corrected

to appear: Recursion (Proceedings of the UMass conference on recursion)

ed. T. Roeper and M. Speas [Oxford University Press]

January 2012

Minimal Recursion: Exploring the Prospects

Noam Chomsky

For our purposes, we can think of recursion as enumeration of a set of discrete objects by a computable finitary procedure, one that can be programmed for an ordinary digital computer that has access to unlimited memory and time.[1] Taking the recursive procedure P to be a function on the integers, its range R = {P(n)}, the set of objects enumerated by P. In the interesting cases, R is infinite, but it could be finite (or null). We are concerned with a special case of recursive procedures, generative grammars G_i, each of which enumerates a set of hierarchically structured expressions, assigning to each a symbolic representation at two interfaces, the sensorimotor interface SM for external realization ER and the conceptual-intentional interface CI for what is loosely termed *thought*: interpreting experience, reflection, inference, planning, imagining, etc. In this respect each G_i can be regarded as an instantiation of the traditional Aristotelian conception of language as sound with meaning (though sound is now known to be only a special case of ER).

The fact that the ER-meaning association is in principle unbounded has occasionally (but not often) been recognized explicitly in the long and rich history of study of language. Darwin wrote that "man differs [from animals] solely in his almost infinitely larger power of associating the most diversified sound and ideas," that is, in having a generative grammar G; "almost infinite" is a traditional phrase to be interpreted ambiguously as "very many" or "infinite." The unbounded character of the association is the basis for the "marvelous invention" that so entranced Galileo, the alphabet, which, in his words, provides the means to construct "from 25 or 30 sounds that infinity of expressions [that] enable us to reveal everything that we think, and all the movements of our soul." Recognition

[1] Adopting what Soare (1996) calls the Recursion Convention, as is common practice, though in his careful conceptual and historical review Soare recommends replacing "recursive" by "computable" when the general framework is Turing computability, the usual case.

of the unbounded character of the language itself, and the deeper concern for the creative character of its normal use, was soon to become a core element of Cartesian science/philosophy. A century ago Otto Jespersen raised the question of how the elements of language "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, over an unbounded range. The task of the linguist, then, is to discover these mechanisms and how they arise in the mind, and 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, marginalizing Jespersen's insights (Falk, 1992).

Reformulating Jespersen's program, the basic task is to investigate the true nature of the interfaces and the generative procedures that relate them, and to determine how they arise in the mind and are used, the focus of concern naturally being on "free expressions"; along with further questions about neural representation, evolution, and much else.

Returning to the Aristotelian formula "sound with meaning," from classical times there has been extensive study of sound in a broad sense of the term, including the entire process of *externalization* that converts the generated structured expressions to ER: phonology, morphology, prosody, etc.; and also study of the much more obscure topic of meaning. But the concept *with* has largely been ignored, partly because the tools to investigate it carefully were not readily available, partly because the matter was regarded as straightforward, just "the use of words." A common view during the cognitive revolution of the 17th-18th century was that the order of words followed the order of thought. Hence there isn't much to say about the notion *with* beyond variations of order and grouping of words into larger units. We can now interpret *with* in terms of recursion, and as we now know all too well, it is by no means an easy complex to unravel.

The modern sciences have largely adopted Galileo's methodological guideline that nature is simple, keeping to the "easiest and simplest rules," as he put it: the task of the scientist is to try to establish the principle, despite the heterogeneity and diversity of phenomena. In our special case, the task is to determine how closely language approximates minimal recursion, within the boundary conditions set by interface satisfaction. That is why a recent collection of essays has the title

Interfaces + Recursion = Language – followed by a big ?, because there is so much that is not understood.[2]

Sometimes such efforts have been considered hopeless, or perhaps completely pointless, not just for language. Not many years ago a prominent molecular biologist could plausibly conclude that the variability of organisms is so free as to constitute "a near infinitude of particulars which have to be sorted out case by case" – "near infinitude" with the usual ambiguity. In much the same vein, Martin Joos (1957) reviewed the prevailing "Boasian tradition" that languages can "differ from each other without limit and in unpredictable ways" so that each language must be approached "without any preexistent scheme of what a language must be," without preconceptions. That of course could not be literally true, but it seemed at the time a reasonable general point of view. Adopting it, linguistic theory consists of procedures of analysis of data, basically segmentation and classification. In general biology, the picture now looks radically different; in linguistics too I think, though more controversially.

From its modern revival, generative grammar largely developed within what came to be called "the biolinguistic framework," which takes the faculty of language to be a component of the individual mind/brain, much like the modules of visual perception, organization of motor action, etc. The move from analytic procedures to the biolinguistic framework greatly enriches the variety of evidence that bears on the study of each individual language to include acquisition, neuroscience, dissociations, and much else, and also what is learned from the study of other languages, on the assumption that the capacity for language relies on shared biological properties. However, it took some time to make a clear transition from the prevailing procedural framework of structural/behavioral linguistics to the biolinguistic framework – in particular, to distinguish clearly between application of procedures of analysis to organize a corpus of data, on the one hand, and on the other, inquiry into language acquisition, the cognitive processes by which the language acquired arises from the data in which the child is immersed.[3]

[2] Sauerland and Gärtner (2007).

[3] One illustration, which has become influential in recent years (in rather misleading ways, I think), is "chunking" of running text into word-like elements. In my *Logical Structure of Linguistic Theory LSLT* (1955), I took for granted that this is done by statistical analysis of transitional probabilities, modifying a proposal of Zellig Harris's for detection of morphemes (not the right units, however, because morphemes are too abstract, lacking the required sequential properties; but a necessary assumption given the principled procedural ban on level-mixing). There is no reason to suppose that such analytic procedures relate at all closely to how a child analyzes discourse, making use of a variety of information, such as recognition of familiar words. It is also now known that the procedure does not work unless supplemented by prosodic information. See Yang (2004), Shukla et al. (2011).

The extensive and very revealing studies of acquisition of language, particularly in the past 30 years, have shown that the process is no trivial matter. That is true even for the first step, not at all well understood: the extraction of language-relevant data from the "blooming buzzing confusion." There is recent evidence that the auditory system of chimpanzees interprets speech much as humans do.[4] Insofar as this is true, internal computational mechanisms must be substantially responsible even for the first step in language acquisition, as well as for later steps that lead to mature knowledge.

It seems to me well established (though much contested) that language "represents a domain-specific mental faculty, one that rests on structural organizing principles and constraints not shared in large part by other mental faculties, and in its processing and computation is automatic and mandatory," the formulation by Susan Curtiss (forthcoming), who has done pioneering work in establishing these conclusions on the basis of dissociations and other evidence. Note that it is the processing and computation that are automatic and mandatory. Use of language extends far beyond these limits. There is no reason today to doubt the fundamental insight of Descartes that use of language has a creative character: it is typically innovative without bounds, appropriate to circumstances but not caused by them – or as far as is known by internal states – and can engender thoughts in others that they recognize they could have expressed themselves. We should bear in mind as well that von Humboldt's now oft-quoted aphorism that language involves infinite use of finite means refers to *use*. There has been great progress in understanding the finite means that make possible infinite use, but the latter remains largely a mystery.

Within the biolinguistic framework, the core concept of language is I-language – "I" signifying individual, internal, and intensional, in that we are concerned with the actual computational procedure, not just its range; other notions of language are derivative, involving considerations of varied kinds.[5] We can take an I-language to

[4] Fitch (2011). As is well-known, human vocal production involves different mechanisms, but the significance of this fact is limited in the light of the use of sign, also in languages created by children without input (Goldin-Meadow and Feldman, 1977).

[5] In Chomsky (1986), I suggested that the term *I-language* be used instead of *grammar*, in one of the senses of this ambiguous term. I added that any other concept of language might be called a variety of *E-language* ("E" for external). The latter term has come to be used to refer to a (finite) corpus of material, or to some infinite object generated by the I-language, usually a set of well-formed formulas WFF. This usage is unfortunate, and should be abandoned, I think. A corpus is not a language in any sense. A set of WFFs is a *formal language*, determined by some finite generative procedure. Apart from its derivative character, it is not even clear that such a set can be coherently identified for human language; see *LSLT*,

be nothing other than the computational procedure P that yields the set of generated expressions, including the paired interface representations.

As in the case of all organic growth and development, several factors interact in acquisition of language: (1) external data, (2) genetic endowment that converts the data to experience and then guides the course of acquisition of I-language, (3) general principles of development that hold more broadly and may even be extra-organic. For a computational system like I-language, we would expect the third factor to include general conditions of computational complexity. While these are not entirely well-defined antecedently, some of their aspects are clear: e.g., less is better than more, minimal search is better than deeper search, etc. Even such simple notions carry us quite far.

The genetic factor (2) in turn has several components, among them (a) those specific to language (Universal Grammar, UG), (b) other cognitive processes, e.g., methods of statistical analysis of data, (c) determinants of structure and organization of the brain and other systems relevant to language. Exactly how these interact is the topic of the study of language acquisition.

Normal scientific rationality requires that at every stage of investigation we adopt the simplest assumptions, unless empirical evidence requires complication, thus weakening explanatory force and setting up still more barriers to an eventual account of evolution of language, if this proves to be a feasible endeavor.[6] This has always been a guiding principle in serious work on language, even if only tacit, including such traditional questions as mode of transcription and search for symmetries in structuralist phonology; and within generative grammar, among other examples elimination of stipulations and complex technology in the course of reduction of phrase structure grammar to X-bar theory and then to bare phrase structure, parallel simplification of transformational technology finally to Move-α, and then unification of the residues of phrase structure and transformational grammar to the simplest computational operation, as discussed below.

Chomsky (1957), and much since. In actual practice dimensions of deviance have been a rich and productive topic of inquiry since the early days of generative grammar.

[6] That much can ever be learned about evolution of language, or of cognition generally, is by no means obvious. For well-argued (and unfortunately, largely ignored) skepticism on the topic, see evolutionary biologist Richard Lewontin (1998). There is a huge and growing literature on evolution of language, but it seems to be almost entirely devoted to a different topic, speculations about evolution of communication; and in fact the alleged target, language, is rarely even characterized. Study of evolution of the eye can proceed only so far as we know what an eye is. The same holds for language or anything else.

In the early days of modern generative grammar, these investigations were generally framed in epistemological terms, as a search for simplicity of theory -- not to be confused with the construction of evaluation measures for I-languages, theory-internal empirical hypotheses about what counts as legitimate linguistic generalization. As the distinction between procedural analysis and the biolinguistic framework became more clear, the epistemological version was reinterpreted in metaphysical terms: as a search for what is true of cognitive systems, language in particular. In a seamless continuation of such endeavors, this search was reformulated as the minimalist program MP about 20 years ago. The primary innovation of MP was to suggest a new research paradigm, sometimes called "approaching UG from below." Pursuing this approach, we ask what the optimal solution would be for I-language, then formulating the *Strong Minimalist Thesis* SMT that holds that each language actually is an optimal solution for the interface conditions. We can then inquire into the (naturally vast) gap between the SMT and what linguistic data appear to require, asking whether the gap can be overcome by a deeper reanalysis of the data, or by reformulating the assumptions about the nature of minimal computation (and other postulated third factor principles, empirical matters that can in principle be tested independently). The MP is entirely theory-neutral; one can choose to pursue these questions or not whatever one's ideas about the nature of language and cognition. MP is not a theory; rather, a research program that falls strictly within the bounds of normal science. Despite repeated efforts at clarification, the MP has often been misinterpreted as a new theory of language. We read, even in professional literature, that the MP has been refuted, or is not minimalist enough, etc. No such critique is intelligible. Research programs may be poorly implemented, or premature, or misguided, but not refuted.

One far-reaching thesis about language acquisition, with many important consequences, is the *Continuity Hypothesis*, which implies that "children are free to try out various linguistic options (compatible with Universal Grammar) before they set these parameters in the same way as the particular human language that is spoken to them."[7] It follows that child language may (and in interesting ways does) exhibit properties different from those of the adult language, though found in other languages compatible with UG. A related view is the *Maturation Hypothesis*,

[7] Crain (forthcoming), focusing primarily on "logical nativism, which is the proposal that humans are biologically equipped with the tools for logical reasoning," a contingent property that allows for very limited variation among languages and is not derivable from assumptions of rationality. The thesis traces to Pinker (1984). A contrasting thesis is Kenneth Wexler's "*Very-Early Parameter-Setting (VEPS)*, holding that from the earliest observable ages (around 18 months), children have set their parameters correctly" (Wexler, 2002); also well-supported empirically.

holding that certain principles of UG are available only at later stages of maturation (Borer and Wexler (1987). Both have substantial empirical support and significant consequences.

As noted, that UG interacts with other cognitive processes in language acquisition has been taken for granted since the origins of modern generative grammar. There is a substantial literature on how such interaction takes place (Yang, 2002). An alternative view, widely held since the 1950s, is that other cognitive processes alone suffice to account for acquisition of language, so that language does not really exist as an object of independent study. I will put all of this aside here, simply assuming (as I think is well established) that the non-existence thesis is no more tenable today than it was in the past.[8]

A finitary computational procedure P will have buried in it in some form an operation – call it *Merge* -- that takes objects already constructed and forms from them a new object, beginning with a set of atomic objects (which may have internal structure). To first approximation, we can take the atomic objects to be lexical items drawn from the lexicon, though this is not an innocent move. We can therefore think of P as having available a *work space* consisting of the lexicon (or some subpart extracted from it for the purpose of the computation) and objects that have already been formed by P. The optimal assumption is that Merge takes the simplest form: a binary operation that applies to X, Y, forming Z = {X, Y}, with X and Y unchanged by the operation (the No-tampering Condition NTC), and also unordered.

This optimal assumption too is far from innocent, and has substantial import. Take linear order. If Merge leaves elements unordered, then it is natural to expect that linear order and other forms of arrangement are reflexes of SM, where they are required, and therefore do not enter into the derivation of the CI interface, that is into core syntax and semantics. The general architecture of P, then, would consist of a *narrow syntax* NS that yields hierarchically-structured but unordered expressions that are transferred to CI, leaving atomic elements unchanged (by virtue of NTC) and with no new elements added (by virtue of sole reliance on Merge – the inclusiveness condition INC). An ancillary procedure of

[8] For detailed critical discussion, see among others Crain, Wexler, *op. cit.* On the failure of the most careful efforts to establish the non-existence thesis, see Berwick et al. (2011); also below.

externalization EXT maps the expressions generated by NS to SM, not affecting the mapping to CI.[9]

There is a still broader thesis that is strongly suggested by these conclusions, call it the thesis T:

(T) Language is optimized relative to the CI interface alone, with externalization a secondary phenomenon.

There are other considerations tending in the same direction, conforming to a traditional view that language is primarily an instrument of thought, with other uses secondary. If so, we should revise the Aristotelian picture of language as sound with meaning; rather, it should be regarded as meaning with sound, a very different conception.

There is familiar evidence supporting a slight revision of INC: unvalued features (including structural case and ϕ-features of verbs and adjectives) are valued in the narrow syntax, a prerequisite for mapping to CI as non-deviant (see note 5). But other than that, there seems to be no strong reason to modify INC. As for ordering, the conceptual argument for restricting it to EXT seems strong, but empirical arguments have been offered to the contrary, particularly in recent work of Richard Kayne's. I will keep to the simplest assumption here, with this caveat.

A Merge-based system is compositional, in the Fregean sense: interpretation of X is a function of the interpretation of its parts and the way they are combined. There is ample evidence in all aspects of language that the principle holds quite generally, with cyclic computation in all components of grammar. A more restrictive condition is *strict cyclicity*: an object already computed can no longer be modified. This condition plainly yields substantial computational saving, and has many consequences. A specific version of the condition is *phase theory*, which designates specific stages in the narrow-syntactic computation where material is transferred to the interfaces, and thus induces limited (but insufficient) compositionality to interface computation. I leave this interesting topic aside here.[10]

[9] Note that this picture could in principle accommodate the more complex and widely explored thesis that CI considerations enter into externalization; as usually formulated, that Spell-Out accesses LF. Questions then arise about cyclicity and other matters.

[10] See Gallego (2012), for various developments of the general idea.

If Merge is binary, then Merge(X, Y) has only two possible cases: either X and Y are entirely distinct (*external Merge EM*), or one is a *term* of the other (*internal Merge IM*, where a term of W is a member of W or a term of a member of W). IM yields the ubiquitous displacement property of natural language, in the form that has been called "the copy theory of movement." If Y is a term of X, then by virtue of NTC, IM(X, Y) = {X, Y}, with two *copies* of Y, one the copy that was already part of X and the other one merged to X. For example, skipping details, if X is (1) and Y is *which boys*, then (2) has two copies of Y, one externally merged within X and the other internally merged to X:

(1) the girls expect which boys to see each other

(2) which boys [the girls expect which boys to see each other]

Further applications of EM yield (3):

(3) I wonder [which boys [the girls expect which boys to see each other]]

Externalization to SM imposes order and also deletes the hierarchically lowest copy, yielding (4):

(4) I wonder which boys [the girls expect to see each other]

If the bracketed clause of (4) appears alone, then *the girls* is the antecedent of *each other*, but not in (4); what is transferred to CI for interpretation is (3), not (4), so that the local antecedent is "which boys," a case of *reconstruction* (the term deriving from earlier more complex and stipulative approaches in which the fronted copy is reinserted by a new operation). The deletion of the internal copy follows from additional principles of minimal computation, presumably third factor properties: mapping to SM is radically simplified by copy-deletion, as is external realization by SM. The hierarchically most prominent copy remains, or there is no evidence that the operation took place.[11] Far more intricate cases proliferate, but minimal computation – optimal Merge and copy deletion – seem to yield the basic properties.

[11] In embedded contexts, under some conditions, a residue of the raised copy remains, providing additional support for the copy theory; see Trinh (2011) for review and analysis. Reflexes of movement through intermediate positions have also been discovered, including verb-raising (Esther Torrego, Spanish), agreement (Sandra Chung, Chamorro; Chris Collins, Ewe), deletion of verbal affix (Doug Saddy, Indonesian). Some accounts of covert movement allow for the internal copy alone to remain after externalization.

Even such simple sentences as (4) have a number of interesting properties. One is that they pose problems for perception: it is necessary to locate the "gap," which may be a difficult task in moderately complex constructions, a familiar problem for parsing systems. If the copy were retained, the problem would be largely overcome. We therefore have a conflict between computational efficiency and efficiency of use (and its subpart, communicative efficiency). In this case, computational efficiency prevails, universally. There are other such cases, and to the extent that they are understood, the same holds: ambiguities, garden path sentences, islands. Throughout, these provide further reasons to suppose that the overarching thesis T is correct.

A further question about (4) is why we do not interpret it to mean something like (5):

(5) I wonder for which boys, the girls expect that those boys will see the other girls

That interpretation, which is a fine thought, would preserve Principle (A) of the binding theory in its simplest form, applying directly to the surface representation. Thus the antecedent of the reciprocal would be the closest appropriate choice in the expression that reaches the ear or eye, whereas the actual principle applies to the abstract form including the copy, visible only to the mind but not to SM. But language design does not permit that simpler option: what reaches SM, again, is of secondary significance for the workings of language.

There are many other examples illustrating the same conclusion. Consider ECP violations, such as (6), close to gibberish, as contrasted with (7), far more acceptable:

(6) (a) how many mechanics do you wonder if fixed the cars

(b) which mechanics do you wonder which cars fixed

(c) which mechanics do you wonder fixed which cars

(7) (a) how many cars do you wonder if the mechanics fixed

(b) which cars do you wonder which mechanics fixed

Here too the thoughts that cannot be expressed by (6) are perfectly straightforward ones, though language design requires some circumlocution. This again provides evidence that computational efficiency trumps efficiency of use (production, perception, communication, etc.), if – and it's a big "if" -- the ECP island condition receives an explanation in terms of optimality of design (as may well be possible; there are some suggestive proposals).

An even simpler case is provided by the much-debated principle of structure-dependence – an informal term, referring to the fact that minimal distance is computed on structural not linear proximity, as in (8):

(8) can eagles that fly swim

The meaning is (9), not (10), though (10) again is a fine thought:

(9) is it the case that eagles that fly can swim

(10) is it the case that eagles that can fly swim

The computational procedure that yields the association ((8), (10)) is far simpler than the one that yields the correct association ((8), (9)), but language does not permit it, in this or any construction in any language in which the issues arise.

The most straightforward explanation for these universal properties of language is that linear order does not enter into core syntactic/semantic computations. That again supports the architectural assumptions sketched earlier, and the broader thesis T. In the case of structure-dependence, a cottage industry has developed within computational cognitive science attempting to show that the facts follow from other cognitive processes, but with no success.[12]

[12] See Berwick, *op. cit*. The critique there is understated, failing to make clear a more fundamental point: the proposed methods, or very similar ones, would work for a pseudo-language that relied on linear order, but no such language exists, surely not an accident. There is evidence from neuroscience that while invented systems conforming to UG (with structure-dependence) activate normal language areas, those that use linear order do not, and are presumably treated as non-linguistic puzzles. See Musso et al. (2003). Note that the problem with (10) cannot be reduced to a principle that predicates do not tolerate gaps; they easily do, as in "(it is a shame to see) how angry John is at the world," etc. More generally, there are, to my knowledge, no accounts of any non-trivial linguistic phenomena in terms of non-linguistic cognitive processes alone, though it is common to claim that they alone suffice. See note 8.

The simplest account of displacement, in terms of IM and NTC, requires a distinction between copies and repetitions. The intuitive basis is clear. In just about every utterance there are likely to be repetitions: of articles, prepositions, verbs, etc. These are unrelated because they are drawn independently from the lexicon (workspace). In the accusative construction (11a) there are two repetitions of *John*; in the unaccusative (11b), two copies:

(11)(a) John saw John

(b) John arrived John

Here too the cases differ in how items are extracted from the workspace; in (11a), *John* is extracted twice, and in (b) only once. Repetitions are unrelated in interpretation; copies in contrast can be regarded as a *discontinuous element* at CI, interpreted as a single element (though the separate copies can make distinct contributions to the unitary interpretation).[13] If IM takes place only at the phase level, then the distinction between copies and repetitions is readily established at the point of transfer, as required. The functioning of discontinuous elements has interesting aspects, which I will have to put aside here.[14]

The architecture suggested above, and the overarching thesis T, correlate with familiar general properties of language. Mapping to the CI interface is generally assumed to be invariant, or virtually so. It is hard to see how it could be otherwise, given the paucity of evidence available for acquisition. There is by now substantial evidence that narrow syntax may also allow only limited variety, virtually none if parametric variation is restricted to the lexicon, or even to functional elements of the lexicon. And simple principles of computational complexity, such as NTC and INC, appear to hold rather generally of narrow syntax, maybe completely. In contrast, externalization radically violates these and other natural principles of efficient computation. It is in general complex, varied, and easily subject to diachronic change, again suggesting that it is an ancillary phenomenon.

The general picture also correlates well with the very little that is known about evolution of language. It appears that there has been no evolution of language (or

---

[13] Separate questions have to do with referring and binding theory, where questions arise relating to modes of conceptualization ("cognitive value"). See Heim, 1998.

[14] See Chomsky (forthcoming).

virtually none; or of cognitive capacities generally) since our ancestors left Africa, perhaps about 50,000 years ago. There are individual differences, but no known group differences. An infant from a tribe in New Guinea that has been isolated for tens of thousands of years, if raised from birth in Boston would have the language of its Boston peers, and conversely.

If we go back roughly 50,000 years before that, there is little evidence that human language existed at all; archaeological evidence suggests that language, and with it complex cognition, emerged within this very narrow window, in what Jared Diamond called a "great leap forward."[15] Doubling the numbers or more changes little; the window remains very narrow in evolutionary time, and many millions of years after separation from other surviving species. These facts suggest that at some point within this narrow range some slight rewiring of the brain occurred yielding the core property of language: Merge with its output linked to the CI interface. Mutations occur in an individual, not a group. The individual endowed with this rewiring would therefore have had a "language of thought" LOT: a means to interpret, reflect, plan, etc., in principle unbounded.[16] An argument would be needed to support the thesis that there is a separate LOT, over and above what the CI interface yields as a mode of representation.

Since there would have been no external pressures, the system arising this way in some individual would have been optimal, determined just by third factor properties, rather like a snowflake, hence conforming to SMT. The mutation could be partially transmitted to offspring, and if it carries some selectional advantage, as seems plausible, it could proliferate through a small hunter-gatherer society. When the capacity is shared, there is some point to externalizing it for interaction, including communication as a special case. But externalization poses a hard cognitive task: it is necessary to link the internal system, perhaps approximating SMT, to a sensorymotor system that had been present for hundreds of thousands of years, so fossil evidence indicates. Accordingly, the problem can be solved in many different ways, each of them complex, and subject to modification. That seems to correspond fairly well with what we know about language. In fact, it is hard to imagine a course of evolution that does not include at least these steps; and it is not easy to find arguments supporting others.

[15] See Tattersall (1998) for general discussion.

[16] It is perhaps worth noting that the general picture seems plausible, maybe even necessary, to leading evolutionary biologists; for quotes from Nobel Laureates, see Chomsky (2005).

Survival of a beneficial trait is a low-probability event, so this might have happened many times and been extinguished, but at least once it was retained, with our ancestors. One might also suspect that the mutation that yields Merge might have taken place many times, not just in the hominid line. But unless it is linked to CI, it would have been of little use, hence unlikely to survive. There have been claims that something of the sort might have happened with songbirds, but simpler interpretationfind of the data that have been advanced (see Beckers et al., 2011). There are other cases of recursive generation in human cognition, the most famous being the natural numbers, a concern that goes back to Wallace and Darwin. But this and other examples that have been suggested might well be piggy-backing on language.[17]

Gallistel (2011) reviews interesting evidence that other primates (and perhaps other species) have representations of the predicate-argument form of human language, and suggests that the "great leap forward" might have been externalization, not recursive generation (let us say just Merge). The latter conclusion seems to me questionable. To yield a LOT that provides the means for free expression and interpretation of thought in the required sense, Merge-generated expressions must be linked to elementary pre-existing mental structures of some kind, perhaps these; perhaps actor-action schemata and others.[18] Hence two steps are necessary to yield externalized human language: Merge and Externalization. The first step carries bounded elementary representations to unbounded and structured LOT. If the first step has been taken, it would be rather surprising if externalization does not follow (as for example happens spontaneously with children acquiring language without input; Goldin-Meadow and Feldman, *op. cit.*). Easy use of other modalities shows that the problem does not reduce to limitations of vocal apparatus. The speculations above seem to me a more parsimonious and plausible account.

It is commonly held that IM is more complex than EM, requiring the operations Form-Copy, Remerge, Copy identification, and Copy deletion. But that is

[17] Reliance on internal computational capacity may yield infinite digital outputs, but that of course does not entail that the output system has its own recursive procedures. Thus externalization of generated I-language expressions does not entail that the vocal apparatus incorporates a recursive procedure. Same with use of the alphabet, or the number system, if it is an offshoot of I-language; or infinite digital visual displays, or other examples that can easily be constructed.

[18] It is worth noting that the most elementary concepts of human languages, such as those used for denoting, seem to be radically different from those found in animal communication systems, raising very serious problems for evolution of language and thought. Perhaps insuperable problems, Lewontin's arguments suggest. Note also that these comments do not touch on far broader and in many respects convoluted issues of what should be regarded as constituting thought.

unwarranted. There are no operations Form-Copy or Remerge, just Merge in its simplest form, satisfying NTC and hence yielding copies. Copy identification is based on a straightforward property, easily detectable at the phase level. Copy deletion follows from elementary third factor considerations.

Prior to the recognition that EM and IM are unified under Merge, it was generally assumed (by me too) that EM and its precursors are somehow natural and that displacement, though ubiquitous, is an "imperfection" of language design, something that requires explanation. But that belief too does not seem tenable. One could on the contrary claim that IM is simpler, because EM requires search through the work space, including the lexicon, while for IM search is restricted to a single object, hence far more economical.[19] Blocking either of the two logically possible cases is a stipulation that requires empirical justification, along with additional justification for whatever new technology is introduced to deal with problems that have a free solution in terms of EM, IM.

In a well-designed system, we would expect that the two logically possible kinds of Merge would correlate with interface effects. At the SM level, that is obvious; IM correlates with displacement, EM with adjacency. At the CI level the distinction correlates closely with duality of semantics: EM yields generalized argument structure, IM yields everything else, specifically discourse-related structure like new/old information, focus, and so on. The correlation is close, but not fully established, and sometimes denied, as in the proposals to reduce control to A-movement.[20] An interesting task is to determine whether the correlation is perfect; if so why, and if not, why not.

As noted at the outset, a recursive procedure need not yield an infinite set of objects (or any at all). We can invent cases in which the range of the I-language would be finite. Imagine, for example, the pseudo-language "truncated English" TE, lacking all the devices of English that permit unbounded length. TE lacks connectives, iterable modifiers ("old old man," "very very old," etc.), possessor constructions ("father of father of Bill"), and so on. A child learning TE will, of course, be unaware of its radical expressive restrictions, and will proceed to acquire it just as if it were normal English based on unbounded Merge, discovering at some point that there is not much to say.[21] It's doubtful that TE could exist.

[19] For arguments on early resort to IM in acquisition, see Roeper (forthcoming).

[20] For extensive review and discussion, see Landau (forthcoming).

[21] Merge is either available by virtue of UG, or unattainable. In general, there no way to learn a system of unbounded computation: addition, Merge or any other. That has been clear at least since David Hume. Separate

Probably the child would simply invent new modes of expression using the resources of UG as in cases already mentioned: isolated acquisition without any input, or the many examples of early acquisition of languages differing from the adult language but conforming to UG . But it is a theoretical possibility. From TE we would learn nothing about UG, or about recursion in language, or about the language faculty generally. It would just be an oddity. It has been proposed that the Amazonian language Pirahä is similar to TE, though the conclusion is contested.[22] The proposal has received enormous publicity, and has been taken to establish far-reaching results about language and cognition, but right or wrong, it is not clear that anything general follows, any more than from TE.

The earliest versions of generative grammar assumed two components -- phrase structure grammar PSG and transformational grammar TG – and identified a number of general properties to be accommodated by the generative system: composition, order, projection, and displacement. The first three were assigned to PSG, the last to TG (along with long-distance morphological relations). PSG was abandoned for substantial reasons in the 1960s, replaced by X-bar theory, which greatly simplified the generative system but imposed new conditions: specifically, universal endocentricity. That aside, the distribution of labor was retained.

We have seen reasons to separate order from compositionality, and to assimilate displacement with composition under Merge. That leaves projection – or in more recent terminology, labeling – which differs from the others in that it is a theory-internal notion, while the others are virtually detectable in the data. It also raises interesting questions, some related to the endocentricity assumption (questionable in my view), but I will put these aside here.[23]

This is the barest sketch of what seem to me to be reasonable conclusions today on the status of minimal recursion, clearly a desideratum for a theory of language, with broad ramifications that should be worth exploration and careful examination.

Bibliography. Minimal Recursion

---

questions have to do with the stage of development at which these properties are exhibited by the child in performance, either perception or (often much later) production.

[22] A further proposal is that this alleged property of Pirahä derives from the cultural features of the community, but these, if they exist, would not be known to a child acquiring the language, hence could have no bearing on UG, just as in the case of TE. On these matters see Everett (2005), Nevins et al. (2009), and subsequent interchange.

[23] For discussion, see Chomsky (forthcoming).

Beckers, G., J. Bolhuis, K. Okanoya, and R. Berwick (2011). “Birdsong neurolinguistics: songbird context-free grammar claim is premature.” *Neuroreport* 2011.

Berwick, R., P. Pietroski, B. Yankama, and N. Chomsky (2011). “Poverty of the Stimulus Revisited.” *Cognitive Science* 35.7.

Borer, H. and K. Wexler (1987). “The Maturation of Syntax.” In T. Roeper and E. Williams, eds. *Parameter Setting*. Reidel.

Chomsky, N. (1955). *Logical Structure of Linguistic Theory*. 1956 revision (partial) published by Plenum and Chicago, 1975.

Chomsky, N. (1957). *Syntactic Structures*. Mouton.

Chomsky, N. (1986). *Knowledge of Language.* Praeger.

Chomsky (2005). “Three Factors in Language Design”. *Linguistic Inquiry* 36.1.

Chomsky, N. (forthcoming). “Problems of Projection.” *Lingua.*

Crain, S. (forthcoming). *Knowledge and Learning in Natural Language*.

Curtiss, S. (forthcoming). “Revisiting Modularity: Using Language as a Window to the Mind.” In M. Piatelli-Palmerini and R. Berwick, eds. *Rich Languages from Poor Inputs.*

Everett, D. (2005). “Cultural Constraints on Grammar and Cognition in Pirahã: Another Look at Design Features of Human Language.” *Current Anthropology.*

Falk, J. (1992). “Otto Jespersen, Leonard Bloomfield, and American Structural Linguistics. *Language* 68.3.

Fitch, T. (2010). “Speech Perception: A Language-Trained Chimpanzee Weighs In.” *Current Biology* 21.14.

Gallego, A., ed. (2012). *Phases: Developing the Framework*. Mouton de Gruyter.

Gallistel, C.R. (2011). “Prelinguistic Thought.” *Language Learning and Development* 7.4.

Goldin-Meadow, S., and H. Feldman (1977). "The Development of Language-Like Communication Without a Language Model." *Science* 197: 401-403.

Heim, I. (1998). “Anaphora and semantic interpretation: a reinterpretation of Reinhart's approach.” In U. Sauerland & O. Percus, eds., *The Interpretive Tract. MIT Working Papers in Linguistics* **25**.

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

Landau, I. (forthcoming). *Control in Generative Grammar: a Research Companion.*

Lewontin, R. (1998). "The evolution of cognition: questions we will never answer." In D. Scarborough and S. Sternberg, eds. *Methods, Models and Conceptual Issues: An Invitation to Cognitive Science*, vol. 4. MIT.

Musso, M., A. Moro, V. Gluache. M. Rijntjes, J. C. Bücheli, and C. Weiller (2003). *Nature Neuroscience* 6**.**774–781.

Nevins, A., D. Pesetsky and C. Rodriguez. (2009). "Pirahã Exceptionality: a Reassessment." *Language* 85.2

Pinker, S. (1984). *Language Learnability and Language Development*. MIT.

Roeper, T. (forthcoming). "*There*-insertion: How Internal Merge Guides the Acquisition Path."

Sauerland, U. and H-M Gärtner, eds. (2007). *Interfaces + Recursion = Language?* Mouton de Gruyter.

Shukla, M., K. S. White, & R. N. Aslin. (2011). "Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants." *Proc. Natl. Acad. Sci. USA*, 108(15).

Soare, R. (1996). "Computability and Recursion." *Bulletin of Symbolic Logic* 2.3.

Tattersall, I. (1998). *The Origin of the Human Capacity* (68th James Arthur Lecture on the Evolution of the Human Brain), American Museum of Natural History.

Trinh, T. (2011). *Edges and Linearization*. MIT Phd dissertation.

Wexler, K. (2002). "Lenneberg's Dream: Learning Normal Language Development and Specific Language Impairment." In J. Schaffer and Y. Levy (Eds.) *Language Competence across Populations: Towards a Definition of Specific Language Impairment.* Erlbaum.

Yang, C. (2002). *Knowledge and Learning in Natural Language.* Oxford.

Yang, C. (2004). "Universal Grammar, statistics, or both?" *Trends in Cognitive Sciences 8*.10.