PHILOSOPHICAL TRANSACTIONS B

Opinion piece

Author for correspondence:
Oren Kolodny
e-mail: okolodny@stanford.edu

The evolution of the capacity for language: the ecological context and adaptive value of a process of cognitive hijacking

Oren Kolodny[1] and Shimon Edelman[2]

[1]Department of Biology, Stanford University, Stanford, CA 94305, USA
[2]Department of Psychology, Cornell University, Ithaca, NY 14853-7601, USA
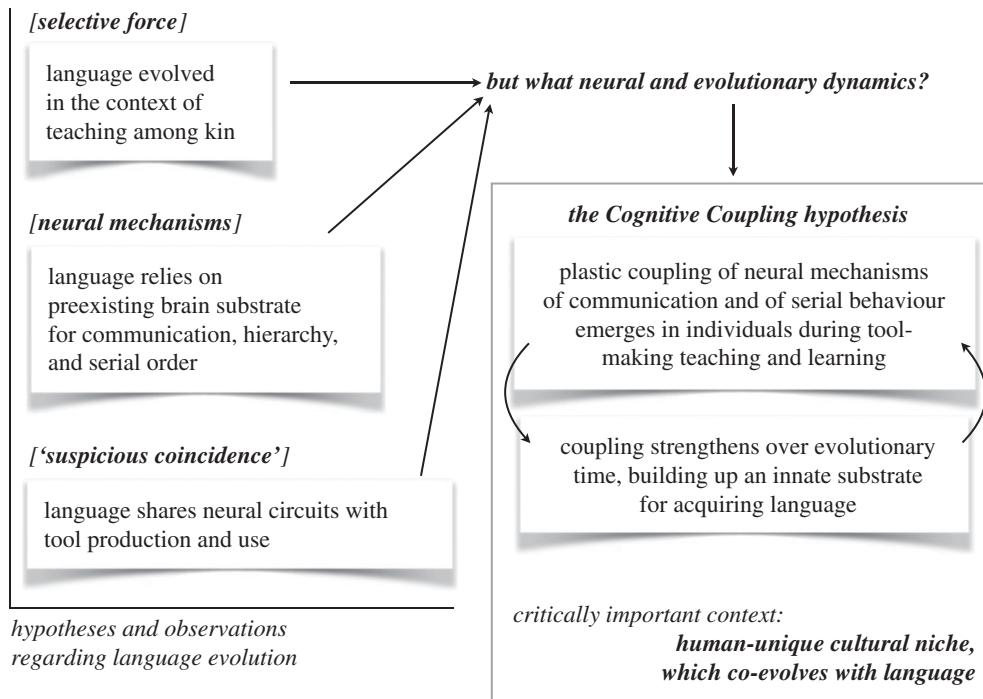
OK, 0000-0002-0095-693X

Language plays a pivotal role in the evolution of human culture, yet the evolution of the capacity for language—uniquely within the hominin lineage—remains little understood. Bringing together insights from cognitive psychology, neuroscience, archaeology and behavioural ecology, we hypothesize that this singular occurrence was triggered by exaptation, or 'hijacking', of existing cognitive mechanisms related to sequential processing and motor execution. Observed coupling of the communication system with circuits related to complex action planning and control supports this proposition, but the prehistoric ecological contexts in which this coupling may have occurred and its adaptive value remain elusive. Evolutionary reasoning rules out most existing hypotheses regarding the ecological context of language evolution, which focus on ultimate explanations and ignore proximate mechanisms. Coupling of communication and motor systems, although possible in a short period on evolutionary timescales, required a multi-stepped adaptive process, involving multiple genes and gene networks. We suggest that the behavioural context that exerted the selective pressure to drive these sequential adaptations had to be one in which each of the systems undergoing coupling was *independently* necessary or highly beneficial, as well as frequent and recurring over evolutionary time. One such context could have been the teaching of tool production or tool use. In the present study, we propose the Cognitive Coupling hypothesis, which brings together these insights and outlines a unifying theory for the evolution of the capacity for language.

This article is part of the theme issue 'Bridging cultural gaps: interdisciplinary studies in human cultural evolution'.

1. The problem of language evolution

Understanding how the capacity for language evolved is a difficult problem, because it is unique to humans, because its computational nature and brain basis are poorly understood, and because the successive stages of its evolution left no fossil record. The problems of the evolution of language and of the cognitive capacities that support it extend over a wide range of disciplines, from comparative ethology to neuroscience, psychology and linguistics, but loom particularly large in evolutionary biology, where the emergence of language has been described as one of a few 'major evolutionary transitions' [1]. What made this transition possible? Answering this question is like finding a missing piece of a puzzle from which many other pieces are also missing, and over which there is no consensus as to how the 'big picture' should look. The shape of the sought-after piece can still be pondered, but to do so meaningfully requires that assumptions be made about the other missing pieces and how they might fit together.

Some of these pieces are highlighted by different questions that may be asked about the origin of language, each portraying a different perspective on the topic.

[selective force]

language evolved
in the context of
teaching among kin

but what neural and evolutionary dynamics?

[neural mechanisms]

language relies on
preexisting brain substrate
for communication, hierarchy,
and serial order

['suspicious coincidence']

language shares neural circuits with
tool production and use

*hypotheses and observations
regarding language evolution*

the Cognitive Coupling hypothesis

plastic coupling of neural mechanisms
of communication and of serial behaviour
emerges in individuals during tool-
making teaching and learning

coupling strengthens over evolutionary
time, building up an innate substrate
for acquiring language

critically important context:
**human-unique cultural niche,
which co-evolves with language**

Figure 1. A schematic illustration of the premises and the reasoning underlying the proposed Cognitive Coupling account of language evolution.

What *is* language? What are the cognitive mechanisms that are involved in language learning and use? What are the roles of cultural exposure and of neural plasticity in the development of an individual's linguistic capacity? What was language's original function? What features of language are common to humans and to other apes' behaviour, and what are different—qualitatively or quantitatively? How much specific evolutionary adaptation was required for language, and how much of it relied on domain-general mechanisms?

In this paper, we address a very specific, yet critically important aspect of the question of language emergence: what was the ecological–behavioural context that made proto-language adaptive and triggered its emergence? We begin by stating and motivating, in §2, our assumptions regarding the nature of language, which differ in important respects from those that underlie the two major classes of linguistic theories. In §3, we discuss some open questions in language evolution in light of our view of language and of the standard methodology in computational cognitive science, which calls for multiple levels of explanation. In §4, we revisit our central question and mention some of the answers that have been proposed for it. Section 5 briefly discusses some of the relevant findings on the brain basis of language. In §6, we lay out our hypothesis, the Cognitive Coupling (COCO) hypothesis: in the context of the teaching of tool use, brain mechanisms involved in hierarchical planning and in sequential control of actions were coupled with brain mechanisms in charge of social communication, giving rise to the capacity for language (figure 1). Section 7 lists evidence that supports this hypothesis. Finally, in §8 we summarize our argument and outline directions for future work.

2. Our background assumptions: what language is 'for' and what it is like

The problem of language evolution looks particularly daunting when seen through the lens of both reigning paradigms in linguistics: the formalist and the functionalist (see [2] for the motivation of these terms). In this section, we briefly state these two approaches, then outline a third, which we favour and which we claim offers a better foundation for an evolutionary understanding of the emergence of language.

(a) The formalist view
On the formalist account, language is a formal system (in the technical sense of the phrase; [3])—a set of generative rules or grammar that licenses a discrete infinity of well-formed strings of symbols or sentences, to each of which it assigns a hierarchical tree-like structure. The primary use of language is taken to be the structuring of thoughts, with its use for communication being considered 'ancillary' [4]. Although both thinking and communication are presumed to involve the generation and analysis of 'meanings', formalist linguistics does not favour a specific approach to semantics, which is seen conceptually as secondary to syntax and which, as a discipline, appears to lag behind theories of syntax (e.g. [5]). For a detailed critical discussion of the formalist approach to language, see Edelman [6,7].

Because the possession of a formal generative grammar implies an ability to generate an infinity of structured sentences that no finite amount of training data can possibly motivate without a 'quantum leap' of generalization,[1] formalists have traditionally shunned questions of evolution altogether, or else resorted to 'singular mutation' accounts (such as the one behind the Merge operation; e.g. [9,10]). Chomsky, in particular, favours viewing the emergence of language as having been due to a unique 'discontinuity' (for a recent recap of that view, see [4]), perhaps acting on his own earlier advice ('It is perfectly safe to attribute this development [of innate mental structure] to 'natural selection', so long as we realize that there is no substance to this assertion'; [11, p. 97]). We note that the continued tendency to regard language evolution as a 'mystery' (as per the title of [12]) is actually a reasonable stance for the formalists, not only because of the problematicity of the all-important

point mutation as the key explanatory move, but also because of the dearth of unequivocal empirical (behavioural and neurobiological) support for the generative hypothesis (see [6], §5a, for a detailed discussion and references).

(b) The functionalist view

On the functionalist account, language is 'for' communication, which makes it inherently social and also elevates semantics and, more generally, concepts and cognition, to a position of primary importance in linguistic theory. Grammar still plays a central role as the tool for encoding meanings into a form that can be easily transmitted via the available gestural (acoustic or other) channel and for decoding the messages on the receiving end. Because they assume that communication between cognitive agents is the overarching goal, functionalist approaches to grammar do, however, tend to be better integrated with theories of cognition, by positing computational mechanisms that are shared between conceptual and linguistic processes (e.g. [13]).

Because functionalist theories still involve the concepts of grammar and well-formedness, they pose the same conceptual problems, and give rise to the same tensions with regard to empirical findings, as do formalist theories (examples include positing intricate rules that involve hidden structures and having to garner behavioural and neural evidence for these rules and structures; see [7] for details and discussion). This problematicity extends to, and is amplified in, the context of the origins of language, where functionalists need to explain not only the emergence of well-formedness (and of course infinite generativity), but also of the kind of code coordination that would make the sharing of meanings possible.

While the evolution of both these traits has been demonstrated in agent-based computational simulations (e.g. [14,15]), these had been geared from the outset towards rewarding effective and properly (i.e. compositionally) structured messaging, leaving open the question of our main concern: the initial ecological context and adaptive value of these traits for agents who prior to the emergence of full-blown language had had no use for either structure or compositional meaning.

(c) Our view

In contrast to the formalist and functionalist approaches, we suggest an alternative view of language and adopt it as a working hypothesis ([6,7]; see [16,17] for similar earlier approaches). It does not postulate the existence of infinitely generative rules (a grammar), nor does it assume that people speak in complete, well-formed sentences that possess a uniquely linguistic kind of recursive structure. Furthermore, we do not accept that language is primarily 'for' either thinking (as per the current Minimalist Grammar version of formalist linguistic theory) or communication, construed as exchanging coded meanings (as per the functionalist approach). Rather, we assume that language constitutes a toolkit that effectively supports an individual in influencing the state and the behaviour of others (and perhaps also self; this notion is discussed, e.g. in [14,18–21]). This system's sophistication transcends anything found in the rest of the animal kingdom, yet its basic computational principles and neural mechanisms, including those that give rise to sequential and hierarchical structure, are closely related to those underlying animal signalling and social interaction.

Importantly, this view highlights the role of language as a means of communication, but—as opposed to some previous

accounts—communication is construed as the interlocutors being able to influence each others' (and their own) cognitive and behavioural processes by linguistic means [22]. This take on communication works not just for the many types of situations that require 'mere' steering of thinking or behaviour in a desirable direction, but also for conveying complex information, such as the design for a complex mechanical contraption or an entire strategic hunting plan. In these latter scenarios, people rely heavily on conceptual—that is, perceptual, abstract and motor—knowledge that their interlocutors possess ahead of time and that need not be (and cannot be) included in the 'messages' that are being exchanged (cf. [16]).

The construal of language as an instance of a system of social interaction suggests that in order to understand the evolution of the capacity for language from, and based on, the cognitive substrates that preceded it, it would be instructive to consider the elements that it shares with communication systems in other animals, particularly apes, which were likely to have been in existence in our most recent common ancestor, as well as those aspects in which language differs from such communication systems (cf. [23]). As noted by Ackermann et al. [24], ape vocalizations lack the kind of sequential (let alone hierarchical) structure that characterizes language; in comparison, the emotional modulation and social uses of vocal gestures is similar in apes and humans, and apes, like humans, are capable of planning and carrying out structured behaviours in domains other than language.

This makes it feasible for the coupling of the capacity for structured behaviour with communication-related traits (such as the intent for affecting the state of others) to have emerged in the hominin line by a sequence of minimal steps (changes in brain circuitry), each of which was adaptive. A relevant observation here is that the primary features of language that distinguish it from ape communication systems are supported by neural mechanisms that play prominent roles in other areas of ape behaviour. In particular, the processing of serial order and hierarchical structure, including long-distance dependencies in space and time, are necessary for behaviours such as primate tool use, navigation, foraging and social action.

Following the hypothesized sequence of adaptations, language became characterized by serial order and hierarchical structure and dependencies, imposed over vocal (or other physical) gestures that may be individually and/or jointly referential [25]. Language also acquired a number of distinctive social characteristics (in addition to those that it shares with other behaviours [26]). Some of these specifically human traits are the intentional use of language to affect others' states and behaviours [27], while tailoring the verbal means to the recipient; propensity for extended dialogic interaction [28]; a vast, learning-intensive, socially acquired component that dwarfs its innate foundation [29]; and a critical role in humans' capacity for innovation [30]. Importantly, elements of each of these, from the so-called *theory of mind* to a combination of innate, learned, and invented behaviours, can be found in non-human ape communication, but are severely constrained and limited compared to their ubiquitous role in humans [31–34].

(d) Implications for the questions about language evolution

Giving up the formalist conceptions of grammar and well-formedness makes it easier for us (as it would for certain

4

other functionalist approaches) to resolve other problematic ideas associated with these conceptions. One such idea that we can safely set aside is that babies learn infinite productivity and structural perfection—traits found in no other species—from finite data; another one is that our species evolved such unique traits from scratch. Furthermore, because we hold that the basic use of language is social in a manner that allows for, but does not necessitate, the exchange of perfectly structured information, our approach is amenable for serving as a bridge between, on the one hand, theories of animal communication and its evolution, and, on the other hand, linguistics and the evolution of language.

3. Structuring the open questions about language and its evolution

With the literature on language evolution growing apace (see the recent special issue of *Psychonomic Bulletin & Review* for a sample, [23], as well as [35]), the range of research questions that are being entertained is very broad. To impose some structure on these, we invoke a methodological consideration that has long been standard in evolutionary, behavioural and cognitive sciences: the notion of levels of explanation. In evolutionary biology, these include Mayr's distinction between proximate and ultimate causes [36,37]. In ethology, there is Tinbergen's [38] fourfold distinction among questions of survival value, ontogeny, evolution and causal powers of a trait. Finally, in computational cognitive science, there are the three levels of explanation identified by Marr & Poggio [39–41]: the levels of problem; representation and algorithm; and implementation (see [42] for a discussion of evolution in this context).

The various levels of explanation are not independent: as one of us pointed out in connection to vision [39,43] and language [6,7], a specific assumption made on one level has implications, sometimes negative, for the directions in which inquiry on other levels proceeds. With regard to language, in particular, getting the problem-level answer or assumption wrong can prevent us from understanding how it works in the computational sense, or how it has evolved [6].

The interdependence of the levels of explanation suggests that the answers to the many questions arising from language evolution are also connected to one another. For instance, a saltatory as opposed to a gradualist account is made more plausible by assuming that there is one key feature (such as the Merge operation that purportedly allows recursion; [4]) that makes language what it is (and different from other animal signalling and communication systems). Given that Merge and recursion, as well as the broader notion of grammar, have no bearing on the question of how gestures can be, or have become, referential, it is not surprising that the formalist linguistic theory built around these concepts posits that language is primarily 'for' internal thought, largely skirting the problem of reference. Rejecting, as we do, the formal notion of grammar (and the conceptual centrality of Merge) thus leads to a completely different set of takes on the questions of saltation/gradualism and reference.

Adopting the popular alternative assumption regarding what language is 'for'—communication, as per functionalist linguistics—suggests other answers to these questions and brings to the fore yet another set of issues. In particular, it becomes critically important to determine how language relates to honest signalling [44–46]. Furthermore, given that communication is by definition a social/cultural activity, understanding the interaction between the genetic evolution of language users and the cultural evolution of language becomes key, giving rise, in turn, to questions about the relation between the evolution of the cognitive capacity for language and the ways in which language evolves (cf. Hurford's [47] discussion of the notion of 'glossogeny', which dates back to Jacobson and earlier work in linguistics). It also brings to mind queries as to which aspects of language acquisition are developmental, dependent on cultural exposure, and which are innate, a product of evolutionary adaptation.

As already noted, our take on language diverges from the standard functionalist one (albeit not so drastically as it differs from that of formalist linguistics). On the abstract computational level, our view holds that language users aim to influence the cognitive state, and accordingly the behaviour of others (and of themselves, when resorting to internal speech). This implies that the key computational problem that the cognitive system needs to address is choosing what to say next (note that this is a subset of the general problem in the control of behaviour, which is deciding what to do next; [7,48]). On the level of representation and algorithms, language production is supported by a system of options structured like a graph, whose nodes correspond to the discrete combinatorial elements of language—phonemes on one level; morphemes or words on another [49–51]. Paths through this graph, which correspond to the (multimodal; [48]) utterances, are constructed by the speaker 'on the fly' during production, by an algorithm that is akin to competitive queuing [52], subject to a number of dynamically applied constraints, which include real-time social feedback; the listener's own version of the graph mediates the processes (and the resulting effects) on the receiving side. Finally, on the level of neural implementation, the use of language, similar to any other behaviour, involves activity that is distributed over most of the brain; some particularly relevant circuits are discussed in §5 below (see also [7]).

We note that our problem-level view of the nature of language places it much closer than other accounts to other behaviours, including those of other species; the same holds for the algorithms that shape language behaviour and the brain circuits that implement them. Thus, we expect our approach to be more amenable to integration with evolutionary theory than formalist and functionalist ones. With the just-stated conception of language in mind, we now turn to the specific question of the adaptive value of language and the context of its emergence.

4. The question we address: what could have been the ecological–behavioural context that made proto-language adaptive?

A popular way of framing this question is by focusing on 'the adaptive value of the first word' ([53, p. 165]). Assuming, as we do, that language is 'for' influencing others [14,16,18,20] implies that the context we are interested in must be a social one, but this leaves open a rather broad range of possible ecological–behavioural contexts. A principled exploration of this range may be facilitated by the explicit

layout of evolutionary considerations that need to be taken into account (e.g. [54]).

First and foremost, the ecological–behavioural context within which the capacity for language evolved (henceforth *the context*) was necessarily one whose occurrence did not rely on the pre-existence of language. In other words: although once language existed it may have turned out to be useful, and perhaps necessary, in many contexts, the fundamental setting in which language evolved must have been one that initially was not facilitated by language at all.

Second, this context must have been one in which alternative outcomes of the involved individuals' behaviour had consequences for their fitness, and this context had to recur on a frequent basis, in the lives of many, and over multiple generations. In order to drive evolutionary change that would give rise to language, these outcomes must have been influenced by the extent to which communication between the individuals was successful.

Third, the context we seek is one in which the interests of the involved individuals are largely aligned. This stems from the fact that language does not have an inherent mechanism that assures honest signalling [44], and it seems highly unlikely that such an elaborate communication system, which is demanding of all individuals involved, and which has the potential to be easily used for deception, would evolve outside of a context in which all involved stand to benefit greatly from successful communication and transfer of knowledge. Although not exclusively, a situation in which most communicators are kin seems the most likely in this context [54–56].

Finally, although not absolutely necessary, it seems reasonable to assume that the context that gave rise to language, given language's flexibility, open-endedness and being itself a cultural construct, would have been one that would change at a rate that requires cultural transmission of knowledge, perhaps because genetic adaptation to the environmental challenge would be too slow. This further suggests that the setting may be one in which niche construction and co-evolution of the context and the language that is involved in its facilitation may take place. It also seems likely that steps along the process occurred in the form of a Baldwin effect: that plasticity in the use of the communication system supported successful coping with the challenge provided by the ecological context, and that selection favoured individuals who happened to be better cognitively pre-adapted to this means of communication, gradually giving rise to the innate capacity for language that characterizes modern humans.

5. What we can learn from how language is implemented in the brain

As is the case with other behaviours, language engages a broad coalition of brain areas and circuits ([57–60]; for a summary, see [7], §5). Although this view is supported by a large and growing set of neuropsychological findings (in addition to the general evidence for massive sharing of brain areas across tasks; [61]), these are commonly still given a conservative interpretation. This interpretation distinguishes between 'core' language areas (such as Broca's), where a lesion results in a major loss of function, and 'optional' ones (e.g. the cerebellum), where lesions may seem to have little, if any effect on normal functioning.

A closer consideration reveals, however, that both these observations are open to challenge. First, lesions in a core language area typically result in a loss of other functions in addition to language (e.g. [62] on Broca's area lesions). Second, lesions in other areas may impair language in ways that standard tests can easily miss (e.g. [63] on cerebellar lesions).[2] These observations suggest that the brain basis of language is much broader than commonly thought—a notion that is entirely in line with Anderson's [61] Neural Reuse hypothesis, according to which 'neural elements originally developed for one purpose are put to multiple uses'. (See also the broader notion of evolutionary exaptation, [64–66].) Language evolution, in particular, is thus seen as having been facilitated by the hijacking of existing brain mechanisms, which came to be reused for new tasks, while retaining their old use [57,67,68].[3]

The category of tasks to which Broca's area, in particular, makes a critical contribution offers a hint as to how such hijacking may have occurred. In addition to its classically documented involvement in the structural (syntactic) aspects of language, Broca's area supports hierarchical structural processing that may or may not be temporal, as, for instance, in task decomposition [62,69]. As such, it serves as a functional hub whose capacity for hierarchical computation can be used both for planning sophisticated action selection for manual tasks as well as sophisticated communication.[4] Furthermore, there is abundant evidence suggesting that action selection and the temporal sequencing of elementary actions is mediated by circuits linking cortical areas (including those in the frontal lobe, such as Broca's) and the thalamus with the basal ganglia (e.g. [52,70]). It has been proposed that these circuits are central to language (e.g. [71–73]). Indeed, the striatum, which is the first destination of the cortical projections to the basal ganglia, appears to be absolutely necessary for language acquisition [74,75].

A detailed two-systems view of language that focuses on the above circuits has been advanced by Ackermann *et al.* [24], who posit one set of mechanisms for the 'digital' content of utterances and the other for the 'analogue' prosodic/affective information that they include. They argue that, although most of the components of each of these two systems are present both in humans and in other primates, non-human primates lack both the kind of coordination that allows independent control over the two, as well as some circuits that may be crucial.

The emergence and subsequent epigenetic and genetic entrenchment of any new circuitry[5] would have been particularly favoured by selection if the early hominin brain was being 'prepared' for it by having the relevant *functional connectivity* ramped up by the behavioural contexts and tasks involving proto-language (cf. the Baldwin effect; [76–78]). Interestingly, in modern humans, learning to read—a task that came into existence recently in human evolution and is unlikely to have had time to significantly influence human cognition via selection on genetic variants—causes, over a period of just several months, a significant shift in the pattern of functional connectivity in the brain [79], demonstrating the potential impact of a novel task that is accommodated by pre-existing circuitry.[6] Similarly, with regard to manual behaviour, changes in functional connectivity were also found in the brains of human subjects who underwent several months of training in stone tool knapping [82].

The preceding observations may help us understand how the human capacity for language could emerge from the primate-general mechanisms. The hypothesis that the brain basis of language was shaped by the hijacking/reuse of existing circuits, which already had well-established uses outside of the context of communication, may be considered in conjunction with two specific such mechanisms (over and above the general pattern of brain connectivity, such as the centrality of the pulvinar, noted in endnote 6). The first is the utilization of the cortical (prefrontal and motor)/basal ganglia circuits for learning visuo-motor sequences [83]—a function that is critical for, among other tasks, the learning of tool use. The second is the use—indeed, the indispensability—of the same circuits for reinforcement learning, which aims to maximize cumulative expected reward [84–86].

In light of these considerations, we may supplement our list of evolutionary considerations with the requirement that the proposed ecological–behavioural context for the evolution of the capacity for language was a setting that would have supported co-option and temporal coupling of the circuits involved in temporal sequencing and hierarchical processing with those involved in communication. Such a context *must* be one that does not initially rely on the coupled action of these mechanisms in order to exist, but which is likely to reliably prompt their simultaneous use again and again throughout an individual's lifetime and across generations. This simultaneous use would have had to be advantageous, such that individuals in whom the mechanisms tended to couple with one another more readily had an advantage.

One scenario that meets these requirements is the *teaching of tool use*—a task whose brain basis, perceptual/motor and social/communicative aspects are close enough to those of language to suggest that it may have made protolanguage adaptive [54,87], thereby facilitating the reuse and temporal coupling of brain mechanisms and driving the evolution of language as we know it. Interestingly, this coupling may have been facilitated by the dual role of the cortical/basal ganglia circuits both in perceptual-motor behaviour and in reinforcement learning, since feedback for effective dyadic communication and successful learning would provide immediate social reward alongside the attendant advantages of social transmission of knowledge. In the next section, we describe this scenario in some detail and discuss its implications.

6. The hypothesis: cognitive coupling in the context of tool production

(a) Hypotheses about the origin of language

The problem of explaining the origin of language has been eloquently posed by Premack [88, p. 282]:

> Human language is an embarrassment for evolutionary theory because it is vastly more powerful than one can account for in terms of selective fitness. A semantic language with simple mapping rules, of a kind one might suppose that the chimpanzee would have, appears to confer all the advantages one normally associates with discussions of mastodon hunting or the like. For discussions of that kind, syntactic classes, structure-dependent rules, recursion and the rest, are overly powerful devices, absurdly so.

A number of suggestions have been put forth regarding the origin of language, framing the question in various ways and accordingly focusing on different aspects that may be related to it [23,45]. One widely cited model, *the gossip theory*, suggests that language evolved in the context of increasing hominid group size, allowing this growth by providing an efficient means to exchange information about non-present individuals [89,90]. According to this approach, this is a necessity that arises when the group reaches a size at which most individuals cannot spend the majority of their time with one another.

Other approaches, along similar or related lines of reasoning, suggest that language emerged as a means of communication that would replace one-on-one grooming or would increase group cohesion through ritual [46,89,91]. Another model, the *hunting coordination theory* [92,93], suggests that language emerged as a means of coordinating complex activities that demanded planning ahead of time and during which the participants could not communicate due to limited physical proximity. Yet another model suggests that language, similar to bird song and courtship displays, evolved as a response to pressures of sexual selection or for the establishment of social status ([94], and see [95,96]).

A detailed discussion of these ideas is beyond the scope of the present paper. We believe that each of these theories has some merit and the mechanisms they posit may have had a role in some phase of language evolution, particularly once language had already developed. However, each of them fails to account for the origin of language, for at least one of a number of reasons. The majority of these theories propose *ultimate* explanations for language, i.e. they explain what its use may have been once it emerged, but ignore the proximate mechanism by which language could have evolved, in such a manner that every step in the process of its evolution was independently advantageous. Thus, it may be useful to talk about a person, object or ambush plan that are out of sight, but it is unclear how such an ability could have gradually evolved [54].

Another point of weakness in some theories such as sexual selection is that it is unclear how and why language would become referential, or why these selective forces, similar to those that many species experience, would lead to language in humans but not in other organisms (discussed in, e.g. [23]). We suggest that the scenarios regarding the origins of language that have been proposed thus far either do not take into account the evolutionary considerations outlined in §4 or do not fulfil the requirement that the context include simultaneous use of brain circuits involved in temporal sequencing of behaviour and in communication.

An ongoing debate is whether language evolution included an early phase of gestural communication that was later replaced by vocal utterances [20,31,92] or whether it arose directly from vocal communication skills that preceded it (discussed in [32]). It has been pointed out that the dichotomy between gestural and vocal origins may be false, because both may have taken place contemporaneously [67]. We return to discuss this topic below, in light of the scenario for the origin of language that we propose.

(b) A possible scenario

Here is how the sequence of events underlying and driving language evolution could have proceeded, according to the scenario we just outlined. At the beginning of the evolutionary

transition in question, our hominin ancestors possessed brain mechanisms that enabled them (a) to engage in the hierarchical planning and sequential control of actions and (b) to perform elementary communication acts consisting of isolated manual and vocal gestures. The main components of the mechanisms supporting (a) and (b) were homologous, respectively, to the cortical/basal ganglia circuits mentioned above and the primate-general 'limbic communication system' [24].

One kind of behaviour that is enabled by hierarchical planning and sequential control and may benefit from concurrent communication is instruction in tool-making and tool use [54,97]. We propose a process that plays out on two timescales. The first occurs along an individual's lifetime: individuals who learned and taught tool-making while at the same time engaging in elementary communication learned and taught more efficiently, produced better tools and enjoyed a selective advantage. Developmentally, this would have led to the coupling of the circuits involved (as per Hebb's idea that neurons that 'fire together, wire together'). Such yoking of the two sets of mechanisms is likely to have occurred unconsciously, reinforced directly by the increase in instructional success. Over an evolutionary timescale, natural selection favoured the lineages of individuals in whose brains the coordination between these two activities was more readily facilitated by changing functional connectivity. Since variation also exists and occasionally arises in the innate aspects of brain architecture, selection also would have favoured lineages in which greater coupling of communication and serial order was genetically determined via newly emerging circuitry.

(c) Some remarks

Related ideas regarding the context in which language arose have been put forward previously, albeit not in a form that links them with the requirement of co-occurrence of communication and temporal sequencing of actions, which could have led to the language-unique coupling of brain circuits. Laland [54] and others [87,97] proposed, on the basis of evolutionary considerations that highlight honest signalling and cultural niche construction, that the context of language evolution was in teaching among kin. Stout and others [67,98,99] hypothesized that tool-making and its instruction gave rise to language; they provide an elaborate analysis of the neural correlates of the two behaviours and use experiments to demonstrate the effects of learning stone knapping on functional connectivity in the brain and of tool production on neural activation.[7] We endorse these accounts, and suggest that combining them leads to a powerful working hypothesis regarding the origins of the capacity for language.

The account we proposed is in line with the evolutionary considerations we have laid out; importantly, it also provides an explanation for the fact that language evolved only in humans, as an outcome (and perhaps also a driver) of humans' increasing reliance on a culturally constructed niche for their survival. This is a context that changes rapidly, relying on learned knowledge, in which teaching and effective information transfer are highly advantageous. Such learning and teaching requirements, coupled with the anchoring of some of the most critical human skills in the physical realm of sequential behaviour in tool production, offer a natural context that bridges communication and sequential and hierarchical processing, which are largely separate from one another in other apes. It also accounts for the referential nature of

language and suggests a way in which it could have arisen gradually, as each incremental improvement in the ability to refer to seen and—later—unseen aspects of the demonstrated process would have been independently advantageous. This is particularly true given that this cultural context is distinguished by the need for precision in some of its components.[8]

The debate regarding the hypothesis of the gestural origin of language can now be reconsidered. The ecological context of language origin that we propose assigns a major role to physical manipulation and motor behaviour, the fundamental tenets of the gestural hypothesis [20,21]. Importantly, these are called upon in our scenario, at least in the primal phase of language evolution, in a functional context without a communicative intent, namely in the service of producing or manipulating tools. However, the recruitment of gestural behaviour, alongside vocal communication, would have been natural and probably occurred early on, for instance in the form of repeating a functional manual gesture multiple times for emphasis, or miming a functional manipulation before it takes place to aid the learner in parsing and interpreting the sequence of actions. In this context, not only are the gestural and vocal approaches not mutually exclusive, they complement each other. The joint involvement of both modalities in the emergence of language is, according to our proposal, to be expected. This notion is also consistent with the multimodal nature of language [48,102].

Although alternative hypotheses of the origin of language invoke different behavioural settings, it would be naive to suggest that any single ecological context such as the teaching of tool production to kin was the exclusive driver of language evolution. First, the teaching of tool production is but one of many social behaviours that share features that we hold to be crucial in supporting the evolution of language. Coordinated manipulation of objects, as in constructing a shelter or using a large hunting net, for example, is one of these (see discussion in [103]). Second, multiple contexts that an individual experiences throughout life are likely to have shaped each of the systems involved in the emergence of language. Thus, for example, the cultural practice of tool use and production, which requires the learning and honing of very precise, highly intentional and sometimes non-intuitive, motor manipulation, and which had probably evolved in the hominin lineage for millions of years before the emergence of language [21], is likely to have primed and selected for brain mechanisms that later—through coupling with the communication circuitry—gave rise to language. Finally, once language or protolanguage began to develop and as it developed, it would have been rapidly applied to additional contexts, which—in turn—would have added to the sum of selective pressures that act on this composite trait [104]. We suggest that the context of teaching tool production and use served as a primal and primary context for the evolution of language and the capacity for it, but additional contexts necessarily played a role even in the early stages of language evolution and even more so as this capacity evolved.

7. The case for tool-making instruction as a key facilitator of the emergence of language

We now proceed to list, and briefly discuss the significance of, sources of evidence in favour of the COCO hypothesis, drawing on studies in archaeology, cognitive and

developmental psychology and neuroscience. Many of the points that we make here recap material already mentioned in §§2–5.

(a) Arguments from archaeology

Language does not fossilize in a manner that leaves tangible archaeological footprints. Accordingly, any link between language and archaeological findings is inferential, and is often deep in the realm of speculation. Various artefacts have been viewed as evidence of complex cognitive abilities and effective means of cultural transmission; some researchers have postulated that certain findings, such as composite tools, cave paintings and artefacts suggestive of symbolic thought such as beads, were unlikely to have been produced in the absence of language [93,105–108]. Even if such inference is taken at face value, it provides, at best, a limit regarding the time frame in which language may have emerged [107,108], but does not tell us much about the context of its origin.

It has been argued that morphological changes in the vocal apparatus or brain structures, whose existence may be inferred from the skull's shape, can serve as evidence regarding language use [107]. We find this unlikely, in light of the fact that language can be carried out fully over a gestural modality as in sign language, and that multiple brain regions are involved in language use while no brain region serves language alone.[9]

However, a new discipline—experimental archaeology—is rapidly bridging the gap between the archaeological record and the neural and behavioural dynamics that it may reflect. This approach involves the study of apes (human and non-human) as they manipulate and produce tools and while they teach and learn these skills [97,110]. Aided by neuroimaging technology, these studies explore the brain mechanisms involved, and, coupled with studies of comparative neuroanatomy, highlight the changes in brain activation and connectivity patterns in response to and in support of tool use, on timescales ranging from seconds to years [82]. They also allow comparison between the brain structures involved in tool use by humans and other apes, fleshing out the possible results of natural selection on this behaviour [111,112].

These studies have provided some of the evidence of the overlap discussed earlier between brain mechanisms that support tool use and those that support language. They are also in line with the idea that producing certain tools, but not others, might be associated with qualitatively different cognitive mechanisms, and that these mechanisms, in turn, may be in accord with a related linguistic capacity: Stout *et al.* [110,113,114] find that Oldowan tool production requires the activation of praxis-oriented neural circuitry in both humans and bonobos, while the production of late Acheulean tools incorporates, in addition to these, circuitry associated with hierarchical organization, abstract action representation, semantic/syntactic integration and long-distance syntactic dependencies. This makes much sense, as the production of Acheulean tools requires the achievement of technological sub-goals, which are interdependent in a manner similar to syntactic dependencies. They also discovered that the same circuits are activated upon observing another individual's production of Acheulean tools, and that in experienced stone knappers who observe another individual's work, neural circuitry associated with intention attribution is activated [67]. We suggest that although these findings cannot prove that language evolved in the context of

technological pedagogy, the coordinated neural activation of the primary elements of linguistic communication, sequential and hierarchical processing and inference of another individual's internal state, during tool production or observation, constitutes compelling support for this hypothesis.

(b) Arguments from cognitive/computational psychology

The behavioural contexts of communal tool use and tool-making jointly activate cognitive mechanisms that are also employed in a range of other situations and tasks, both in humans and, to various extents, in other primates. These include, on the one hand, sequentially and hierarchically structured action control and, on the other hand, communication and social manipulation (see [7], for behavioural and computational analyses, concrete examples of relevant tasks, and extensive references). Socially shared tool-related activities, and especially tool-making instruction, are therefore prime candidates for the kind of context that first brought together previously unrelated abilities that had existed in the hominin line, making their integration and the subsequent emergence of modern capacity for language initially and continually adaptive.

(c) Arguments from developmental psychology

The tool-making pedagogy hypothesis of language origin fits naturally with the view of language as a skill (e.g. [115,116]). None of the skills that distinguish humans from other animals—with the exception of the skill to *learn* language and other complex behaviours—are available at birth: all must be learned. Just like the learning of other complex skills, language acquisition involves social interaction [117–119] and possibly even instruction [120]. The presumably innate foundation over which the specifically linguistic skills are built during ontogeny, the ease with which humans combine communicative influence and reference with complex sequential hierarchical structuring of actions, may be considered an 'evolutionary fossil' of the social skill-transmission context in which the two components of this combination first came together. A less indirect type of evidence for the existence of a common neural substrate for gesturing, object manipulation, tool use and communication can be obtained by studying the trajectories of development of these behaviours throughout childhood, with particular stress on early phases of their acquisition (e.g. studies discussed in [98]).

(d) Arguments from neuroscience

The evolutionary convergence, as well as the behavioural, computational and developmental commonalities between language and other advanced human skills, would be difficult to explain were the brain basis of language confined to the cortical Broca's and Wernicke's areas, as most introductory textbooks still have it. The understanding that emerges from dozens of studies is, however, that most brain areas, both cortical and subcortical, participate in supporting most tasks that have been considered [61], with language processing being no less distributed ([58,121]; see [7], §5, for a brief synthesis). The critical question is, of course, what made the brain mechanisms underlying various relevant behaviours that had been in place prior to the emergence of (proto-) language particularly (and, among all species, uniquely) suitable for supporting such emergence. The answer, based on the discussion of the brain

basis of language in §5 above, is twofold. First, the relevant behaviours, as per the tool-making pedagogy hypothesis, included sequential hierarchical planning and control, together with social engagement and attempted influence or 'communication'. Second, the circuits that controlled these behaviours (as mentioned in §5) became functionally interdependent, a development that facilitated their subsequent greater anatomical affinity, as documented in modern humans (see [67] for a detailed survey of the relevant circuits and a discussion of the significance of the anatomical and functional findings for the 'technological' hypothesis of language origin).

8. Predictions and summary

A number of predictions can be derived from the COCO hypothesis, which may direct analysis and interpretation of empirical findings and can suggest directions for further exploration.

(1) We predict structural similarities (of the kind that exists between tutor and juvenile songs in the zebra finch; [122]) between (i) transcripts of manual skill pedagogy and (ii) child-directed speech (CDS). In particular, various structural characteristics of CDS, such as the prevalence of variation sets (e.g. [123]) are known to change over time as infants grow older and become more proficient in language (e.g. [124]). Given the parallels between linguistic development and skill instruction mentioned earlier, we expect a similar developmental trajectory to be found in both cases. Ideally, both language and skill pedagogy episodes should be analysed in their full multimodal complexity, in full detail [125], and using state of the art computational tools [126].

(2) If some component of the cognitive coupling has not become innate and remains dependent on experience, we should expect that exposure to linguistic sequences that are coupled with dynamics in the visuo-motor modality, such that these circuits' activation is coordinated, would be a particularly effective way of supporting language learning. This may be testable experimentally.

(3) The COCO hypothesis is founded on the assertion that communication that is coupled with tool use would facilitate its social transmission. This can be tested in experiments, particularly with children. Dimensions of interest would include whether (and how) children choose to incorporate coupled communication in their demonstration of a physical skill, when instructed to teach the skill or to help another child acquire it; whether doing so increases transmission efficacy; how communication that accompanies the physical demonstration changes along transmission chains that include multiple children; what kind of communication is added, e.g. gestural or vocal, linguistic or non-linguistic; and how this changes as a function of age, earlier experience or the details of the instructions provided.

(4) It may be useful to study the manner in which human cognition accommodates evolutionarily novel tasks other than language and how new skills emerge when provided with a cultural context that makes them beneficial or that reinforces them socially. Such are the skills related to reading and writing, complex mathematics and the perception of virtual reality settings and the behaviour in them (e.g. [127,128]). In particular, it may be interesting to study whether different timing of the exposure to these settings along the individual's cognitive developmental trajectory affects the manner in which they are accommodated by neural circuitry.

(5) The COCO hypothesis ascribes a prominent role in language to coupling between neural circuits, which may occur with greater ease if linguistic exposure begins early in cognitive development. In this light, perhaps some differences between first and second languages, which are learned at a later age, stem from different efficacy of coupling between circuits that were evolutionarily involved in non-linguistic communication and circuits related to serial order and hierarchical processing. This possibility suggests that there should be prominent differences between first and second languages in the ease and fluency with which individuals process and produce utterances that include complex or high-order syntactic dependencies. An indication that this is the case may be found in existing studies (e.g. [129,130], but see also a different result, perhaps due to the influence of explicit language instruction, in [131]), but targeted exploration of this possibility may yield interesting insight.

To summarize, our hypothesis on the origin of language, the COCO hypothesis, rests on the notion that language is a means by which individuals influence one another's state, through communicating information in a manner that is characterized by serial order and hierarchical structure, imposed over referential vocal (or other physical) gestures. This view of language highlights its reliance on two sets of pre-existing neural mechanisms, those involved in communication and those involved in serial behaviour, giving rise to the central tenet of our theory: that the coupling of these mechanisms triggered the emergence of language. Adopting an evolutionary perspective in searching for a plausible proximate mechanism and an ecological context in which language could have emerged, we suggest that the most likely setting that induced this neural coupling was instruction in tool use and tool production.

Our COCO hypothesis provides an important missing link connecting three previously suggested takes on language and its evolution: theories that propose that language evolved in the context of teaching among kin [54,67,87,97]; theories that highlight the reliance of language on pre-existing neural substrates of communication and of hierarchical and serial order processing (e.g. [57,61,98]); and observations showing that these neural circuits are common to language and to tool production and use (e.g. [67,98]). These three takes, respectively, suggest a selective force that could favour efficient communication, provide an invaluable mechanistic description of the solution that evolved in response to this challenge and point out the 'suspicious coincidence' that is inherent in the sharing of neural circuits between language and tool production. However, even when taken together, these ideas leave open the question of the concrete step by step evolutionary dynamics that (i) plausibly led to language as we know it and not to some other solution to the evolutionary challenge,[10] and (ii) would offer a causal explanation of the 'suspicious coincidence' on the level of neural implementation.

The COCO hypothesis fills this lacuna: it proposes an account of specific dynamics that would have taken place while teaching a structured skill such as tool production and

that would include *neural coupling* of existing initially independent neural mechanisms—a coupling that occurred in individuals during development, via neural plasticity. Heritable variation among individuals in the efficiency of this coupling would have allowed natural selection to gradually increase the *innate* coupling between the mechanisms or select for factors that increase the rapidity of the developmental coupling. This would eventually give rise to a communication system that, completely *non*-coincidentally, shares neural correlates of structure and serial order with tool production: human language.

Finally, the human-unique reliance on a cultural niche in which interaction with the environment is heavily mediated by tools, and which entirely depends on culturally acquired knowledge, may explain why language evolved only in the hominin lineage. Language not only stems from this unique cultural niche; it also supports it, suggesting that the two have co-evolved in a tight loop of positive feedback, playing a prominent role in the shaping of humanity as we know it. The emergence of language from the humble origins of simple tool production has thus set off the process that boosted the cognitive ability of one species over all others and has allowed it, for good and for bad, to transform the face of the planet.

Endnotes

[1]Cf. Chomsky [8, p. 380]: 'Human language is based on an elementary property that also seems to be biologically isolated: the property of discrete infinity, which is exhibited in its purest form by the natural numbers 1, 2, 3, . . . Children do not learn this property of the number system. Unless the mind already possesses the basic principles, no amount of evidence could provide them; and they are completely beyond the intellectual range of other organisms.'

[2]Cf. Koziol *et al.* [63, p. 156]: 'Expressive language impairments include word finding difficulties and abnormal syntax with agrammatism, long latency and brief responses, reluctance to engage in conversation. Verbal fluency is decreased, affecting phonemic (letter) more than semantic (category) naming. Mutism occurs following acute injury such as surgery involving the vermis, mostly in children but also to varying degrees in adults. Poor control of volume, pitch and tone can produce high-pitched, hypophonic speech.'

[3]Neural reuse also occurs, necessarily, in dealing with tasks that did not accompany our species throughout most of its evolutionary history and that are unlikely to have had time to affect our innate brain structures. Examples of such tasks are reading and writing, carrying out complex mathematical calculations and accommodating settings such as virtual reality, from watching television to actively engaging in it, as in gaming.

[4]As in 'animal communication'; our use of the term does not constitute an endorsement of the construal of language as message-passing.

[5]The specific hypothesis of Ackermann *et al.* [24], which we do not necessarily endorse, is that the new circuit consisted of a monosynaptic connection between cortical 'language' areas and the motor area that controls the laryngeal muscles.

[6]Notably, in addition to the cortex and the midbrain visual attention circuits, changes following learning to read were found in the pulvinar—a higher-order thalamic nucleus, which is reciprocally connected to the entire cortex [80], as well as to many subcortical areas, including the basal ganglia [81], and which can therefore mediate large-scale functional plasticity that eventually becomes translated into new connections.

[7]Cf. Engels [100], writing in 1876: 'First labour, after it and then with it speech—these were the two most essential stimuli under the influence of which the brain of the ape gradually changed into that of man.'

[8]Notably, hypotheses which highlight the advantages conferred by propensity to learn and by social status associated with linguistic proficiency, are compatible with this scenario [101].

[9]Additionally, recent findings show that previously, the productive abilities of monkeys' vocal tracts significantly underestimated their vocal capacity. In fact, monkey vocal tracts appear to be speech-ready [109].

[10]That some of the subject matter of communication during teaching was structured (tool production) does not imply that the communication itself had to be structured identically. Conceivably, for example, a communication system could be improved through an increase in referentiality, without incorporating structure. Even communicating about hierarchically structured tools is possible using non-hierarchical referential sequences. Such communication may indeed have been an intermediate phase of language evolution [132].

References

1. Szathmary E, Smith JM. 1995 The major evolutionary transitions. *Nature* **374**, 227. (doi:10.1038/374227a0)

2. Newmeyer FJ. 2017 Form and function in the evolution of grammar. *Cogn. Sci.* **41**(Suppl. 2), 259–276.

3. Smullyan R. 1961 *The theory of formal systems*, Annals of Mathematics Studies, vol. 47. Princeton, NJ: Princeton University Press.

4. Everaert MBH, Huybregts MAC, Chomsky N, Berwick RC, Bolhuis JJ. 2015 Structures, not strings: linguistics as part of the cognitive sciences. *Trends Cogn. Sci.* **19**, 729–743. (doi:10.1016/j.tics.2015.09.008)

5. Pietroski PM. 2003 The character of natural language semantics. In *Epistemology of language* (ed. A. Barber), pp. 217–256. Oxford, UK: Oxford University Press.

6. Edelman S. In preparation. Verbal behavior without syntactic structures: beyond Skinner and Chomsky. In *Chomsky's legacy* (ed. C. Behme)

7. Edelman S. 2017 Language and other complex behaviors: unifying characteristics, computational models, neural mechanisms. *Lang. Sci.* **62**, 91–123. (doi:10.1016/j.langsci.2017.04.003)

8. Chomsky N. 2004 Language and mind: current thoughts on ancient problems. In *Variation and universals in biolinguistics* (ed. L Jenkins), pp. 379–405. Amsterdam, The Netherlands: Elsevier.

9. Bolhuis JJ, Tattersall I, Chomsky N, Berwick RC. 2014 How could language have evolved? *PLoS. Biol.* **12**, e1001934. (doi:10.1371/journal.pbio.1001934)

10. Chomsky N. 2005 Three factors in language design. *Linguist. Inq.* **36**, 1–22. (doi:10.1162/0024389052993655)

11. Chomsky N. 1972 *Language and mind*. Cambridge, UK: Cambridge University Press.

12. Hauser MD, Yang C, Berwick RC, Tattersall I, Ryan MJ, Watumull J, Chomsky N, Lewontin RC. 2014 The mystery of language evolution. *Front. Psychol. Front.* **5**, 401. (doi:10.3389/fpsyg.2014.00401)

13. Langacker RW. 2016 Working toward a synthesis. *Cogn. Linguist.* **27**, 465–477. (doi:10.1515/cog-2016-0004)

14. Scott-Phillips TC, Kirby S. 2010 Language evolution in the laboratory. *Trends Cogn. Sci.* **14**, 411–417. (doi:10.1016/j.tics.2010.06.006)

15. Kirby S. 2017 Culture and biology in the origins of linguistic structure. *Psychon. Bull.*

Rev. **24**, 118–137. (doi:10.3758/s13423-016-1166-7)

16. Ramscar M, Baayen H. 2013 Production, comprehension, and synthesis: a communicative perspective on language. *Front. Psychol.* **4**, 233. (doi:10.3389/fpsyg.2013.00233)

17. LaPolla RJ. 2015 On the logical necessity of a cultural and cognitive connection for the origin of all aspects of linguistic structure. In *Language structure and environment: social, cultural, and natural factors* (eds R De Busser, RJ LaPolla), pp. 31–44. Amsterdam, The Netherlands: John Benjamins.

18. Sperber D, Origgi G. 2012 A pragmatic perspective on the evolution of language. In *Meaning and relevance* (eds D Wilson, D Sperber), p. 331. Cambridge, UK: Cambridge University Press.

19. Sperber D, Wilson D. 1986 *Relevance: communication and cognition*. Oxford, UK: Blackwell.

20. Arbib MA. 2005 From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behav. Brain Sci.* **28**, 105–124. (doi:10.1017/S0140525X05000038)

21. Arbib MA. 2017 Toward the language-ready brain: biological evolution and primate comparisons. *Psychon. Bull. Rev.* **24**, 142–150. (doi:10.3758/s13423-016-1098-2)

22. Scott-Phillips TC. 2017 Pragmatics and the aims of language evolution. *Psychon. Bull. Rev.* **24**, 186–189. (doi:10.3758/s13423-016-1061-2)

23. Fitch WT. 2017 Empirical approaches to the study of language evolution. *Psychon. Bull. Rev.* **24**, 3–33. (doi:10.3758/s13423-017-1236-5)

24. Ackermann H, Hage SR, Ziegler W. 2014 Brain mechanisms of acoustic communication in humans and nonhuman primates: an evolutionary perspective. *Behav. Brain Sci.* **37**, 529–546. (doi:10.1017/S0140525X13003099)

25. Everett DL. 2016 Grammar came later: triality of patterning and the gradual evolution of language. *J. Neurolinguist.* **43**, 133. (doi:10.1016/j.jneuroling.2016.11.001)

26. Seyfarth RM, Cheney DL. 2017 Precursors to language: social cognition and pragmatic inference in primates. *Psychon. Bull. Rev.* **24**, 79–84. (doi:10.3758/s13423-016-1059-9)

27. Mercier H, Sperber D. 2011 Why do humans reason? Arguments for an argumentative theory. *Behav. Brain Sci.* **34**, 57–74. (doi:10.1017/S0140525X10000968)

28. Du Bois JW. 2014 Towards a dialogic syntax. *Cogn. Linguist.* **25**, 359–410. (doi:10.1515/cog-2014-0024)

29. Tomasello M. 2006 Acquiring linguistic constructions. In *Handbook of child psychology* (eds R Siegler, D Kuhn), pp. 1–48. Oxford, UK: Wiley.

30. Clark A. 1998 Magic words: how language augments human computation. In *Language and thought: interdisciplinary themes* (eds P Carruthers, J Boucher), pp. 162–183. Cambridge, UK: Cambridge University Press.

31. Cartmill EA, Beilock S, Goldin-Meadow S. 2012 A word in the hand: action, gesture and mental representation in humans and non-human primates. *Phil. Trans. R. Soc. B* **367**, 129–143. (doi:10.1098/rstb.2011.0162)

32. Byrne RW, Cochet H. 2017 Where have all the (ape) gestures gone? *Psychon. Bull. Rev.* **24**, 68–71. (doi:10.3758/s13423-016-1071-0)

33. Byrne R, Whiten A. 1989 *Machiavellian intelligence: social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford, UK: Oxford University Press.

34. Leavens DA, Hostetter AB, Wesley MJ, Hopkins WD. 2004 Tactical use of unimodal and bimodal communication by chimpanzees, *Pan troglodytes*. *Anim. Behav.* **67**, 467–476. (doi:10.1016/j.anbehav.2003.04.007)

35. Steele J, Ferrari PF, Fogassi L. 2012 From action to language: comparative perspectives on primate tool use, gesture and the evolution of human language. *Phil. Trans. R. Soc. B* **367**, 4–9. (doi:10.1098/rstb.2011.0295).

36. Mayr E. 1961 Cause and effect in biology. *Science* **134**, 1501–1506. (doi:10.1126/science.134.3489.1501)

37. Laland KN, Sterelny K, Odling-Smee J, Hoppitt W, Uller T. 2011 Cause and effect in biology revisited: is Mayr's proximate-ultimate dichotomy still useful? *Science* **334**, 1512–1516. (doi:10.1126/science.1210879)

38. Tinbergen N. 1963 On aims and methods of ethology. *Ethology* **20**, 410–433.

39. Edelman S. 2012 Vision, reanimated and reimagined. *Perception* **41**, 1116–1127. (doi:10.1068/p7274)

40. Marr D. 1982 *Vision: a computational investigation into the human representation and processing of visual information*. San Francisco, CA: WH Freeman.

41. Marr D, Poggio T. 1977 From understanding computation to understanding neural circuitry. *Neurosci. Res. Progr. Bull.* **15**, 470–488.

42. Poggio T. 2012 The levels of understanding framework, revised. *Perception* **41**, 1017–1023. (doi:10.1068/p7299)

43. Edelman S. In press. Perception of object shapes. In *The Oxford handbook of computational perceptual organization* (eds S Gepshtein, L Maloney). New York, NY: Oxford University Press.

44. Lachmann M, Szamado S, Bergstrom CT. 2001 Cost and conflict in animal signals and human language. *Proc. Natl Acad. Sci. USA* **98**, 13 189–13 194. (doi:10.1073/pnas.231216498)

45. Számadó S, Szathmáry E. 2006 Selective scenarios for the emergence of natural language. *Trends Ecol. Evol.* **21**, 555–561. (doi:10.1016/j.tree.2006.06.021)

46. Knight C. 1998 Ritual/speech coevolution: a solution to the problem of deception. In *Approaches to the evolution of language* (eds JR Hurford, M Studdert-Kennedy, C Knight), pp. 68–91. Cambridge, UK: Cambridge University Press.

47. Hurford JR. 1990 Nativist and functional explanations in language acquisition. In *Logical issues in language acquisition* (ed. IM Roca), pp. 85–136. Dordrecht, The Netherlands: Foris Publications.

48. Kolodny O, Edelman S. 2015 The problem of multimodal concurrent serial order in behavior. *Neurosci. Biobehav. Rev.* **56**, 252–265. (doi:10.1016/j.neubiorev.2015.07.009)

49. Solan Z, Horn D, Ruppin E, Edelman S, McClelland JL. 2005 Unsupervised learning of natural languages. *Proc. Natl Acad. Sci. USA* **102**, 11 629–11 634. (doi:10.1073/pnas.0409746102)

50. Edelman S. 2008 *Computing the mind: how the mind really works*. Oxford, UK: Oxford University Press.

51. Kolodny O, Lotem A, Edelman S. 2015 Learning a generative probabilistic grammar of experience: a process-level model of language acquisition. *Cogn. Sci.* **39**, 227–267. (doi:10.1111/cogs.12140)

52. Bullock D. 2004 Adaptive neural models of queuing and timing in fluent action. *Trends Cogn. Sci.* **8**, 426–433. (doi:10.1016/j.tics.2004.07.003)

53. Bickerton D. 2009 *Adam's tongue: how humans made language, how language made humans*. Basingstoke, UK: Macmillan.

54. Laland KN. 2017 The origins of language in teaching. *Psychon. Bull. Rev.* **24**, 225–231. (doi:10.3758/s13423-016-1077-7)

55. Fitch WT. 2004 Kin selection and 'mother tongues': a neglected component in language evolution. In *Evolution of communication systems: a comparative approach* (eds D Oller, U Griebel), pp. 275–296. Cambridge, MA: MIT Press.

56. Fitch WT. 2010 *The evolution of language*. Cambridge, UK: Cambridge University Press.

57. Anderson ML. 2016 Précis of after phrenology: neural reuse and the interactive brain. *Behav. Brain Sci.* **39**, e120. (doi:10.1017/S0140525X15000631)

58. Hagoort P. 2014 Nodes and networks in the neural architecture for language: Broca's region and beyond. *Curr. Opin. Neurobiol.* **28**, 136–141. (doi:10.1016/j.conb.2014.07.013)

59. Friederici AD. 2012 The cortical language circuit: from auditory perception to sentence comprehension. *Trends Cogn. Sci.* **16**, 262–268. (doi:10.1016/j.tics.2012.04.001)

60. Thompson-Schill SL. 2005 Dissecting the language organ: a new look at the role of Broca's area in language processing. In *Twenty-first century psycholinguists: four cornerstones* (ed. A Cutler), pp. 173–189. Hillsdale, NJ: Lawrence Erlbaum Associates.

61. Anderson ML. 2010 Neural reuse: a fundamental organizational principle of the brain. *Behav. Brain Sci.* **33**, 245–266. (doi:10.1017/S0140525X10000853)

62. Koechlin E, Jubault T. 2006 Broca's area and the hierarchical organization of human behavior. *Neuron* **50**, 963–974. (doi:10.1016/j.neuron.2006.05.017)

63. Koziol LF *et al.* 2014 Consensus paper: the cerebellum's role in movement and cognition. *The Cerebellum* **13**, 151–177. (doi:10.1007/s12311-013-0511-x)

64. Gould SJ, Vrba ES. 1982 Exaptation—a missing term in the science of form. *Paleobiology* **8**, 4–15. (doi:10.1017/S0094837300004310)

65. Dehaene S, Cohen L. 2007 Cultural recycling of cortical maps. *Neuron* **56**, 384–398. (doi:10.1016/j.neuron.2007.10.004)

66. Dehaene S et al. 2010 How learning to read changes the cortical networks for vision and language. Science 330, 1359–1364. (doi:10.1126/science.1194140)

67. Stout D, Chaminade T. 2012 Stone tools, language and the brain in human evolution. Phil. Trans. R. Soc. B 367, 75–87. (doi:10.1098/rstb.2011.0099)

68. Fitch WT. 2011 The evolution of syntax: an exaptationist perspective. Front. Evol. Neurosci. 3, 9. (doi:10.3389/fnevo.2011.00009)

69. Botvinick MM. 2008 Hierarchical models of behavior and prefrontal function. Trends Cogn. Sci. 12, 201–208. (doi:10.1016/j.tics.2008.02.009)

70. Jin X, Tecuapetla F, Costa RM. 2014 Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. Nat. Neurosci. 17, 423–430. (doi:10.1038/nn.3632)

71. Lieberman P. 2002 Human language and our reptilian brain: the subcortical bases of speech, syntax, and thought. Boston, MA: Harvard University Press.

72. Longworth CE, Keenan SE, Barker RA, Marslen-Wilson WD, Tyler LK. 2005 The basal ganglia and rule-governed language use: evidence from vascular and degenerative conditions. Brain 128, 584–596. (doi:10.1093/brain/awh387)

73. Ullman MT. 2006 Is Broca's area part of a basal ganglia thalamocortical circuit? Cortex 42, 480–485. (doi:10.1016/S0010-9452(08)70382-4)

74. Sidtis DVL, Pachana N, Cummings JL, Sidtis JJ. 2006 Dysprosodic speech following basal ganglia insult: toward a conceptual framework for the study of the cerebral representation of prosody. Brain Lang. 97, 135–153. (doi:10.1016/j.bandl.2005.09.001)

75. Darkins AW, Fromkin VA, Benson DF. 1988 A characterization of the prosodic loss in Parkinson's disease. Brain Lang. 34, 315–327. (doi:10.1016/0093-934X(88)90142-3)

76. Baldwin JM. 1896 A new factor in evolution. Am. Nat. 30, 441–451. (doi:10.1086/276408)

77. Weber BH, Depew DJ. 2003 Evolution and learning: the Baldwin effect reconsidered. Cambridge, MA: MIT Press.

78. Iriki A, Taoka M. 2012 Triadic (ecological, neural, cognitive) niche construction: a scenario of human brain evolution extrapolating tool use and language from the control of reaching actions. Phil. Trans. R. Soc. B 367, 10–23. (doi:10.1098/rstb.2011.0190)

79. Skeide MA, Kumar U, Mishra RK, Tripathi VN, Guleria A, Singh JP, Eisner F, Huettig F. 2017 Learning to read alters cortico-subcortical cross-talk in the visual system of illiterates. Sci. Adv. 3, e1602612. (doi:10.1126/sciadv.1602612)

80. Sherman SM. 2016 Thalamus plays a central role in ongoing cortical functioning. Nat. Neurosci. 16, 533–541. (doi:10.1038/nn.4269)

81. Barron DS, Eickhoff SB, Clos M, Fox PT. 2015 Human pulvinar functional organization and connectivity. Hum. Brain Mapp. 36, 2417–2431. (doi:10.1002/hbm.22781)

82. Hecht EE, Gutman DA, Khreisheh N, Taylor SV, Kilner J, Faisal AA, Bradley BA, Chaminade T, Stout D. 2015 Acquisition of Paleolithic toolmaking abilities involves structural remodeling to inferior frontoparietal regions. Brain Struct. Funct. 220, 2315–2331. (doi:10.1007/s00429-014-0789-6)

83. Nakahara H, Doya K, Hikosaka O. 2001 Parallel cortico-basal ganglia mechanisms for acquisition and execution of visuomotor sequences—a computational approach. J. Cogn. Neurosci. 13, 626–647. (doi:10.1162/089892901750363208)

84. Báez-Mendoza R, Schultz W. 2013 The role of the striatum in social behavior. Front. Neurosci. 7, 233. (doi:10.3389/fnins.2013.00233)

85. Sutton RS, Barto AG. 1998 Introduction to reinforcement learning, vol. 135. Cambridge, MA: MIT Press.

86. Fareri DS, Delgado MR. 2014 The importance of social rewards and social networks in the human brain. Neuroscientist 20, 387–402. (doi:10.1177/1073858414521869)

87. Csibra G, Gergely G. 2011 Natural pedagogy as evolutionary adaptation. Phil. Trans. R. Soc. B 366, 1149–1157. (doi:10.1098/rstb.2010.0319)

88. Premack D. 1985 'Gavagai!' or the future history of the animal language controversy. Cognition 19, 207–296. (doi:10.1016/0010-0277(85)90036-8)

89. Aiello LC, Dunbar RIM. 1993 Neocortex size, group size, and the evolution of language. Curr. Anthropol. 34, 184–193. (doi:10.1086/204160)

90. Dunbar R. 1998 Grooming, gossip, and the evolution of language. Cambridge, MA: Harvard University Press.

91. Dunbar RIM. 2009 Why only humans have language. Oxford Scholarship Online (doi:10.1093/acprof:oso/9780199545872.001.0001)

92. Hewes GW et al. 1973 Primate communication and the gestural origin of language [and comments and reply]. Curr. Anthropol. 14, 5–24. (doi:10.1086/201401)

93. Jaynes J. 1976 The evolution of language in the late Pleistocene. Ann. NY Acad. Sci. 280, 312–325. (doi:10.1111/j.1749-6632.1976.tb25496.x)

94. Miller G. 2000 The mating mind: how sexual choice shaped the evolution of human nature. New York, NY: Doubleday.

95. Deacon TW. 1998 The symbolic species: the co-evolution of language and the brain. New York, NY: WW Norton & Company.

96. Darwin C. 1871 The descent of Man and selection in relation to sex. London, UK: John Murray.

97. Morgan TJH et al. 2015 Experimental evidence for the co-evolution of hominin tool-making teaching and language. Nat. Commun. 6, 6029. (doi:10.1038/ncomms7029)

98. Greenfield PM. 1991 Language, tools and brain: the ontogeny and phylogeny of hierarchically organized sequential behavior. Behav. Brain Sci. 14, 531–551. (doi:10.1017/S0140525X00071235)

99. Stout D, Chaminade T. 2009 Making tools and making sense: complex, intentional behaviour in human evolution. Cambridge Archaeol. J. 19, 85–96. (doi:10.1017/S0959774309000055)

100. Engels F. 1895 The part played by labour in the transition from ape to man. Stuttgart, Germany: Die Neue Zeit. English translation available at https://www.marxists.org/archive/marx/works/1876/part-played-labour/index.htm.

101. Tallerman M. 2013 Kin selection, pedagogy, and linguistic complexity: whence protolanguage. In The evolutionary emergence of human language (eds R Botha, M Everaert), pp. 77–96. Oxford, UK: Oxford University Press.

102. Vigliocco G, Perniss P, Vinson D. 2014 Language as a multimodal phenomenon: implications for language learning, processing and evolution. Phil. Trans. R. Soc. B 369, 20130292. (doi:10.1098/rstb.2013.0292).

103. Reynolds PC. 1993 The complementation theory of language and tool use. In Tools, language and cognition in human evolution (eds KR Gibson, T Ingold), pp. 407–428. Cambridge, UK: Cambridge University Press.

104. Lotem A, Halpern J, Edelman S, Kolodny O. 2017 The evolution of cognitive mechanisms in response to cultural innovations. Proc. Natl Acad. Sci. USA 114, 7915–7922. (doi:10.1073/pnas.1620742114)

105. Montagu A. 1976 Toolmaking, hunting, and the origin of language. Ann. NY Acad. Sci. 280, 266–274. (doi:10.1111/j.1749-6632.1976.tb25493.x)

106. Isaac GL. 1976 Stages of cultural elaboration in the Pleistocene: possible archaeological indicators of the development of language capabilities. Ann. NY Acad. Sci. 280, 275–288. (doi:10.1111/j.1749-6632.1976.tb25494.x)

107. Klein RG. 2017 Language and human evolution. J. Neurolinguist. 43, 204–221. (doi:10.1016/j.jneuroling.2016.11.004)

108. Tattersall I. 2017 How can we detect when language emerged? Psychon. Bull. Rev. 24, 64–67. (doi:10.3758/s13423-016-1075-9)

109. Fitch WT, de Boer B, Mathur N, Ghazanfar AA. 2016 Monkey vocal tracts are speech-ready. Sci. Adv. 2, e1600723. (doi:10.1126/sciadv.1600723)

110. Stout D, Chaminade T. 2007 The evolutionary neuroscience of tool making. Neuropsychologia 45, 1091–1100. (doi:10.1016/j.neuropsychologia.2006.09.014)

111. Hecht EE, Gutman DA, Bradley BA, Preuss TM, Stout D. 2015 Virtual dissection and comparative connectivity of the superior longitudinal fasciculus in chimpanzees and humans. Neuroimage 108, 124–137. (doi:10.1016/j.neuroimage.2014.12.039)

112. Hecht EE, Gutman DA, Preuss TM, Sanchez MM, Parr LA, Rilling JK. 2013 Process versus product in social learning: comparative diffusion tensor imaging of neural systems for action execution–observation matching in macaques, chimpanzees, and humans. Cereb. Cortex 23, 1014–1024. (doi:10.1093/cercor/bhs097)

113. Stout D. 2011 Stone toolmaking and the evolution of human culture and cognition. Phil. Trans. R. Soc. B 366, 1050–1059. (doi:10.1098/rstb.2010.0369)

114. Stout D, Toth N, Schick K, Chaminade T. 2008 Neural correlates of Early Stone Age toolmaking: technology, language and cognition in human evolution. Phil. Trans. R. Soc. B 363, 1939–1949. (doi:10.1098/rstb.2008.0001)

115. Moerk EL. 1990 Three-term contingency patterns in mother-child verbal interactions during first-language acquisition. *J. Exp. Anal. Behav.* **54**, 293–305. (doi:10.1901/jeab.1990.54-293)

116. Chater N, McCauley SM, Christiansen MH. 2016 Language as skill: intertwining comprehension and production. *J. Mem. Lang.* **89**, 244–254. (doi:10.1016/j.jml.2015.11.004)

117. Bruner J. 1981 The social context of language acquisition. *Lang. Commun.* **1**, 155–178. (doi:10.1016/0271-5309(81)90010-0)

118. Goldstein MH *et al.* 2010 General cognitive principles for learning structure in time and space. *Trends Cogn. Sci.* **14**, 249–258. (doi:10.1016/j.tics.2010.02.004)

119. Nelson K. 2015 A bio-social-cultural approach to early cognitive development: entering the community of minds. In *Emerging trends in the social and behavioral sciences* (eds R Scott, S Kosslyn), pp. 1–14. New York, NY: John Wiley and Sons.

120. Moerk EL. 1996 Input and learning processes in first language acquisition. *Adv. Child Dev. Behav.* **26**, 181–228. (doi:10.1016/S0065-2407(08)60509-1)

121. Silbert LJ, Honey CJ, Simony E, Poeppel D, Hasson U. 2014 Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proc. Natl Acad. Sci. USA* **111**, E4687–E4696. (doi:10.1073/pnas.1323812111)

122. Menyhart O, Kolodny O, Goldstein MH, DeVoogd TJ, Edelman S. 2015 Juvenile zebra finches learn the underlying structural regularities of their fathers' song. *Front. Psychol.* **6**, 571. (doi:10.3389/fpsyg.2015.00571)

123. Onnis L, Waterfall HR, Edelman S. 2008 Learn locally, act globally: learning language from variation set cues. *Cognition* **109**, 423–430. (doi:10.1016/j.cognition.2008.10.004)

124. Huttenlocher J, Vasilyeva M, Cymerman E, Levine S. 2002 Language input and child syntax. *Cogn. Psychol.* **45**, 337–374. (doi:10.1016/S0010-0285(02)00500-5)

125. Steffensen SV. 2016 Cognitive probatonics: towards an ecological psychology of cognitive particulars. *New Ideas Psychol.* **42**, 29–38. (doi:10.1016/j.newideapsych.2015.07.003)

126. Anderson DJ, Perona P. 2014 Toward a science of computational ethology. *Neuron* **84**, 18–31. (doi:10.1016/j.neuron.2014.09.005)

127. Lee KM, Jung Y. 2005 Evolutionary nature of virtual experience. *J. Cult. Evol. Psychol.* **3**, 159–176. (doi:10.1556/JCEP.3.2005.2.4)

128. Reiner M. 2004 The role of haptics in immersive telecommunication environments. *IEEE Trans. Circuits Syst. Video Technol.* **14**, 392–401. (doi:10.1109/TCSVT.2004.823399)

129. Johnson JS, Newport EL. 1989 Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cogn. Psychol.* **21**, 60–99. (doi:10.1016/0010-0285(89)90003-0)

130. Newport EL. 1990 Maturational constraints on language learning. *Cogn. Sci.* **14**, 11–28. (doi:10.1207/s15516709cog1401_2)

131. Dąbrowska E, Street J. 2006 Individual differences in language attainment: comprehension of passive sentences by native and non-native English speakers. *Lang. Sci.* **28**, 604–615. (doi:10.1016/j.langsci.2005.11.014)

132. Jackendoff R, Wittenberg E. 2017 Linear grammar as a possible stepping-stone in the evolution of language. *Psychon. Bull. Rev.* **24**, 219–224. (doi:10.3758/s13423-016-1073-y)

13

rstb.royalsocietypublishing.org *Phil. Trans. R. Soc. B* **373**: 20170052