for similar ideas). This is the ideal, of course. Like SMT generally, how closely it can be approximated is an empirical matter, and there remain many challenges.

Parameter specification – i.e., the P&P conception of "learning" as the specification of values for the variables in i-language – can be schematized as a decision tree (parameter hierarchy) which is governed by minimax economy: minimizing formal features (feature-economy) coupled with maximizing accessible features (input-generalization). Traversal of a hierarchy – a conditional-branching Turing machine program – is inevitably economical in that the shortest (in binary) and most general parameter settings are necessarily "preferred" in the sense that the sooner the computation halts, the shorter the parameter settings. For instance, to specify word order, a series of binary queries with answers of increasing length and decreasing generality (microparameters) is structured thus:

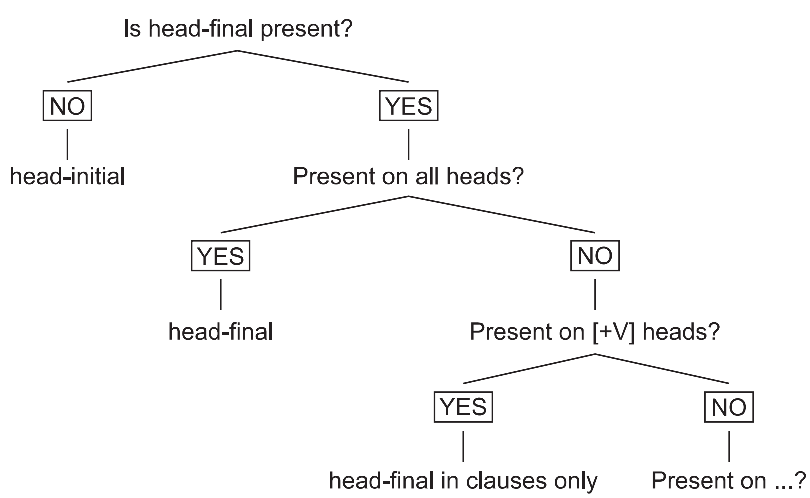


FIGURE 15.1 Parameter Hierachy

For compatibility with computability theory and Boolean logic, the parameter hierarchy can be translated as follows:

Hierarchy: *H*
State *T*: Decision Problem
Yes: 0/1 (0 = transition to state $T + 1$) (1 = halt and output parameter specification for *H*)
No: 0/1 (0 = transition to state $T + 1$) (1 = halt and output parameter specification for *H*)

Hierarchy: Word Order

State 1: Is head-final present?
Yes: Output 0 (transition to State 2)
No: Output 1 (halt and output "head-initial")

State 2: Present on all heads?
Yes: Output 1 (halt and output "head-final")
No: Output 0 (transition to State 3)

State 3: Present on [+V] heads?
Yes: Output 1 (halt and output "head-final in clause only")
No: Output 0 (transition to State 4)

So in P&P, the logic is not "enumerate and evaluate" with stipulative (theory-internal) simplicity measures; it is "compute all and only what is necessary", which implies the language-independent reality of economy in that, as with the parameter hierarchies, the process answers all and only the questions it needs to. It is not that there is any explicit instruction in the genetic endowment to prefer simple answers: it is simply otiose and meaningless to answer unasked questions (i.e., once the parameters are set, the computation halts). ETI grammars, too, might therefore grow along the lines of the P&P method.

Moreover, the "answers" to "questions" can be represented in binary. Indeed, binary is a *notation-independent* notion necessary and sufficient to *maximize* computation with *minimal* complexity (hence, it is optimal for terrestrial and extraterrestrial computation): functions of arbitrarily many arguments can be realized by the composition of binary (but not unary) functions – a truth of minimax logic with "far-reaching significance for our understanding of the functional architecture of the brain" (Gallistel and King 2010: x) – for the physical realization of intelligence anywhere in the universe. The mathematical and computational import of binary was rendered explicit in the theories of Turing (1936) and Shannon (1948), the former demonstrating the necessarily digital – hence, ultimately binary – nature of *universal computation* (a universal Turing machine being the most general mathematical characterization of computation); the latter formalizing *information* in terms of *bits* (binary digits). The consilience of these ideas is our Economy Thesis: human language is based on simple representations (i.e., bits) and strong computations (i.e., the binary functions of Turing machines) – and the "economy of a basis may be said to be the ratio of its *strength* to its simplicity" (Goodman 1943: 111, emphasis in original).

As one of the "general considerations of conceptual naturalness that have some independent plausibility", economy would be a factor that obtains of any optimally "designed" (natural or artificial) computational system. In terms of universality, if the ETI language were optimal in the sense of conforming to virtual conceptual necessity, then it might be surprisingly similar to human language. In point of fact, we ought not to be too surprised. It is now well established by biologists that *convergence* is a common theme in any evolutionary process: "the number of evolutionary end-points is limited: by no means is everything possible. [Because of evolutionary convergence,] what is possible usually has been arrived at multiple times, meaning that the emergence of the various biological properties is effectively inevitable" (Conway Morris 2013: xii–xiii); indeed, the distinguished Cambridge paleontologist Simon Conway Morris argues that human-style intelligence was effectively inevitable, given the initial conditions of evolution on Earth. And there is no reason a priori to assume that the principle of evolutionary convergence is unique to the biology of a particular planet. Quite the contrary, if we accept the rational form of inquiry in which the principle is understood abstractly in a computational framework. The idea is that *any* computational system *anywhere* made of *anything* is governed by *laws* of computation (see Gallistel and King 2010: 167). Given this universality of the functional, mathematical architecture of computation, it is possible that we may need to rethink how uniquely human or even uniquely biological our modes of mental computation really are. One interesting implication is that we must rethink any presumptions that extraterrestrial intelligence or artificial intelligence would really be all that different from human intelligence.

So we assume that human language is a computational process that can be characterized by a Turing machine (see Watumull 2015). It is possible to explore the space of all possible Turing machines (i.e., the space of all possible computer programs), not exhaustively of course, but with sufficient breadth and depth to make some profound discoveries. Marvin Minsky and Daniel Bobrow once enumerated and ran some thousands of the simplest Turing machines (computer programs with minimal numbers of rules) and discovered, intriguingly, that out of the infinity of possible behaviors, only a surprisingly – and intriguingly – small subset emerged (Minsky 1985). These divided into the trivial and the nontrivial. The boring programs either halted immediately or erased the input data or looped indefinitely or engaged in some similar silliness. The remainder, however, were singularly interesting: *all* of these programs executed an effectively *identical* counting function – a primitive of elementary arithmetic. In fact, this operation reduces to a form of

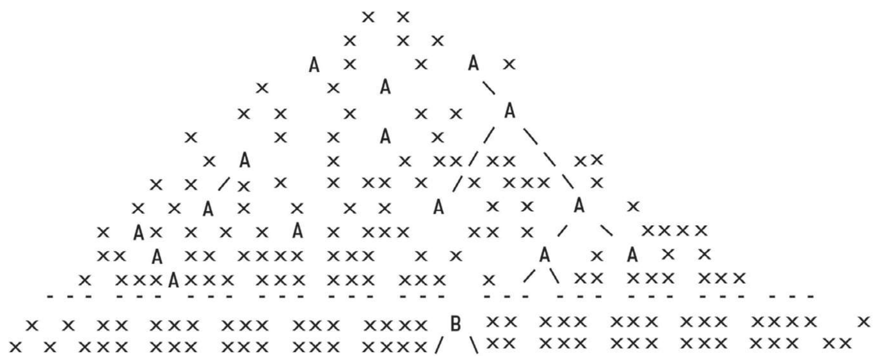


FIGURE 15.2 Sampling of the universe of possible Turing machines

Merge (see Chomsky 2008). More generally, these "A-machines" (*A* for *arithmetic*) prove a point:

> [I]t seems inevitable that, somewhere, in a growing mind some A-machines must come to be. Now, possibly, there are other, really different ways to count. So there may appear, much, much later, some of what we represent as "B-machines" – which are processes that act in ways which are similar, but not identical to, how the A-machines behave. But, our experiment hints that even the very simplest possible B-machine will be so much more complicated that it is unlikely that any brain would discover one before it first found many A-machines.
>
> *(Minsky 1985: 121)*

This is evidence that arithmetic, as represented in an A-machine, is an *attractor* in the *phase space* of possible mathematical structures (Figure 15.2):

> any entity who searches through the simplest processes will soon find fragments which do not merely resemble arithmetic but *are* arithmetic. It is not a matter of inventiveness or imagination, only a fact about the geography of the universe of computation.
>
> *(Minsky 1985: 122, emphasis added)*

This thesis obviously generalizes beyond arithmetic to all "simple" computations (see Wolfram 2002 for countless examples). "Because of this, we can expect certain 'a priori' structures to appear, almost always, whenever a computational system evolves by selection from a universe of possible processes" (Minsky 1985: 119). Analogously, we submit that it is not implausible that an evolutionary search through the simplest computations will soon find something like Merge. Merge is an operation so elementary as to be subsumed somehow in every more complex computational procedure: take two objects X and Y already constructed and form the object Z without modifying X or Y, or imposing any additional structure on them: thus Merge (X, Y) = {X, Y}. This simple assumption suffices to derive in a principled (necessary) way a complex array of otherwise arbitrary (contingent) phenomena such as the asymmetry of the conceptual-intentional and sensory-motor interfaces (entailing the locus of surface complexity and variety), the ubiquity of dislocation, structure dependence, minimal structural distance for anaphoric and other construals, and the difference between what reaches the mind for semantic interpretation and what reaches the apparatus of articulation and perception (see Chomsky 2017).

As implied by our Economy Thesis, simplicity can be defined in algorithmic information theory (or the theory of program-size complexity): the complexity of a program is measured by its maximally compressed length in bits so that the simplest program is that with the shortest description. A search of the phase space of possible programs, whether conducted consciously (e.g., by us, ETIs, etc.) or unconsciously (e.g., by modern computers, evolution, etc.), automatically proceeds in size order from the shortest and increasing to programs no shorter than their outputs (these incompressible programs are effectively lists); many complex programs would subsume simpler programs as the real numbers subsume the natural numbers. And, as demonstrated logically and empirically, "any evolutionary process must first consider relatively simple systems, and thus discover the same, isolated, islands of efficiency" (Minsky 1985: 122). Thus it may well be that, given the universal and invariant laws of evolution, convergence on systems – Turing machines – virtually identical to those "discovered" in our evolutionary history is inevitable.[3] Hence our questioning the proposition "Martian could be different".

The fact that simple computations are attractors in the phase space of possible computations goes some way to explaining why language should be optimally designed (insofar as SMT holds) in that an evolutionary search is likely to converge on it, which leads us to consideration of the origin of language.[4] The evolution of language is mysterious (see Hauser et al. 2014), but SMT is consistent with the limited archaeological evidence that does exist on the emergence of language, evidently quite recently and suddenly in the evolutionary time frame (see Tattersall 2012). Furthermore, there is compelling evidence for SMT in the design of language itself. For instance, it is a universal truth of natural language that the rules of syntax/semantics are structure-dependent (see Berwick et al. 2011): hierarchy, not linearity, is determinative in the application of rules and interpretation of expressions. Thus, the most reasonable speculation today – and one that opens productive lines of research – is that from some simple rewiring of the brain, Merge emerged, naturally in its simplest form, providing the basis for unbounded and creative thought – the "great leap forward" evidenced in the archaeological record and in the remarkable differences distinguishing modern humans from their predecessors and the rest of the animal kingdom (see Huybregts 2017; Berwick and Chomsky 2016 for in-depth discussion of these topics).

If this conjecture can be sustained, we could answer the question why language should be optimally designed: optimality would be expected under the postulated conditions, with no selectional or other pressures operating, so the emerging system should just follow the laws of nature such as minimal computation and more "general considerations of conceptual naturalness that have some independent plausibility, namely, simplicity, economy, symmetry, nonredundancy, and the like" (Chomsky 1995: 171, 1) – quite like the way a snowflake forms. If this is correct, then, contrary to what was once presumed, there *would* be a priori reasons to expect that human language will have the properties it does; the "principles, conditions, and rules that are elements or properties of all human languages" (Chomsky 1975: 29) *would* be *logically* necessary, deriving from laws of nature. Remarkably, it could be in language, not physics, that we first discover, in Wheeler's words, "an idea so simple, so beautiful, that [. . .] we will all say to each other, how could it have been otherwise?" (Wheeler 1986: 304).

One idea in language so simple that perhaps it could not have been otherwise is Merge. As we have discussed, it is in some sense an attractor in the phase space of possible generative functions: its irreducibility and conceptual necessity render it "inevitable" in the design of any computational system. In this most elementary of forms, Merge functions as follows. Given a workspace WS of syntactic objects $\{X_i, \ldots, X_m\}$, let $\sum$ be the shortest sequence $(X_1, \ldots, X_n)$ such that X*i* is accessible and $\sum$ exhausts WS. Thus, Merge($\sum$) = $\{\{X_1, X_2\}, X_3, \ldots, X_m\}$. By this formulation, which conforms to the simplest, necessary principles of computability theory, we map WS to WS′ by taking any two accessible elements X and Y in WS, and replace X and Y with {X, Y} in WS′. It is manifest that two and only two legitimate forms of Merge follow from this formulation. (1), if X_1 is a term of X_2 (or conversely), then Merge replaces X_1 and X_2 by $\{X_1, X_2\}$: Merge($\sum$) = $\{\{X_1, X_2\}, X_3, \ldots, X_m\}$. (2) If neither X_1 nor X_2 is a term of the other, then, again, Merge($\sum$) = $\{\{X_1, X_2\}, X_3, \ldots, X_m\}$. Call (1) Internal Merge and call (2) External Merge. The outputs of (1) and (2) are identical, consistent with the Galilean ideal, "general considerations of conceptual naturalness that have some independent plausibility, namely, simplicity, economy, symmetry, nonredundancy, and the like" (Chomsky 1995: 171, 1).

The simple, legitimate form of Merge we proffer has implications for notions of universality. This "merger as replacement" formulation is formally equivalent to a 2-tag system of the general form given by Post (1943) (whose formulation of rewrite rules influenced early work on generative grammar), which is interesting because tag systems form "a class of systems with a particularly simple underlying structure" (Wolfram 2002: 93). In a tag system, given a sequence of elements, a set of rules removes a fixed number of elements from the beginning of the sequence and replaces them with a fixed number of elements to the end of the sequence.[5] For instance, consider a 1-tag system with the rewrite rules (1, . . .) => (. . . , 1, 0) and (0, . . .) => (. . . , 0, 1): If the sequence begins with a 1, remove it, and replace it with 1, 0 at the end of the sequence; If the sequence begins with a 0, remove it, and replace it with 0, 1 at the end of the sequence. Representing 1 with black and 0 with white, we can express the rewrite rules as follows (Figure 15.3).

And now we can see the complex patterns that emerge from applying such simple rules to simple initial conditions (Figure 15.4).

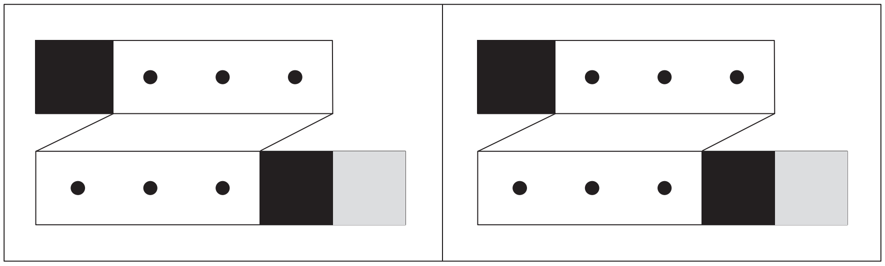


FIGURE 15.3 Merge encoded in rules of a 2-state, 3-color universal Turing machine—provably the "smallest" universal Turing machine

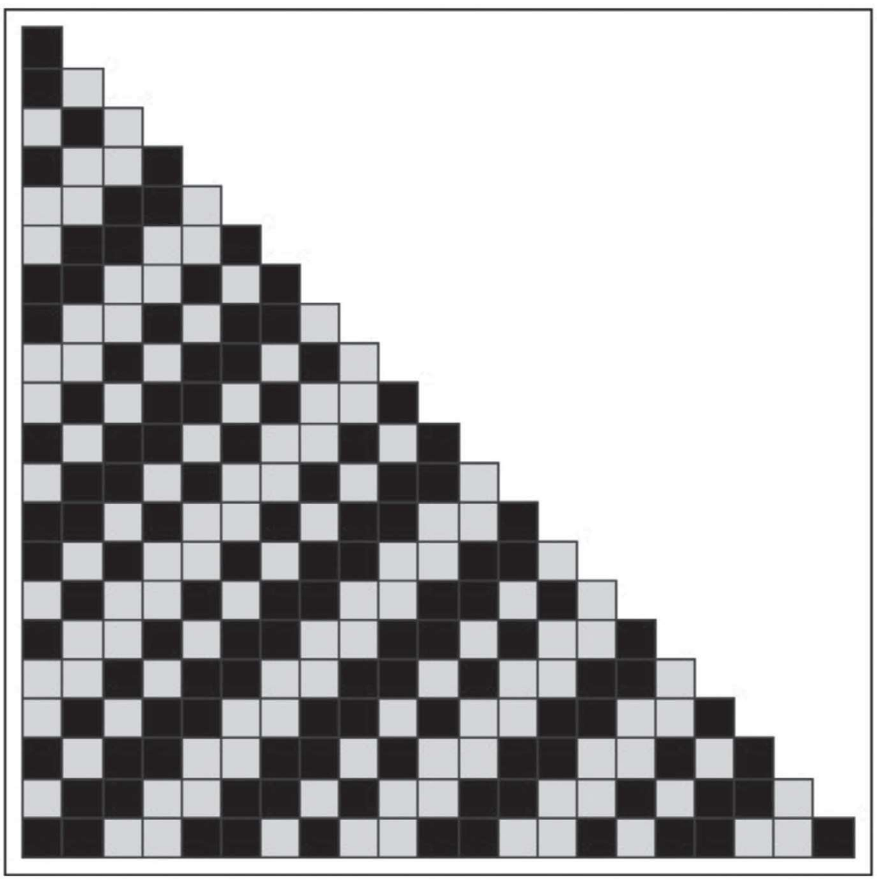


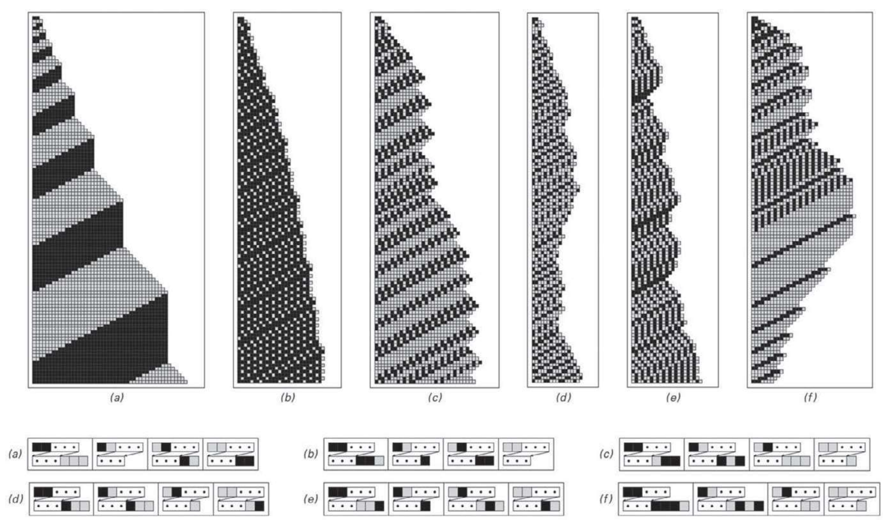
FIGURE 15.4 Merge formalized as a universal 2-tag system

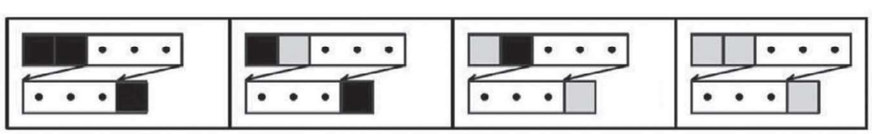
The patterns that emerge with 2-tag systems become increasingly complex (Figure 15.5).

Merge (Figure 15.6) is the simplest 2-tag system with the following rules (n.b., other permutations are possible).

What is most interesting is that this 2-tag system is provably equivalent to a *universal Turing machine* (see the proofs in Minsky 1961; Cocke and Minsky 1964; Davis 1958; Wolfram 2002; Watumull 2015): it can compute anything that is computable. Therefore, it would not be absurd to conject that, by virtue of Merge, the human mind is a universal Turing machine. However we will not defend the conjecture here – we proffer it simply in the spirit of exploration:

> It is [. . .] quite possible that we, as a species, have crossed a cognitive threshold. Our capacity to express anything, through the recursive syntax and compositional semantics of natural language, might have taken us into a cognitive realm where anything, everything is possible.
>
> *(Roberts 2017: 181–182)*





Let us suppose that the human mind is a universal instrument, a universal Turing machine. Immediately, we must answer questions of scope and limits. With respect to scope, a Turing-universal mind could arguably explain and understanding everything, in principle. The argument is simple: a universal Turing machine can emulate any other Turing machine (i.e., a universal computer can run any program); a program is a kind of theory (written to be readable/executable by a computer); thus, a universal Turing machine can compute any theory; and thus, assuming that everything in the universe could in principle be explained by and understood within some theory or other (i.e., assuming no magic, miracles, etc.), a universal Turing machine – a Turing-universal mind – could explain and understand everything. QED, perhaps. It is an intriguing conclusion, and not obviously false.

However, notwithstanding the universal logic of computation, it is obviously necessary that there exist *constraints* on the mind if it is to have any *scope* at all. If a Turing-universal mind is to be a universal explainer, it should not generate all possible explanations – true and false – because that would be merely to restate the problem of explaining Nature: deciding which in an infinite set of explanations are the true (or best) explanations is as difficult as constructing the best explanations in the first place. There must be "a limit on admissible hypotheses", in the words of Charles Sanders Peirce (Peirce 1903, see Chomsky 2006: xi). This interdependence of scope and limits has been expounded by many creative thinkers and analyzed by (creative) philosophers of aesthetics: the beauty of jazz emerges not by "playing anything", but only when the improvisation is structured, canalized; the beauty of a poem is a function of its having to satisfy the constraints of its form, as the eminent mathematician Stanislaw Ulam (1976: 180) observed

> When I was a boy I felt that the role of rhyme in poetry was to compel one to find the unobvious because of the necessity of finding a word which rhymes. This forces novel associations and almost guarantees deviations from routine chains or trains of thought. It becomes paradoxically a sort of automatic mechanism of originality.

Thus from science to art, we see that the (hypothesized) infinite creativity of the Turing-universal human mind is non-vacuous and useful and beautiful only if it operates within constraints – constraints discoverable by any (evolved) intelligence. Might this imply that, endowed with universal linguistic Turing machines (i.e., Turing-universal minds) analogous to ours, ETIs would share our sense of beauty? We think it likely.

Conclusion: Language and Mind Across the Universe

In the foregoing, we have argued that that the human language faculty is optimally designed for maximal computational simplicity and is instantiated in the fundamental operation Merge, and that the likelihood is that any recognizably intelligent entity would possess the same system. The arguments from simplicity and optimality take us close to the actual conceptual necessity of this conclusion. What does this mean in the context of potential contact?

Let us assume the existence of intelligent ETIs who have developed a technological civilization to at minimum a human-level technological sophistication such that they would be capable of conceiving the existence of civilizations alien to them, such as us, and able to contemplate sending/receiving a signal of some form. To be precise, we assume "human-level of technological sophistication" to include an understanding of fundamental mathematics (particularly computability theory), fundamental physics, whatever mechanisms (if any) enter into biological (or its equivalent) evolution in their environment, etc. The considerations we have raised in the foregoing mean that it is all but logically impossible that such a level of knowledge could be attained without a "language" for generating, storing, and communicating information. Languages are based on grammars: systems of primitives, principles, parameters, and procedures encoded in cognitive mechanisms that determine the possible structural properties of languages. Modern linguistic theory proposes the existence of a species-specific cognitive capacity, UG, that predisposes all human children (in normal environments) to acquire the language(s) to which they are exposed. The central question for xenolinguistics is the degree to which UG, as we currently understand it, would resemble ETI grammar. We have argued that the essential architecture of ETI UG must be virtually identical to that of human UG. Fundamental to the human – and ergo ETI – architecture are three axioms: (I) a "linguistic Turing machine" – a computable function – generative of an infinite set of hierarchically structured expressions that interface (at minimum) two extra-linguistic systems; (II) a conceptual-intentional system; (III) an externalization system for communicating conceptual-intentional information. We have argued that human evolution converged on (I) as a globally optimal solution to generating the conceptual structures represented in (II); there is also a system for connecting them to externalization in (III), whose details we have largely left aside here. The computational optimality/efficiency of (I) derives from a recursive structure-building operation, capable of constructing structures of unbounded complexity by iteratively applying to its own output. In human UG, this operation is Merge. The overwhelming likelihood is that ETI UG would also be based on Merge. Thus, the greatest difficulty in communicating with ETIs would be posed not by their grammar, but in understanding their externalization system; however, we submit that this is an engineering problem which should pose no difficulties in principle (although its practical nature is hard to foresee).

Ultimately, this theory of UG for humans and ETI (AI) can enkindled a more general, grander, unified theory of Life, Information, Language, and Intelligence (see Watumull and Chomsky, "To appear").[6]

Notes

1 The machine learning systems (e.g., deep learning neural networks powering large language models) so popular in the current "AI spring" are *weak AI*: brute-force systems laboriously trained to "unthinkingly" associate patterns in the input data to produce outputs that approximate those data in a process with no resemblance to human cognition (thus betraying Turing's original vision for AI). These systems will never be truly intelligent, and are to be contrasted with the *strong – anthronoetic – AI* Turing envisioned: a program designed to attain human-level competence with a *human-style* typified by *syntactic generativity* and *semantic fluidity* – to think *the way* a human thinks. See Copeland (2004) for more. Today such programs, based on generative grammars, are finally being built at Oceanit.

2 Such an inefficient and unintelligent technique is the *modus operandi* of many machine learning (weak AI) systems.

3 Indeed, we might speculate that were we to "wind the tape of life back" and play it again, in Stephen Jay Gould's phrasing, not only would something like Merge re-emerge, but something like humans could well be "inevitable", as serious biologists have suggested (see Conway Morris 2013).

4 Convergence is a consequence of constraints. As with intelligence, evolution and development are possible only by coupling scope with constraints. Stated generally, the scope of any creative process is a function of its operating within limits. In the context of evolution, for instance, Stuart Kauffman (1993: 118) observes,

> Adaptive evolution is a search process – driven by mutation, recombination, and selection – on fixed or deforming fitness landscapes. An adapting population flows over the landscape under these forces. The structure of such landscapes, smooth or rugged, governs both the evolvability of populations and the sustained fitness of their members. The structure of fitness landscapes inevitably imposes limitations on adaptive search.

The analogy to mind is deeply nontrivial, for "intellectual activity consists mainly of various kinds of search" (Turing 1948: 431).

5 This remove/replace formulation mirrors that of the original remove/replace formulation of Merge (see Chomsky 1995), which we have revised here so as not to stipulate a separate "remove" step. In our formulation, there is simply "replace".

6 www.youtube.com/watch?v=gqTyg_W_yHI

References

Berwick, Robert C., and Noam Chomsky. 2016. *Why Only Us*. Cambridge, MA: MIT Press.

Berwick, Rober C., Paul Pietroski, Berach Yankama, and Noam Chomsky. 2011. "Poverty of the stimulus revisited." *Cognitive Science* 35: 1207–1242.

Chomsky, Noam. 1951. *The Morphophonemics of Modern Hebrew*. Master's thesis. University of Pennsylvania, Philadelphia.

Chomsky, Noam. 1975. *Reflections on Language*. New York: Pantheon.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.

Chomsky, Noam. 2000. *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Chomsky, Noam. 2002. *On Nature and Language*. Cambridge: Cambridge University Press.

Chomsky, Noam. 2005. "Three factors in language design." *Linguistic Inquiry* 36: 1–22.

Chomsky, Noam. 2006. *Language and Mind*. Cambridge: Cambridge University Press.

Chomsky, Noam. 2008. "On phases." In *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, edited by Robert Freiden, Carlos P. Otero, and Maria Luisa Zubizarreta, 133–166. Cambridge: MIT Press.

Chomsky, Noam. 2013. "Problems of projection." *Lingua* 130: 33–49.

Chomsky, Noam. 2016. *What Kind of Creatures Are We?* New York: Columbia University Press.

Chomsky, Noam. 2017. "The language capacity: Architecture and evolution." *Psychonomic Bulletin and Review* 24: 200–203.

Chomsky, Noam, Ángel J. Gallego, and Dennis Ott. 2017. "Generative grammar and the faculty of language: Insights, questions, and challenges." *Catalan Journal of Linguistics*.

Cocke, John, and Marvin Minsky. 1964. "Universality of tag systems with P = 2." *Journal of the Association for Computing Machinery* 11: 15–20.

Conway Morris, Simon. 2013. *Life's Solution: Inevitable Humans in a Lonely Universe*. Cambridge: Cambridge University Press.

Copeland, Jack B., ed. 2004. *The Essential Turing*. Oxford: Clarendon Press.

Davis, Martin. 1958. *Computability and Unsolvability*. New York: McGraw-Hill Book Company.

Descartes, Rene. 1637 [1995]. *The Project Gutenberg eBook of a Discourse on Method.* Translated by John Veitch. https://www.gutenberg.org/files/59/59-h/59-h.htm

Gallistel, Charles Randy, and Adam Philip King. 2010. *Memory and the Computational Brain: Why Cognitive Science Will Revolutionize Neuroscience*. New York: Wiley-Blackwell.

Goodman, Nelson. 1943. "On the simplicity of ideas." *Journal of Symbolic Logic* 8: 107–121.

Hauser, Marc D., Charles Yang, Robert C. Berwick, Ian Tattersall, Michael J. Ryan, Jeffrey Watumull, Noam Chomsky, and Richard C. Lewontin. 2014. "The mystery of language evolution." *Frontiers in Psychology* 5.

Huybregts, Riny. 2017. "Phonemic clicks and the mapping asymmetry: How language emerged and speech developed." *Neuroscience and Biobehavioral Reviews*.

Kauffman, Stuart. 1993. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford: Oxford University Press.

Minsky, Marvin. 1961. Recursive unsolvability of Post's problem of "tag" and other topics in the Theory of Turing machines. *Annals of Mathematics* 74: 437–455.

Minsky, Marvin. 1985. "Why intelligent aliens will be intelligible." In *Extraterrestrials: Science and Intelligence*, edited by Edward Regis, 117–128. Cambridge: MIT Press.

Obata, Miki, Samuel Epstein, and Marlyse Baptista. 2015. "Can crosslinguistically variant grammars be formally identical? Third factor underspecification and the possible elimination of parameters of UG." *Lingua* 156: 1–16.

Peirce, Charles. 1903. *Collected Papers of Charles Sanders Peirce* (Volumes I–VI, edited by C. Hartshorne and P. Weiss; Volumes VII–VIII, edited by A.W. Burks). Cambridge: Harvard University Press.

Post, Emil. 1943. "Formal reductions of the general combinatorial decision problem." *American Journal of Mathematics* 65: 197–215.

Roberts, Ian G. 2012. "Macroparameters and minimalism: A programme for comparative research." In *Parameter Theory and Linguistic Change*, edited by Charlotte Galves, Sonia Cyrino, Ruth Lopes, Filomena Sândalo, and Juanito Avelar, 320–335. Oxford: Oxford University Press.

Roberts, Ian. 2017. *The Wonders of Language or How to Make Noises and Influence People*. Cambridge: Cambridge University Press.

Roberts, Ian. 2019. *Parameter Hierarchies and Universal Grammar*. Oxford: Oxford University Press.

Roberts, Ian, and Jeffrey Watumull. 2015. "Leibnizian linguistics." In *50 Years Later: Reflections on Chomsky's Aspects*, edited by Ángel J. Gallego & Dennis Ott, 211–222. Cambridge: MIT Press.

Shannon, Claude E. 1948. "A mathematical theory of communication." *Bell System Technical Journal* 27: 379–423, 623–656.

Tattersall, Ian. 2012. *Masters of the Planet: The Search for Our Human Origins*. New York: Palgrave Macmillan.

Turing, Alan. 1936. "On computable numbers, with an application to the Entschidungsproblem." *Proceedings of the London Mathematical Society* 42: 230–265.

Turning, Alan. 1948. *Intelligent Machinery*. Reproduced in Copeland (ed.) 2004. Page numbers as in Copeland.

Ulam, Stanislaw. 1976. *Adventures of a Mathematician*. New York: Scribner's.

Watumull, Jeffrey. 2015. *The Linguistic Turing Machine*. Doctoral dissertation. University of Cambridge, Cambridge.

Watumull, Jeffrey, and Noam Chomsky. "To appear." *Life, Information, Language, Intelligence*.

Wheeler, John Archibald. 1986. "How come the Quantum?" *Annals of the New York Academy of Sciences* 480: 304–316.

Wolfram, Stephen. 2002. *A New Kind of Science*. Champagne, IL: Wolfram Media.