DOI: 10.1111/mila.12160

WILEY

BOOK REVIEW

The emergence of language

Mark Steedman

Informatics, University of Edinburgh, Edinburgh, UK

Correspondence
Mark Steedman, Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh EH8 9AB, UK.
Email: steedman@inf.ed.ac.uk

Funding information
Australian Research Council Discovery Grant/Award number DP 160102156; European Research Council Horizon 2020 Advanced Fellowship/Award number GA 742137 SEMANTAX

This paper argues that the faculty of language comes essentially for free in evolutionary terms, by grace of a capacity shared with some evolutionarily quite distantly related animals for deliberatively planning action in the world. The reason humans have language of a kind that animals do not is because of a qualitative difference in the nature of human plans rather than anything unique to language.

KEYWORDS
combinatory categorial grammar, language evolution, minimalist program, plans and the structure of behavior, syntax–semantics interface

With admirable succinctness, Berwick and Chomsky sum up the main arguments of their (2016) book *Why Only Us* (hereafter *WOU*) as follows (p. 87):

> In some completely unknown way, our ancestors developed human concepts. At some time in the very recent past, apparently sometime before 80,000 years ago, if we can judge from associated symbolic proxies, individuals in a small group of hominids in East Africa underwent a minor biological change that provided the operation Merge—an operation that takes human concepts as computational atoms and yields structured expressions that, systematically interpreted by the conceptual system, provide a rich language of thought. These processes might be computationally perfect, or close to it, hence the result of physical laws independent of humans. The innovation had obvious advantages, and took over the small group. At some later stage, the internal language of thought was connected to the sensorimotor system, a complex task that can be solved in many different ways and at different times. In the course of these events, the human capacity took shape, yielding a good part of our "moral and intellectual nature," in Wallace's phrase.

There is a lot that one can agree with in the above statement. It seems entirely reasonable to argue that the human languages we actually encounter have as their evolutionary and developmental precursor a universal conceptual language of mind whose origins lie in the need to reason about the

 © 2017 John Wiley & Sons Ltd

world, rather than in any more primitive form of communication, and that the attachment of this language of mind to the forms of individual languages is logically secondary. As Chomsky himself has pointed out since 1965 (hereafter *Aspects*), since languages differ in structure, and the input to the child language learner contains few markers if any of structure (the "poverty of the stimulus"), the only plausible account of the child's ability to learn language is that he or she has access to an underlying universal representation, ultimately related to semantics or a conceptual representation of the world around him/her.

It follows, as *WOU* argues elsewhere, following Lenneberg (1967, pp. 234–239), that we should not look for continuity in evolution between the communication systems of animals, such as birds and whales, on the one hand and language on the other. We should rather expect that a precursor developed for some unrelated human cognitive function has shaped the communicative apparatus to its own ends.[1]

One might even also agree with *WOU*'s rather gnomic statement about the language of mind being the reflection of "physical laws independent of humans" if by that is meant that it is shaped by the way the world is (although that world of course also includes other humans, so the aforesaid application to communication with them is rather immediate).

The only points of disagreement with *WOU* that seem possible are relatively minor. One concerns the extremely short time span that they allocate to this process, which they believe occurred around 100,000 years ago or, at any rate, in the "very recent past." This claim rests, as they are careful to say in the passage quoted above, on the assumption that the emergence of durable secondary symbolic artefacts like cave paintings and musical instruments are immediate and inevitable concomitants of the core symbolic activity of language use.

The other point of possible contention concerns the identification of the Minimalist Programmatic operation called Merge (Chomsky, 1995) with the crucial evolutionary innovation that made this process possible. This identification is considered to follow from the claim that Minimalism is the simplest and most explanatory theory of linguistic competence and that its introduction therefore qualifies as the best candidate we have for the "minor biological change" referred to by *WOU* above.

To discuss these two finer points, we must begin by briefly reviewing what is known from the paleoanthropological record concerning the evolution of human cognition.

1 | WHAT WE KNOW ABOUT THE EVOLUTION OF HUMAN COGNITION

Human cognition and language seem very different from analogous functions in animals. Yet humans are descended from animals through a process of evolution, sharing a common ancestor with our closest-related existing species, the chimpanzees, only around 10 or so million years ago.

The fossil record shows that some of the earlier species on the evolutionary line that probably led directly from that common ancestor to humans, the australopithecines, were characterized by upright posture and bipedal locomotion at least 4 million years ago, but seem otherwise to have been ape-like, particularly in terms of cranial capacity. The genus *homo* seems to have arisen between 3 and 2 million years ago. It is not clear how many distinct species it comprises, but the genus is characterized by a steady increase in cranial capacity relative to body weight across all forms until modern humans. From at least 2 million years ago, all forms have been found to be associated with shaped stone tools that are not found with other genera. About 1.5 million years ago,

[1] Darwin (1871), p. 57, cited in *WOU*, makes the same point.

there was a technological refinement with the appearance of more finely worked tools, which coexisted with the simpler type for the next half million years. Other evidence of technological innovation, such as permanent hearths, and evidence of symbolic activity, such as cave-painting and musical instruments, is much more recent, from around 100,000 years ago (Deacon, 1997; Lewontin, 1998; Tattersall, 1995).

What does the theory of evolution tell us about the ontology of human language and cognition over this period? As Lewontin points out, very little. The technological advances these authors describe cannot be identified with genetic mutation or evolutionary change. In particular, while it is reasonable to argue that evidently symbolic activities like painting and music draw upon the same cognitive substrate as language, it does not follow that their emergence can be considered a proxy or temporal marker for the emergence of language itself. For an individual or a society to be able to support activities like drawing or music, including preparatory activities like grinding pigments and inventing and making flutes, requires a stable way of life with assured surpluses of the bare necessities for life that allows such expenditure of labor. To make paintings that have lasted for tens of thousands of years also implies the use of media for those pigments, requiring the sacrifice of valuable materials such as fat, eggs, and bone marrow that could be used otherwise as food and fuel. Since suitable mineral pigments are rare and may not be obliging enough to be at hand where food and shelter are most plentiful, they also require time to be spent on travel to collect them and possibly even the development of trade networks.

There are, in fact, only two concrete pieces of evidence for evolutionary events specifically related to the ontogeny of distinctively human cognition. First, there is no parallel to even the earliest and most primitive human manufacture of tools in any animal, living or extinct. Many animals, including apes, but also some of the more recently evolved birds, can solve problems requiring tools. In particular, apes use stones as hammers and anvils and can improvise complicated plans to obtain out-of-reach bananas using piles of crates, sticks, etc. (Köhler, 1925). They can also modify available tools such as branches by stripping them by hand of leaves and twigs or bend a wire into a hook. However, if the tools are not immediately available in the problem situation, but are (say) in an adjacent room, then the animals have immense difficult in solving the problem (even if they have previously seen them there). Similarly, plans that require the use of tools such as stones to make the tool that is actually needed, such as a blade, seem to be beyond them. Animals seem to have to rely on whatever tools the immediate environment affords. We can conclude that the genus *homo* has been cognitively quite different from other apes from the very start.

Of course, we cannot draw any conclusion about any corresponding use of language from that fact alone. However, the second piece of evidence is more suggestive of the presence of language. The vocal tract of modern human beings shows clear signs of an evolutionary adaptation to the production of the very wide range of vocal gestures that characterize human speech, as opposed to chimpanzee vocalization. (This evolutionary adaptation has been so rapid and extreme as to leave adult humans alone among animals in not being able to swallow and breathe at the same time, a change that would otherwise seem to be maladaptive as it can cause them to die prematurely by choking on food.) The adaptation in question involves a number of anatomical changes that allow the vocal tract to lengthen by lowering the larynx. Lieberman, Laitman, Reidenberg, and Gannon (1992) and Lieberman and McCarthy (2007) have shown from fossil evidence that over at least the last 2 million years of human evolution, including at least the development of the *erectus* and later *neanderthalis* forms, there has been a steady process of lengthening of the vocal tract and lowering of the larynx in *homo.*

There are other possible biological explanations for lengthening in animal vocal tracts. Fitch and Reby (2001) have argued convincingly that the larynx is descended in male red deer as an adaptation with the function of attracting mates by exaggerating the impression of male physical size. Fitch goes on to suggest that the lowering of the human larynx noted by Lieberman had a similar evolutionary function and has only recently been recruited to linguistic purposes. It is of course perfectly possible that our ancestors went through a laryngeal arms race of this kind, though it is odd that it should have applied equally to males and females and that neither modern humans nor extant apes of other genera seem to seek reproductive success in anything similar to this way.

It seems at least as likely, as Deacon argues (Deacon, 1997, pp. 354–365), that the changes that Lieberman describes reflect the presence of an essentially fully human language, but that its original medium was primarily gestural and only gradually recruited and adapted the vocal apparatus (because of its superior properties when communicating around corners or in the dark) by incremental lowering of the larynx to the point where the role of gesture became minor.[2]

Under such assumptions, we would seek a single evolutionary event that would support *both* the ability to make plans requiring the manufacture of situationally nonafforded tools *and* support the use of language.

Such an evolutionary innovation is not hard to imagine: making a plan to skin a quarry with a yet-to-exist knife seems logically quite similar to making a plan to get someone else who is not yet involved to cooperate in the same activity (Steedman, 2002; Steedman & Petrick, 2007). The only question seems to be which of the two capabilities carried a direct evolutionarily selectional advantage that allowed the other to piggyback upon it. Since evolution is generally characterized by the recombination of previously independently evolved mechanisms, and we know of other related instances of cooperation among primates such as tamarins (Hrdy, 2007), it seems most likely that it is the human faculty for cooperative action that was actively selected, rather than tool manufacture as such.

One concrete example of such cooperation for which the evolutionary population dynamics has been quite thoroughly explored (Hamilton, 1964) has been suggested by Fitch (2007), Burkart, Hrdy, and Van Schaik (2009), and Hrdy (2009) to be cooperative child rearing, the evolutionary pressure that may in turn stem from selection for a further characteristic of human evolution, namely neoteny, or a prolonged period of infant development (Haldane, 1932). Both human tool use and human communication would then be a simple extension of nonhuman cognition to planning using conceptual knowledge of actions involving nonself agencies.[3]

Lewontin argues forcefully that these are questions to which we cannot know the answer because the evidence is simply not there, a point that *WOU* repeatedly endorses.

However, in such a situation, it is standard to appeal to Ockham's Razor. A theory in which there is *one* evolutionary event that might be related *both* to the human conception of the world and to human language, of a kind that has been seen elsewhere in the nearby evolutionary landscape, might be preferred on the grounds of simplicity to one in which the same evolutionary endpoint requires *two* unrelated evolutionary events, one for the development of human concepts and one that is claimed to be entirely specific to language.

Interestingly, *WOU*'s argument for the second of these evolutionary events is of exactly this kind, based on parsimony. They argue that because the Minimalist theory defines human language

[2] See Bolhuis, Tattersall, Chomsky, and Berwick (2014) and related exchange for a contrary view on the countervailing advantage of this evolutionary development.

[3] This possibility does not conflict with *WOU*'s assumption that a language of mind representing the way the world works is logically prior to its use for communication. It merely allows that some of that cognitive substrate for humans is social, as Tomasello (1999) has argued.

in terms of a single primitive operation Merge, it places the least burden on the theory of evolution to come up with it.

2 | THE EMERGENCE OF MERGE

Merge is defined by Chomsky (2001/2004) and *WOU* as a composite of two operations, External Merge and Internal Merge. External Merge is simply the formation of a structure, say [NP,VP], from two elements NP and VP. (Thus, external merge is closely related to the notion of a context-free phrase–structure rule such as S → NP VP. The technical difference need not detain us.) In semantic terms, External Merge is similarly simple: one of the elements NP and VP is a predicate (or, more generally, a function), and the other is an argument.

The structure can then be interpreted as the application of the function to the argument. (For example, we might think of the VP as a function such as *walks* that applies to people like *gilbert* to yield propositions like *walks (gilbert)*, the meaning of "Gilbert walks" in the Language of Mind. Less intuitively, we could think of "Gilbert" as denoting a higher-order function, applying to the same function *walks* to yield the same result.)

The other component, Internal Merge, is very different in character. It used to be called Move because it corresponds to the phenomenon of "displacement" in natural language, where elements that belong together at the level of meaning as function and argument are separated in the sentence, as *book* and *reading* are in *Which book is Gilbert reading?*

The authors (WOU:175,n.9) deny that any difference or "natural break" exists between Internal and External Merge, pointing out that "the simplest wh-question [like the above], which is still context-free, invokes the internal version of Merge." However, that just means that for such simple *wh*-questions, there is, as Gazdar (1981) demonstrated, an alternative, entirely context-free analysis in terms of External Merge alone. They cannot intend to argue that Internal Merge *in general* actually is interpreted as a kind of Generalized Phrase Structure Grammar (GPSG)-style hypercyclic recursive application of context-free External Merge because, in the general case, Internal Merge, unlike External Merge, generates constructions that are *not* context-free. Perhaps they are thinking of the related, more expressive feature-passing mechanism of Head-driven Phrase Structure Grammar (HPSG), as has been proposed recently by Neeleman and van de Koot (2010). In any case, something more than application seems to be needed.

Semantically, Internal Merge corresponds to the operation of abstraction, or formation of a new function from a proposition, using a variable, say x, in place of an argument somewhere in it and designating x as the argument of the new function. Thus, *Which book is Gilbert reading* is turned into something like the property "book □ such that Gilbert is reading □" (c.f., *WOU*:100).

A theory that consisted only of the two rules Apply and Abstract would be a completely general applicative system and would be very expressive indeed. In the absence of any specifics about how Internal Merge is implemented (particularly the implicit copying process that it assumes), it is not clear in what sense its evolutionary introduction can be claimed to be "minor" (as in the quotation from *WOU* with which we began.)

WOU goes to considerable lengths to defend the Merge theory against the suspicion that this kind of all-in-one appearance of a fully formed, otherwise-unprecedented innovation is not the way that evolution usually works. It is well-known that the speed of evolutionary change is nonuniform: the fossil record consists of long periods of equilibrium where nothing much changes, punctuated by periods of very rapid change where new species suddenly appear and old ones equally suddenly vanish.

This phenomenon, sometimes referred to as "saltation," is widely understood as arising through a number of mechanisms. If the environment is unchanging, then the effect of natural selection is to actively limit drift away from the local optimum it has reached so far, but if the environment changes drastically, then the old optimum is likely to be optimal no longer, and variation, especially if accelerated and distributed over a group by sex, is selected positively rather than negatively. (This is the pattern in the great extinctions, such as the one that led to the elimination of the dinosaurs and the rise of the mammals.) Another mechanism is reuse of previously evolved structures as macroinstructions. The development of segmented creatures, such as modern millipedes, seems to treat the segment as a macro.

Another powerful mechanism is symbiosis (Dyson, 1985; Margulis, 1981), whereby two autonomous organisms merge to their mutual evolutionary benefit. The incorporation of algal cells in primitive eukaryotes to form the mitochondria and chloroplasts seen in the cells of every animal and plant was an event of this kind. This process also seems to have played a part in the even earlier rapid evolution of the primitive eukaryotes, or cells with nuclei, formed by the symbiosis of a bacterium with an archeal host cell (Lane, 2015). A still more primitive version of this process is the direct incorporation of genetic material from one organism into the genome of another, known as lateral gene transfer, as is common among bacteria.

However, despite the availability of these various mechanisms for accelerated evolution, biologists feel it incumbent upon them to show how the developments for which they are invoked could have taken place in the time period defined by the fossil record and from known or inferable precursors. For example, Lane (2015, p. 194) says of the eukaryotic development:

> To suggest that the nucleus, for example, somehow just popped into existence is to confound genetic saltation with adaptation. The nucleus is an exquisitely adapted structure, … the product of natural selection acting over extended periods of time, [in] a purely Darwinian process. But that does not mean it had to happen slowly in geological terms.

Lane devotes the rest of his book telling us exactly how this could have happened in the known time period and whether it only happened once or if it may have happened repeatedly.

WOU energetically invokes the notion of saltation as a justification for identifying Merge as a unitary mutation, following Berwick (1997) and citing Lane in support. However, it does not explain how this remarkable event could happen in evolutionary terms.[4]

WOU attempts to deal with this difficulty by defining Minimalism, quite correctly, as a "Theory of the Computation," in the sense of Marr (1977, 1982), who defines three levels of representation at which a computation can be defined: the Theory, the Algorithm, and the Machine.[5]

The universe in which an organism evolves may be such that computing a certain function, such as the Fourier Transform or an applicative system, may help it survive and multiply. In that sense, the Theory of Computation may be thought of as a ghost in the evolutionary machine. (For example,

---

[4] *WOU* repeatedly cites Lenneberg (1967) in this connection, in support of discontinuity in the evolutionary record. However, Lenneberg's concern (1967, pp. 227–239) seems to have been exclusively with evolutionary noncontinuity of human language with animal communication systems.

[5] Marr uses the example of the Fourier Transform, which among other things can be applied to decompose complex waveforms into the sum of a number of simple frequency components to illustrate the three levels of analysis. The Theory of Computation is the most abstract level of description, which in the case of the Fourier Transform is a mathematical equation. The second level is that of the algorithm, where there are several quite different methods for actually computing the function so defined, including the Fast Fourier Transform and optical methods. The last of Marr's levels is the machine on which the algorithm is implemented, which again varies along dimensions such as digital, analog, neural, and so on and which may determine the choice of algorithm.

the world we live in makes vision so advantageous that the lens-based cameral eye has been independently evolved through quite different routes in vertebrates and cephalopods.)

However, as Johnson (2016) points out, the Theory of Computation itself can play no direct part in the evolutionary process as it is just what is “out there,” an invisible hand determining fitness, as yet with no internal representation in the organism.

Evolution by natural selection has to work at the level of the somatic machine and the algorithm, by a process of generating variations on whatever representations are already implicit in the organism’s existing genetic code. Among the other mechanisms of saltation mentioned earlier, these variations can be at the level of structures or macroinstructions rather than codons, so they can induce large phenotypic changes.[6]

Even if we accept that Merge is a unitary rule and that, as such, it constitutes the simplest Theory of Computation for human linguistic competence (which would require us to also take into consideration all the constraints and minimality conditions that limit its application), that theory itself tells us very little about the specific solution that evolution has been able to reach in order to actually compute that function.

In particular, it does not follow that the simplest evolutionary route to implementation of the theory involves a correspondingly unitary mechanism or even a monolithic rule of internal merger as such. If a number of components, such as the composition and type-raising operators of Combinatory Categorial Grammar (CCG, Steedman, 2000)—independently evolved for other reasons—can be combined to achieve the same result, then they may offer a shorter route (or even the only accessible one).

This is often the way discontinuity in evolution works: for example, insect wings and feathers are thought to have originally evolved with thermoregulatory functions before the physics that was “out there” variously afforded their use in flight. The following section proposes a way that composition and type raising could have evolved separately for quite different functions, only later to be co-opted by language.

3 | PLANS AND THE STRUCTURE OF BEHAVIOR

It is, in fact, quite likely that an operation-like composition should have been independently evolved for other reasons before being co-opted by the nascent language faculty. Chimpanzees can, as was noted earlier, make elaborate plans involving tools, for example, placing a box under otherwise

[6] *WOU*:175-6,n.9 accuses Steedman (2014) of describing the process of evolution as consisting in complete search of all possible codeable sequences, citing a remark there that evolution “essentially works by trying every possible variation on every viable variation so far.” The authors seem to have missed the words “possible,” “viable,” and “so far,” together with a footnote saying “It is a little more complex than this. Even genomes must be structured programs, and even evolution needs the occasional mass-extinction to escape overfitting.”

Of course, the fact that all stages of evolutionary development must be viable places stringent limits on the solutions that are reachable. My point was that, within those limits, the search can be exhaustive as the number of variations generated, most of which are lethal, is essentially unlimited and can find extremely singular solutions. For example, the genetic code itself, consisting of 64 triplets of two base-pairs encoding the 20 amino acids that make up all proteins found in terrestrial life, is a near-optimally efficient outlier for its function of reliable transmission of the genome (Freeland & Hurst, 1998; Freeland, Knight, & Landweber, 2000; see Morris, 2004, pp. 17–19; Di Giulio, 2005).

WOU:175 also suggests that Steedman (2014), claims that “[evolution] has solved the acquisition problem offline,” whatever that could mean. However, it is quite clear from the passage the authors quote that it contrasts the capability of evolution to develop the language of mind that Berwick and Chomsky themselves assume with language acquisition in the child, where the poverty of the language stimulus means that the child must have prior access to it.

See Kwiatkowski, Goldwater, Zettlemoyer, and Steedman (2012) and Abend, Kwiatkowski, Smith, Goldwater, and Steedman (2017) for a fully online algorithmic model of child language acquisition of the kind that my colleagues and I actually advocate.

inaccessible bananas and continuing to stack boxes until climbing the stack allows access to the bananas.

Actions like stacking a box on something and climbing on it are functions that map states of the world into other states of the world. Imagining taking one such action in a given state is *application* of that function to that state. Thus, any animal that can do that much has External Merge.[7]

A plan like placing boxes on boxes until climbing on them allows you to obtain the bananas is also a function mapping states onto other states, but to imagine this composite function, you have to mentally compose elementary actions of stacking boxes until climbing will get you high enough. Only then can you know that executing the plan will result in a state where you have the bananas. Thus, composition already does some of the work of abstraction and, hence, of Internal Merge.

Similarly, your chances of actually finding such a plan, given that there is a potentially very large number of mostly irrelevant or impossible actions you could consider in this planning process, will be enhanced if you organize your knowledge of actions around the objects that afford those actions. That way, if there are boxes around, you think about plans involving actions of stacking and climbing, whereas if there are sticks, you think about those involving hitting and poking. The replacement of an object by a function over the actions that it affords into their results is closely akin to the operation of type raising. Type raising and composition alone are known to support a restricted form of applicative system (Smullyan, 1985). Thus, it is also reasonable to think that the ape language of mind may include type raising as well as composition. If so, apes already have some elementary operations of an applicative system of the kind we see in language and, hence, some direct precursors of Merge.

Not surprisingly, since apes nevertheless do not evince human language, the language of mind that is implicit in such ape plans still lacks many characteristics of the precursor to human language that *WOU* envisages in the passage that this paper began by quoting. The language of plans is essentially a finite-state language of "*while*-loops," of the kind Piaget (1936) called "circular reactions" and Miller, Galanter, and Pribram (1960) called "TOTE Units."

Nevertheless, this finite-state language of plans already has what *WOU*:1 calls the "Basic Property" of linguistic competence, which they define as "a finite computational system yielding an infinity of expressions, each of which has a definite interpretation in semantic-pragmatic and sensory-motor systems" since the number of plans that can be expressed in it and executed is, in principle, unbounded.

It is presumably this chimpanzee language of mind to which sign systems, like the one taught to Washoe by Gardner and Gardner (1969) and even some communicative gestures native to chimpanzees, attach. (Such sign systems are heavily oriented toward actions and action sequences, McNeill, 1980, p. 148.)

However, the expressive limitations on such communicative systems and differences from human language use (Premack, 1986; Premack & Premack, 1983) are well known. Apes' use of sign language should therefore be seen as evidence that if you have a language of mind of any kind, then a sensory motor communicative system of some sort can be attached to it, rather than as constituting an evolutionary precursor to human language in its own right, of the kind proposed by Bickerton (1992), Jackendoff (1999), and Progovac (2014, 2015).

Nevertheless, it is worth noticing that in order to arrive at the piling and climbing plan, the ape has to carry out a *search* through a tree of possible future states as each state typically permits more

[7] In fact, it isn't clear that you can do *anything* we would want to call thinking—that is, anything deliberative, going beyond purely reactive reflex action—without function application. (In particular, you obviously can't apply operators such as abstraction or Copy/Internal Merge without application/External Merge, as Brody (1995) and Nunes (2004) point out.)

than one action. Such search requires something equivalent to a stack machine or push-down automaton (PDA) (actually, since memory is in fact bounded, a *simulation* of such a machine) to keep track of the actions that remain to be explored in each state reached by each possible action.[8]

Whatever the evolutionary pressure was for human social collaboration, it requires a different kind of planning ability. Preparing food may require finding somebody who is willing to mind the baby. But that willingness doesn't exist: you need to cause it to come into being. You may need to go and find somebody who will commit to it. Moreover, in order to do that, you need to be able to imagine that if someone would do so, whoever it was, you would be able to get on with dinner (c.f., Hamburger & Crain, 1987).

This kind of planning already goes beyond chimpanzee planning abilities. It is parallel to seeing that, if you had a box or two, you could reach the banana and making a plan to first obtain the boxes, a problem at which we have noted apes generally fail. It is also the kind of planning involved in recognizing that if one made a knife, one would be able to skin the animal *with that currently nonexistent knife* (or conversely, that if you make a knife first, someone will be able to skin a yet-to-be-obtained animal).

This kind of planning is well known in the AI literature as "planning with variables" (Etzioni et al., 1992; Petrick & Bacchus, 2002). If I want to visit you, and I don't know where you live, then a very bad plan is to start knocking on doors until the person who answers is you. A better plan involves the use of a variable we might call *address*, making a plan to set *address* to your address (say by looking it up in my address book) and then going to *address*, wherever that is, and knocking on the door there.

Planning with variables allows you to make more complicated plans, such as *finding some kid who will find an adult who will mind the baby*, involving multiple agents. Making these plans also require graph search, in this case, for a situation where you are free to get on with dinner. Graph search again requires simulation of a recursive process of function composition with a PDA.

However, the language of human plans has more structure than in the ape case. As far as the Theory of Computation goes, it is a language with an unbounded number of constituent types as its elements can, in principle, be functions over arbitrary numbers of individuals. It is therefore a language of greater expressivity than context-free.

Interestingly, constituents of just this multivalent character can be seen mirrored in the verbal clusters of verb-final languages like Dutch and Zurich German, which were used to provide the first formal proofs of the trans-context-free expressivity of human grammars (Huybregts, 1984; Shieber, 1985). (In such languages, the plan that in English is expressed as "get some kid to find an adult to mind the baby" is glossed with the order "some kid an adult the baby [get find mind]," with the verb cluster homomorphic to a plan involving three plan variables.)

It therefore seems reasonable to speculate that the crucial evolutionary event leading to human language was a shift to a way of life requiring the generalization of the prehuman ability for composing sequences of elementary actions into plans to the ability to do so for plans involving multiple agents. This ability constituted a language of mind capable of supporting human language.

[8] A related point is made by Hauser, Chomsky, and Fitch (2002), p. 1578. Of course, to describe this mechanism as a PDA is to define it at the level of the Theory of Computation. The simulation may of course be in terms of counters, dynamic programming, etc., using whatever machinery is available. Thus, this proposal does not constitute "evolutionary incremental 'tweaking' of a stack-like architecture" as *WOU*:175,n.9, suggests concerning the related proposal in Steedman (2014). Nor is it the case, as *WOU*:131-2 rather bizarrely argues, that because ants have memory and therefore embody a machine of a type that *could* be used to simulate a Turing machine to some approximation, that ants therefore already embody the endpoint of such evolutionary tinkering. Clearly, the issue is not what can, in principle, be simulated with an FSM but what actually is being computed by ants, apes, and people and how evolution has worked to make their computations rather different.

If so, we might tell a slightly different story from the quotation from *WOU* with which we began, which might go as follows:

> In some unknown way, but in a development with known parallels in other evolutionarily unrelated lineages, our non-human primate ancestors developed the capacity for deliberative planning of sequential action, requiring search through hypothetical state spaces, using sequential operators of merger such as application and composition of actions/functions representing state-change, sometimes afforded by tools that the current situation includes, of a kind we see other animals, including their non-human descendants, applying today, constituting structured symbolic plans.
>
> More recently, somewhere between 2 and 3 million years ago, in some equally unknown way, the human line became distinguished by access to a new kind of cooperative action concept, requiring planning with variables representing non-present agents and resources, constituting a good part of our moral and intellectual nature.
>
> The language of mind that was implicit in such plans logically preceded but immediately supported connection to the sensory motor system to form communicative languages, a task which can be solved in many different ways and at different times, and was initially heavily gestural.
>
> We can have no direct evidence for the form of such languages, but we have no reason not to suppose them to have had all the characteristics of full human language. The same development in planning ability also immediately supported the manufacture of tools.
>
> After a long period during which the only strictly linguistic evolutionary development consisted in continuous adaptation of the vocal tract to the advantageous medium of vocal speech, the human lineage got lucky enough for long enough to apply the planning abilities that were already supporting language and tool manufacture to the development of more sedentary ways of life, supported by larger groups and resources like leisure, trade and agriculture, eventually allowing the emergence of less evanescent symbolic practices related to language, including painting and music.

But that is just another story, and as Lewontin (1998, pp. 129–130) wisely (and somewhat testily) reminds us, concerning the evolution of human cognition, “there is no end to plausible storytelling.”

ACKNOWLEDGEMENTS

Thanks to Geoff Pullum, John Torr, and Bonnie Webber for reading the draft. The work was supported in part by ARC Discovery Grant DP 160102156 and ERC H2020 Advanced Fellowship GA 742137 SEMANTAX.

REFERENCES

Abend, O., Kwiatkowski, T., Smith, N., Goldwater, S., & Steedman, M. (2017). Bootstrapping language acquisition. *Cognition*, *164*, 116–143.

Berwick, R. (1997). Syntax facit saltum: Computation and the genotype and phenotype of language. *Journal of Neurolinguistics*, *10*, 231–249.

Berwick, R., & Chomsky, N. (2016). *Why only us: Language and evolution*. Cambridge, MA: MIT Press.

Bickerton, D. (1992). *Language and species*. Chicago, IL: University of Chicago Press.

Bolhuis, J., Tattersall, I., Chomsky, N., & Berwick, R. (2014). How could language have evolved? *PLoS Biology*, *12*, e1001934.

Brody, M. (1995). *Lexico-logical form: A radically minimalist theory*. Cambridge, MA: MIT Press.

Burkart, J., Hrdy, S. B., & Van Schaik, C. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, *18*, 175–186.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, N. (2001/2004). Beyond explanatory adequacy. In A. Belletti (Ed.), *Structures and beyond: The cartography of syntactic structures* (Vol. 3, pp. 104–131). Oxford: Oxford University Press.

Darwin, C. (1871). *The descent of man*. London: John Murray.

Deacon, T. (1997). *The symbolic species*. New York, NY: Norton.

Di Giulio, M. (2005). The origin of the genetic code: Theories and their relationships, a review. *Biosystems*, *80*, 175–184.

Dyson, F. (1985). *Origin of life*. Cambridge, UK: Cambridge University Press.

Etzioni, O., Hanks, S., Weld, D., Draper, D., Lesh, N., & Williamson, M. (1992). An approach to planning with incomplete information. In *Proceedings of the 2nd International Conference on Knowledge Representation and Reasoning (KRR-2)* (pp. 115–125). Menlo Park, CA: AAAI.

Fitch, T. (2007). Evolving meaning: The roles of kin selection, allomothering and paternal care in language evolution. In C. Lyon, C. Nehaniv, & A. Cangelosi (Eds.), *Emergence of communication and language* (pp. 29–51). Berlin: Springer.

Fitch, T., & Reby, D. (2001). The descended larynx is not uniquely human. *Proceedings of the Royal Society of London B: Biological Sciences*, *268*, 1669–1675.

Freeland, S., & Hurst, L. (1998). The genetic code is one in a million. *Journal of Molecular Evolution*, *47*, 238–248.

Freeland, S., Knight, R., & Landweber, L. (2000). Measuring adaptation within the genetic code. *Trends in Biochemical Sciences*, *25*, 44–45.

Gardner, A., & Gardner, B. (1969). Teaching sign language to a chimpanzee. *Science*, *165*, 664–672.

Gazdar, G. (1981). Unbounded dependencies and coordinate structure. *Linguistic Inquiry*, *12*, 155–184.

Haldane, J. (1932). *The causes of evolution*. London: Macmillan.

Hamburger, H., & Crain, S. (1987). Plans and semantics in human processing of language. *Cognitive Science*, *11*, 101–136.

Hamilton, W. (1964). The genetical theory of social behavior, i and ii. *Journal of Theoretical Biology*, *7*, 1–52.

Hauser, M., Chomsky, N., & Fitch, T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*, 1569–1579.

Hrdy, S. B. (2007). Evolutionary context of human development: The cooperative breeding model. In C. Salmon & T. Schackelford (Eds.), *Family relationships: An evolutionary perspective* (pp. 39–68). New York, NY: Oxford University Press.

Hrdy, S. B. (2009). *Mothers and others*. Cambridge, MA: Belnap/Harvard University Press.

Huybregts, R. (1984). The weak inadequacy of context-free phrase-structure grammars. In G. de Haan, M. Trommelen, & W. Zonneveld (Eds.), *Van Periferie naar Kern* (pp. 81–99). Dordrecht: Foris.

Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences*, *3*, 272–279.

Johnson, M. (2016). Marrs levels and the minimalist program. *Psychonomic Bulletin & Review*, *24*, 171–174.

Köhler, W. (1925). *The mentality of apes*. New York, NY: Harcourt Brace and World.

Kwiatkowski, T., Goldwater, S., Zettlemoyer, L., & Steedman, M. (2012). A probabilistic model of syntactic and semantic acquisition from child-directed utterances and their meanings. In *Proceedings of the 13th Conference of the European Chapter of the ACL (EACL 2012)* (pp. 234–244). Avignon: ACL.

Lane, N. (2015). *The vital question: Energy, evolution, and the origins of complex life*. New York: WW Norton & Company.

Lenneberg, E. (1967). *The biological foundations of language*. New York, NY: John Wiley.

Lewontin, R. (1998). The evolution of cognition: Questions we will never answer. In *An invitation to cognitive science* (Vol. 4, *107–132*). Cambridge, MA, MIT Press.

Lieberman, P., Laitman, J., Reidenberg, J., & Gannon, P. (1992). The anatomy, physiology, acoustics and perception of speech: Essential elements in analysis of the evolution of human speech. *Journal of Human Evolution*, *23*, 447–467.

Lieberman, P., & McCarthy, R. (2007). Tracking the evolution of language and speech. *Expedition*, *49*, 15–20.

Margulis, L. (1981). *Symbiosis and cell evolution*. San Francisco, CA: Freeman.

Marr, D. (1977). Artificial intelligence: A personal view. *Artificial Intelligence*, *9*, 37–48.

Marr, D. (1982). *Vision*. New York, NY: Henry Holt and Co., Inc.

McNeill, D. (1980). Sentence structure in chimpanzee communication. In T. Sebeok & J. Umiker-Sebeok (Eds.), *Speaking of apes* (pp. 145–160). Berlin: Springer.

Miller, G., Galanter, E., & Pribram, K. (1960). *Plans and the structure of behavior*. New York, NY: Holt.

Morris, S. C. (2004). *Life's solution: Inevitable humans in a lonely universe*. Cambridge, UK: Cambridge University Press.

Neeleman, A., & van de Koot, H. (2010). A local encoding of syntactic dependencies. *Syntax*, *13*, 331–372.

Nunes, J. (2004). *Linearization of chains and sideward movement*. Linguistic Inquiry Monograph 43. Cambridge, MA: MIT Press.

Petrick, R., & Bacchus, F. (2002). A knowledge-based approach to planning with incomplete information and sensing. In *Proceedings of the Sixth International Conference on Artificial Intelligence Planning and Scheduling (AIPS-2002)* (pp. 212–221). Menlo Park, CA: AAAI Press.

Piaget, J. (1936). *La naissance de l'intelligence chez l'enfant,* Paris: Delachaux et Niestle*, translated 1953 as The Origin of Intelligence in the Child*. London: Routledge and Kegan Paul.

Premack, D. (1986). *Gavagai!*. Cambridge, MA: MIT Press, Bradford Books.

Premack, D., & Premack, A. J. (1983). *The mind of an ape*. New York, NY: Norton.

Progovac, L. (2014). Degrees of complexity in syntax: A view from evolution. In *Measuring grammatical complexity* (pp. 83–102). New York, NY: Oxford University Press.

Progovac, L. (2015). *Evolutionary syntax*. Oxford: Oxford University Press.

Shieber, S. (1985). Evidence against the context-freeness of natural language. *Linguistics and Philosophy*, *8*, 333–343.

Smullyan, R. (1985). *To mock a mockingbird*. New York, NY: Knopf.

Steedman, M. (2000). *The syntactic process*. Cambridge, MA: MIT Press.

Steedman, M. (2002). Plans, affordances, and combinatory grammar. *Linguistics and Philosophy*, *25*, 723–753.

Steedman, M. (2014). Evolutionary basis for human language: Comment on "toward a computational framework for cognitive biology: Unifying approaches from cognitive neuroscience and comparative cognition" by Tecumseh Fitch. *Physics of Life Reviews*, *11*, 382–388.

Steedman, M., & Petrick, R. (2007). Planning dialog actions. In *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue, Antwerp, Sept., ACL* (pp. 265–272). ACL Stroudsberg, PA.

Tattersall, I. (1995). *The fossil trail: How we know what we think we know about human evolution*. Oxford: Oxford University Press.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

How to cite this article: Steedman M. The emergence of language. *Mind Lang*. 2017;32:579–590. https://doi.org/10.1111/mila.12160