CrossMark

ORIGINAL ARTICLE

Cumulative Cultural Evolution and the Origins of Language

Kim Sterelny[1]

Abstract In this article, I present a substantive proposal about the timing and nature of the final stage of the evolution of full human language, the transition from so-called "protolanguage" to language, and on the origins of a simple protolanguage with structure and displaced reference; a proposal that depends on the idea that the initial expansion of communicative powers in our lineage involved a much expanded role for gesture and mime. But though it defends a substantive proposal, the article also (perhaps more importantly) defends and illustrates a methodological proposal too. I argue that language is a special case of a more general phenomenon—cumulative cultural evolution—and while we rarely have direct information about communication, we have more direct information about the cumulative cultural evolution of technical skill, ecological strategies, and social complexity. These same factors also enable us to make a reasonable estimate of the intergenerational social learning capacities of these communities (on which rich communication depends) and of the communicative demands these communities face. For example, we can, at least tentatively, identify forms of cooperation that are stable only if third party information is transmitted widely, cheaply, and accurately. So we can use these more direct markers of information accumulation to locate, in broad terms, the period in our evolutionary history during which we became lingual.

Introduction: Aims and Assumptions

Views on language evolution are profoundly constrained by views on its nature, and as a consequence there are two broad traditions of thought and work on the evolution of language. One tradition is framed around Chomsky's conception of language. This view takes the most central, defining characteristic of language to be its computational architecture; a recursive procedure that generates sentences from words, and from structured combinations of words. An important aspect of this generative view of language is that sentences are hierarchical organized structures, not just strings of words. In virtue of this computational competence, languages are unbounded, despite their finite lexicons.[1] As this tradition sees it, the decisive difference between language-enabled minds and language-less minds is computational. This view of the essential nature of language is taken to have the following corollaries: (1) It is universalist. The different languages do not differ in fundamental ways; nor (except for rare, pathological individuals) does individual competence vary in significant ways. Variation between speakers and languages is minor noise, compared to what they have in common. (2) It is individualist: language is an internal cognitive competence of individual agents; it is not essentially social. (3) The

✉ Kim Sterelny
  Kim.Sterelny@anu.edu.au

[1] School of Philosophy, Australian National University, Canberra, Australia

---

[1] Though Hurford (2004b) points out that while recursive syntax is sufficient to make a system unbounded, it is not necessary: an iterative syntax combined with a finite stock of reusable elements can likewise be the basis of an unbounded system.

primary, first-order effects of this cognitive competence are on thought. Language has been co-opted for communication, but its core properties are not explained by its role in facilitating communication. (4) Since the essential feature of language is a computational procedure ("merge") that specifies the structure of sentences, and since this procedure is simple and general, there is no need to build (and perhaps no possibility of building) an incremental model of the evolution of language (see for example: Berwick 2011; Berwick et al. 2013; Bolhuis et al. 2014; Berwick and Chomsky 2016; Chomsky 2016).

A second, somewhat more heterogeneous research tradition is organized around a conception of language as an essentially communicative tool, and so as a public and social phenomenon.[2] A couple of consequences of this guiding thought are: (1) The most striking feature that distinguishes language from other communication systems is its expressive power. That might in part depend on the unbounded character of language central to the first tradition, but a critical part of the explanatory challenge posed by language is to give an account of meaning and its evolution, and perhaps the theory of mind capacities that make meaning possible. (2) Language is a complex system of coadapted elements, involving memory; executive control; theory of mind; capacities to represent the environment in abstract and amodal ways; and fast, accurate, online processing of complex serial inputs. (3) In view of this complex, coadapted character, we need an incremental account of the evolution of language; or, on an alternative version of this broad research tradition, an incremental account of the social intelligence and cognition that, once a threshold is passed, make linguistic meaning possible (Scott-Philips 2015a, b). (I shall return to these two different ways of developing an incrementalist model of the evolution of language at the end of the next section.)

This paper is intended as a contribution to the second of these research traditions, and for that reason, I shall develop two more features of that second framework below, before laying out the specific objectives of this article. I am skeptical of the first tradition, but explaining and defending that skepticism would be a paper in itself, so I am just going to set it aside. That said, a defender of the Chomskian tradition might see the co-option of language for communication as incremental, even though the evolution of language itself is abrupt. Consider, for example, the intense cognitive demands imposed by conversation. Agents produce long and exact sequences of phonemes or gestures, while monitoring and interpreting others' sequences (Christiansen and Chater 2015), and at the same time being sensitive to the social and physical

environment, and to the common knowledge that makes conversation work, even as that knowledge changes as a conversation unfolds. Even if language itself emerged abruptly, it is surely plausible that the scaffolds that turn it into a means of communication evolved gradually. So conceived, some of what follows is relevant to the Chomskian perspective.[3]

Most researchers who take language to be a complex, coadapted system of communication, and hence a system that evolved incrementally, sign on to two further commitments; commitments this article shares. First, the evolution of language depended on social learning and intergenerational transmission. Hominin social lives have long depended on reliable, large bandwidth social learning, for hominin lifeways came to depend on informational capital inherited from the previous generation, and transmitted with reasonable fidelity to the next generation. Thus by 500 kya (and perhaps much earlier), hominin lifeways depended on skills—the control of fire, skilled stonework, natural history understanding—that no individual could learn for him/herself. That was true of language and its various precursors too. These depended on cultural inheritance and cultural evolution. The storehouses of specific signals in language-like systems were built by individual innovation, as new signs were coined and caught on, and were transmitted by social learning to the next generation. Maintaining and extending these storehouses depended on some form of high-fidelity transmission. The same is true of language-specific features of syntax, morphology, and phonology, even if the generative bases of these subsystems of language evolved abruptly, via some large-effect genetic mutation.

It has been suggested that high-fidelity transmission, in turn, depends on a specific form of social learning: imitation learning (strictly defined)–learning how to solve a problem by observing the means others use to solve that problem (Tomasello 1999). I am somewhat skeptical of a general link between imitation and fidelity. Emulation and other forms of socially supported learning can support high fidelity, and one tradition in the cultural evolution literature has shown that redundancy and repetition can compensate for somewhat noisy one-on-one interactions (Richerson and Boyd 2005; Henrich 2016). However, there is a good case for thinking that imitation plays an important role in the transmission of arbitrary signals. Because they are arbitrary, it is hard to reverse engineer their meaning with

---

[2] See, for example: Tomasello (2008), Bickerton (2009), Cloud (2015), Scott-Philips (2015b), Tamariz and Kirby (2015), Henrich (2016).

[3] Berwick and Chomsky do in fact endorse an incrementalist view of the externalization of language, but place that process much more recently than the timings suggested in this paper. They do so through a view of the archaeological record that I suggest (at the beginning of "The Social Scaffolds of Cumulative Culture" section) is outdated, so accepting an earlier dating would be consistent with their main claims about language and its nature.

the help of a few social clues. That said, while the account of language evolution developed in this article depends on the centrality of social learning to hominin life, it is not committed to specific claims about the cognitive foundations of social learning.

There is no circularity in a model of the evolution of language presupposing rich social learning. In Sterelny (2012a, b), I built a detailed account of the incremental construction of learning environments that support the social acquisition of complex skills, even when there are no specific genetic adaptations for the acquisition of those very skills. According to the framework developed there, the early stages of the expansion of social learning depended on models tolerating novices' attention, but not on teaching or on rich forms of communication. That said, when a skill is, and has long been, central to the life prospects of agents over a broad range of the environments they experience, we would expect selection to favor genetic changes that make acquisition more reliable or less costly (Deacon 1997; Avital and Jablonka 2000; West-Eberhard 2003; Zollman and Smed 2010). These might in turn affect the capacity for further cultural elaboration and transmission of the emerging system (Avital and Jablonka 2000). So the evolution of language may well have involved coevolution between cultural learning and genetic response.[4] But even if gene-culture coevolution played an important role in the evolution of language, cultural innovation came before adapting genetic change.[5] There will be selection for genes with specific positive effects on an agent's capacity to learn and use a communication system only when that system is an established and important feature of the local environment.

Second, in company with many of those theorizing about the evolution of language, I accept that an important intermediate stage in language evolution was the establishment of protolanguage (though see Mithen 2005).[6] Our picture of protolanguage comes from pidgins, adult migrant versions of a new language, trading lingua franca, and similar limited human communication systems that arise when people are thrown together over substantial periods and must communicate, but have no common language (Lieberman 1998; Jackendoff 1999; Bickerton 2002, 2009). These pidgin-like systems typically have quite extensive vocabularies,[7] but have little or no grammatical or morphological structure, and their word order is often quite variable. They are face-to-face communication systems, typically somewhat restricted in their expressive power, with mutual understanding depending heavily on context. In the next three sections, I argue that by 500 kya, our ancestors had built a minimal version of protolanguage, but no more. The large, rich, readily expandable lexicon came later; I explain why in "The Social Scaffolds of Cumulative Culture" and "The Changing Communicative Landscape" sections.

The argument that follows depends almost completely on archaeological phenomena and their implications. Some of those working on the evolution of language have given significant weight to evidence from developmental psychology, and from neuroanatomical studies of living humans; of evidence, for example, of neuroanatomical overlap in the control of language and of skilled motor activity (Stout and Chaminade 2012). These connections are certainly suggestive, but human brains seem to be very plastic on both ontogenetic and phylogenetic timescales, and that makes me reluctant to rely evidentially on these connections (Malafouris 2010; Anderson 2014). It is hard to overstate the differences between the developmental environments of ancient and of recent humans. Recent humans, in contrast to ancient humans, develop in largely human-built environments, densely packed with people, with material inscriptions, and with language in use. Even if there are defaults and biases in neural developmental trajectories, these differences are so great that such defaults may well have been very different. Thus I embed my arguments in the material record and its implications.

From Gesture to Protolanguage

In this section, I sketch an account of the evolutionary foundations of the simplest version of protolanguage. In previous work, I have argued in some detail that hominins evolved as cooperative, skilled, tool-using foragers, as a result of positive feedback between ecological cooperation, information sharing, and reproductive cooperation

---

[4] Some theoretical work suggests, however, that cultural evolution is often too rapid for there to be much genetic response. See Chater et al. (2009), Thompson et al. (2016).

[5] Our capacity to speak obviously depends on genetically supported structural features of the mouth, tongue, and larynx, together with very elaborate mechanisms of control (Fitch 2010). But it is not obvious that adaptations for vocal control evolved for language. Music is another possibility (Mithen 2009); so too are much simpler precursors to language.

[6] Stephen Mithen argues that hominin communication until sapiens remained holistic, with a repertoire of calls that encoded information and/or instructions, but where the significance of the call as a whole did not derive from independently meaningful elements from which the calls are built. Animal signaling systems like the vervet warning calls are holistic in this sense. But the call repertoires are very limited.

[7] The referential elements in these pre-language protolanguages are not lexical items as generative grammars represent lexical items. They are not (I presume) tagged with morpho-syntactic features like number, gender, or transitivity. On the protolanguage-first view of the evolution of language, these syntax-guiding features of lexical items came later.

(Sterelny 2007, 2012a, b). As this new lifeway emerged, it selected for enhanced communicative capacities, both in planning and coordination, and in social learning. There is now some evidence that hominins were successful large-game hunters as early as 1.7 mya (Bunn 2007; Bunn and Pickering 2010; Pickering 2013; Bunn and Gurtov 2014). That evidence is compelling for the very large-brained hominins of approximately 500 kya (see, for example, Smith 2012). Hunting large game with short-range weapons[8] with reasonable levels of risk requires both cooperation and coordination, hence communication. So perhaps as early as the erectines (approximately 1.7 mya) and certainly by the Heidelbergensians,[9] the presumptive common ancestor of Neandertals and *Homo sapiens*, hominin technical capacities improved, their lives became much more cooperative, and this built an adaptive platform for improved communication.

It is likely that this extension of hominin communicative capacities included, and depended on, an expanded role for gesture (Tomasello 2008; Corballis 2009, 2011). Great apes, and so presumably early hominins, have top-down control of gesture, and their specific repertoires are shaped by individual and social learning (Genty et al. 2009; Genty and Zuberbuhler 2014; Hobaiter and Byrne 2014). On the framework presented here, this expansion of gestural communication[10] was facilitated by the evolution of technical skills far more elaborate than anything found in great ape lives, skills most obviously manifest in Acheulian technology. For the evolution of technical skills brought long, complex, and precise motor sequences under executive control, and, through selection for social learning, made those sequences salient to others. To the extent that these skills were important and transmitted socially, there was selection to attend to, and parse, these sequences. So, to the extent that expanded communication included an expanding role for gesture and mime, the expansion of technical ability built critical cognitive capacities needed for protolanguage, by bringing complex motor sequences under the control of inner templates rather than external

stimuli, and by selecting for improved memory and executive control. Stone toolmaking selects for focus, and for precise control of motor sequences, for the core must be struck sharply and precisely. Lack of precision is dangerous, for sharp chips of stone can fly off in unpredictable directions, threatening fingers, limbs, and eyes (Hiscock 2014). The changing hominin ecology also selected for an enhanced memory, another ingredient needed to support a larger signal repertoire. As hominins became obligatorily bipedal, their range size expanded, and as they became dependent on their tools, they needed to keep track of a broader range of resources (Jeffares 2014). Larger territories, more detailed maps: greater memory requirements.

Most importantly, the evolution of new technical capacities helps explain the emergence of one of the critical differences between language and animal communication systems. Animal signals are stimulus bound: the famously distinct vervet signals of different predators are responses to threats of predation in the here and now. As a consequence, others can learn their significance through standard mechanisms of associative learning. The stimulus-bound character of the vervet leopard call implies that it is not even roughly equivalent to our word "leopard." For the vervetese call is not used as a meaningful part of more complex utterances, and nor does it refer to leopards in general. In contrast to the vervet signal, most utterances of "leopard" are not produced in confrontation with leopards, and hence word meanings cannot be learned associatively (Deacon 1997; Hurford 2004a). If hominin communication initially expanded through a large role for gesture and mine, it is much easier to explain the emergence of structured signals composed of independently meaningful parts. Mimes and demonstrations are structured by default. Elements of a demonstration are independently significant, and have the potential to be recruited as elements with the same significance in another demonstration.

On this analysis, stimulus-independent signals piggyback on enhanced technical capacities (Sterelny 2012b). Middle Pleistocene hominins mastered complex stoneworking techniques (and possibly fire control and ignition), and these skills selected for top-down control of complex and precise action sequences. As these skills were difficult and expensive to acquire by individual trial-and-error learning, there was also selection on naive subjects to attend to, analyze, and remember the complex action sequences of other agents. Indeed, Peter Hiscock has recently argued that Acheulian skills were actively taught. Acheulian craftwork was both highly skilled and expensive to acquire, given the dangers of undirected trial-and-error learning (Hiscock 2014). These are just the conditions in which we expect teaching to evolve: when it is inexpensive, while reducing otherwise high learning costs of

---

[8] Even using throwing spears, kill ranges may have been as close as eight meters. See Barham (2013, pp. 211–212 and pp. 254–259) for an informative discussion of the effective range of javelins, spear-throwers, and the bow and arrow.

[9] *Homo Heidelbergensis* evolved somewhere between 1 mya and 500 kya, probably somewhat closer to the more recent date (for a recent review, see Manzi 2011).

[10] It is quite possible that vocal communication expanded with gestural communication. In previous work, I have defended a more exclusively gesture-first view of expanding hominin communication (Sterelny 2012b); a view that depended in part on the idea that great ape vocalizations were reflexlike responses to charged situations, and their repertoires seem to be quite insensitive to learning (Tomasello 2008). But that may well be false (Slocombe and Zuberbühler 2007; Crockford et al. 2012).

critical skills, especially in a social environment in which cooperation, enhanced communication, and theory of mind capacities are evolving for other reasons (Thornton and Raihani 2008). If Hiscock is right about teaching in the Middle Pleistocene, these hominins had the ability to take elements of these sequences offline, in demonstration, practice, and perhaps even mental rehearsal (Ron Planer has pointed out to me that there is suggestive evidence that such rehearsal improves performance; Driskell et al. 1994).

Further on this analysis, Middle Pleistocene hominins had the capacity to use inner templates; that is, explicit representations both of the goal of an action sequence, and of the structure of that sequence itself, to initiate and control complex motor sequences.[11] Hominins who can execute complex action sequences from memory, in the absence of their normal physical substrate, have most of the cognitive machinery needed to produce a stimulus-independent mime of that activity: they just need to reframe the social context and point of the action. For they can produce, say, a sequence of hand actions used to ignite fire without actually holding the fire-starting kit they normally use. To turn vacuum practices and demonstrations into a mime, they need a new trigger to initiate the sequence, and a new way of interpreting others' practice-like performances. They need communicative intentions and a theory of mind.

So, if it is to explain displaced reference, inner template control needs to be linked to improved theory of mind. There is reason to suppose that an improved theory of mind was becoming part of the mid-Pleistocene cognitive repertoire, as other aspects of mid-Pleistocene life selected for improved theory-of-mind capacities. The technical skills that depended on inner templates evolved in support of cooperative foraging. As noted above, there is evidence that mid-Pleistocene hominins were effective, cooperative hunters as long ago as 1.7 mya, probably by ambush hunting (Pickering 2013). This form of cooperative foraging required coordination, and hence theory of mind capacities. In face-to-face encounters with large and dangerous animals, each member of the group will need to anticipate what the other will do. They will need to anticipate and respond to others. These agents were equipped (1) with cooperative intentions and expectations; (2) with template-driven control of action sequences; (3) with reasonably advanced theory-of-mind capacities; and

(4) with the capacity to focus on, interpret, and remember an action sequence, as an aid to skill acquisition. These agents had what they needed to interpret a sequence as a message, rather than as practice. Stimulus-independent gestural signals are delivered by inner template control of action sequences and communicative goals, plus enhanced theory-of-mind capacities.

On this view, a minimal protolanguage emerges through linking amplified great ape gestural communication with inner-template-controlled, structured action sequences (evolving through gene-culture coevolution for enhanced technical skills) and with improved theory of mind (evolving under selection for cooperative foraging). I noted in the first section that there is an alternative way of viewing the evolution of language; one in which the incremental changes are changes in social intelligence, not changes in hominin communication systems. In this idea, human language is not a much-modified version of great ape communication. The idea derives from an analysis of meaning and communication first put forward by H. P. Grice, and recently developed by Dan Sperber, Thom Scott Phillips, and Michael Tomasello.[12] The core claim is that genuinely meaningful utterances—the bedrock phenomena of language—are acts committed with overt communicative intentions, and requiring sophisticated theory of mind. On this view, social intelligence and mind reading does indeed evolve incrementally, and when and only when a threshold is reached, acts of meaningful communication become possible. Animal communication systems (including those of great apes) are associative codes, with very limited flexibility, and these cannot gradually morph into language-like systems. In contrast, human communication depends on inferences based on overt intentions to communicate. Speakers both have communicative intentions, and the intention to provide evidence about the existence and content of those intentions. So language-like communication is an evidence-inference interaction, mediated on both sides by advanced theory of mind and by common knowledge. This gives them their great flexibility. With the right stage setting, my pointing to my nose can let you know that I thought last week's talk was appalling. On this view, sophisticated social intelligence evolves without fundamental change in communicative capacity until a threshold is reached. That gives agents the capacity to make meaning in a flexible but ad hoc way. Flexible symbol use comes first; systematic and conventionalized symbol use then follows.

Richard Moore gives a clear depiction of the essential structure of this view of ostensive, overt intentional

---

[11] For a detailed analysis of the task complexity of Acheulian technology, and the complexity and flexibility of control the mastery of that technology demanded, see Stout (2010, 2011). My notion of an explicit inner representation is borrowed from Andy Clark: a representation is explicit to the extent that it can be recruited to help control a variety of activities (Clark 1992), and the idea here is that these inner templates do just that; they control practice and demonstration, not just toolmaking itself.

[12] See Grice (1957), Sperber and Wilson (1986), Scott-Phillips (2015a, b). The analysis that follows relies heavily on Scott-Phillips' recent defenses of this viewpoint.

communication (Moore 2015). A sender S means something by a signal **u** if and only if S sends **u** to R intending:

1. R to produce a particular response r, and
2. R to recognize that S intends (1).

The intended response r determines what **u** means. The fact that R's intention is overt—R wants the target audience to know what he/she is doing (via clause 2)—is what makes **u** meaningful. My pointing to my nose in response to your question about last week's talk is meaningful because I want you to understand that my pointing to my nose is intended to tell you something.

In the standard version of this view, r *is itself* a cognitive response; the audience is intended to represent a complex state of the speaker's mind. One might well suspect that this is an implausibly rich conception of speaking and understanding. But even if we were to accept this richly metarepresentational account of conversational interaction, we can still give an incremental account of the evolution of overtly intentional communication in this rich sense, from intentional communication in a much less rich sense; one within the range of earlier hominins and great apes. For we can give an increasingly rich account of what it is for S's production of **u** to be overt. In the initial stage of the transition to Gricean meaning, the overt production of **u** is just the fact that S's production of **u** is not deceptive. It is public information, and the probability that R will respond to **u** with r would not be reduced, were R to be aware that S produced **u** (and aware that S wanted R to r). In the second stage of the transition, the transaction between S and R is *explicitly cooperative*: S expects R's recognition of S's production of **u** to boost the probability of r. So,

(1)   S produces **u** intending R to r.
(2)   S signals to R his/her production of r.

Moore (2015) argues that great ape gestures are probably overtly communicative in this sense. In the final stage of this transition, S expects r *to depend* on R's recognition of S's intention. What it is for an intention to be overt has transitioned from one in which an agent's goal would not be undermined by the audience's recognition of its presence to one in which it critically depends on that recognition. There is a relatively smooth pathway from intentional, signal-like acts that do not depend on rich metalizing capacities to fully Gricean speaker meaning (this line of argument is developed in much more detail in Sterelny 2017). Moreover, on this view, communication and theory of mind evolve together. On the alternative view, selection drives enhanced theory of mind, despite the fact that the social environment is not posing more complex communication and coordination challenges.

The Limits of Heidelbergensian Conversation

In the last section, I explained why I think Heidelbergensians had a fairly simple gesture- and mime-based protolanguage. In this section, I explain why I think that is all they had. In many ways, Heidelbergensians were impressively humanlike. From the neck down, their physique was humanlike, and they were very large brained, though probably not, on average, quite as large-brained as sapiens or Neandertals. They had impressive technical capacities. They controlled fire (Attwell et al. 2015), and mastered difficult stoneworking techniques. Some late Acheulian handaxes are beautifully made, showing striking control of the material substrate. It is likely that they regularly and successfully hunted large- and medium-size game. Given the physical similarities between us and Heidelbergensians, birth imposed real physical stresses on the mother at the time, and their children were long dependent. Their life history patterns may not have been exactly like ours; their children may not have been dependent as long; they may not have had our life expectancy. But hominin life history had by then evolved towards sapiens patterns, away from the shorter-lived great apes with their less helpless young. Almost certainly, reproduction involved complex webs of cooperation between parents; between the mother and her relatives at and across generations; and within the mother's focal social group (Hrdy 2009; Isler and van Schaik 2012).

In short, there is good reason to believe that their social environment was not just cooperative; cooperation included teamwork, hence coordination, hence communication. Heidelbergensians communicated well enough to support cooperative foraging in challenging environments and with challenging targets. They cooperated to support, nurture, and educate their young. Did they, as Dediu and Levinson (2013) suggest, use language? If so, sapiens and Neandertals inherited language from their common ancestor, and language is a deep feature of human social life. I suspect not. Rather, I shall argue that the social world of archaic sapiens (and probably the later Neandertals) was very different from that of the Heidelbergensians, and that those differences in the social environment (a) explain the differences between Heidelbergensian material and ideological culture, and the cultures of more recent hominins, and (b) imply that Heidelbergensians were unlikely to have a lexically rich protolanguage, or anything approximating full language.

Sophisticated though it was, Heidelbergensian social life and technical achievements were quite different from those of hominins that lived (say) 100 kya. Thus:

(1)   Technology was limited at a location, and there seems to have been limited variation between

locations. Importantly, we do not see any signs of the ability to reliably *retain, fine-tune, and transmit* technical and ecological *innovations*. Thus Heidelbergensians and their immediate successors used a narrower range of tools, and exploited a more limited range of resources.

(2) These earlier hominins show no overt signs of an ideological life. We see no signs of ritual practices in the disposal of the dead; no figurines or other objects made for non-utilitarian purposes.[13] There is no jewelry made from shells, coral, teeth, or ivory; all of this is much later. Ochre is not yet used, so there is no indirect signal that these agents modified the default appearance of their bodies, their shelter, or their gear. These practices all leave traces. If they were standard features of mid-Pleistocene hominin life, it is likely that we would see those traces.

The archaeological signs of a more complex and varied material culture and of an ideological life appeared in the later Pleistocene (the exact dates are controversial), and are taken to indicate the arrival of behaviorally modern hominins. These features of modernity probably appeared incrementally, unevenly, and unstably; there are, for example, microliths from Africa over 200,000 years old, even though this is usually taken as a signature technology of modernity (McBrearty and Brooks 2000; Hiscock and O'Conner 2006; McBrearty 2007). The uneven and fragile arrival of these new techniques and technologies has led most researchers to the view that these signatures of modernity probably did not depend on the evolution of new, genetically canalized cognitive capacities (Roberts 2015). The record does not look as if a threshold was crossed, once and for all (Hiscock and O'Conner 2006; O'Connell and Allen 2007), though this is certainly not universally accepted (see, for example, Klein and Steele 2013). So there is no moment at which hominins became modern; more on this in the next section. Even so, there are very substantial differences between the Heidelbergensians and those hominins that lived in the last 150,000 years.

The Heidelbergensians had to cooperate and coordinate, but that coordination was in a small social world, over a fairly narrow range of potential options, and probably over fairly short time frames. If our picture of their foraging niche is right, they needed communicative capacities significantly richer than those of great apes, but even so, they needed no more than some version of basic protolanguage (perhaps still with substantial gestural elements).[14] To build a significantly richer system, the Heidelbergensian social world would have had to support cumulative cultural evolution; one in which lexical innovations were made and retained, thus allowing the system to become richer over time. There is good reason to doubt that the Heidelbergensians lived in such a world. One of the puzzling features of hominin evolution is the apparently slow pace of technical change until the last 150,000 years or so. Of course, much is invisible; soft materials technologies leave little trace. But in the technology that we can see, that of stoneworking, innovation was very slow (Foley and Lahr 2003). More exactly, the rate at which innovations were made, were taken up in local bands, and then became established as regional practice, was very low.

The record seems to show that for much of hominin history, a small set of core skills—a "core culture"—was reliably retained and transmitted. But innovations rarely established securely enough to become a stable part of local lore, and then part of regional practice. The record of fire, for example, does not become systematic until about 400 kya, though there is clear evidence of control of fire at sites between 800 kya and 700 kya, and more ambiguous dates back to about 1.5 mya. No doubt this patchy record is in part due to trace destruction over time, but it also seems likely that the control of fire (perhaps especially its ignition) was difficult to incorporate within a stable but constrained core culture. The record suggests that there were a number of false starts and partial successes (Gowlett and Wrangham 2013; Twomey 2013; Attwell et al. 2015). This technological record makes it very unlikely that there could have been the cultural evolution of language, or a lexically rich protolanguage, in the Heidelbergensian social world or its predecessors. Such a hypothesis requires that those hominins had a great capacity to retain and transmit communicative innovation, despite their fragile capacity to retain and transmit technical innovation. The rate at which individual agents innovated may have been low compared to later hominins; innovation may depend in part on specialization, or on the very long periods of adolescent learning that may be part of the life history only of our species. But even if (as is possible) individual agents innovated at rates similar to those of later hominins, the social environment made it harder for innovations to establish.

---

[13] Their elegantly symmetrical handaxes may be an exception, if, as has been suggested, they were made to display skill rather than for mundane use (Kohn and Mithen 1999). But even if handaxes were made as sexual display, this would not show an ideological life of norms, prohibitions, myths, rituals.

[14] Once protolanguage was established, there would be selection pressure for it to become more dependent on voice. For one thing, as Liz Irvine has pointed out to me, gesture-based systems are very demanding on visual attention, when three or more are interacting. For another, as Matt Spike points out, voice escapes a line-of-sight constraint; a serious constraint in cluttered environments. I outline an incremental framework for a shift from gesture to voice in Sterelny (2012b).

The evidence seems to suggest, then, that the social world of the Heidelbergensians was not conducive to cumulative cultural evolution. It was not an environment in which cognitive capital was transmitted with the volume, reliability, and precision that regularly allowed innovations on, and expansions of, the core skill set to be retained and to be available as a basis for further innovation. This capacity to retain innovation reliably itself came on stream hesitantly, without a clear point or moment of origin. Moreover, as I shall argue, there are many features of full human language that would not have been of critical value in the social world of the Heidelbergensians, but which are naturally seen as responses to the more complex social and economic environment of the late Pleistocene. These arguments reinforce one another. The reliability with which an item of cultural capital is transmitted to the next generation is sensitive to its centrality and salience in social life. Rarely used skills are much more likely to be lost than those that are part of daily life. The transition to something approximating the expressive richness of contemporary language probably did not begin until the last 200,000 years or so.

The Social Scaffolds of Cumulative Culture

The European archaeological record once seemed to show that there had been an "Upper Palaeolithic Revolution," a dramatic and abrupt transformation in human culture and technology at about the time our ancestors displaced the Neandertals. The traces of our past suddenly showed evidence of music and art, a much wider toolkit, and the use of new materials (ivory, bone). Anatomically modern humans arrived 250 kya; "behaviorally modern humans" only after this revolution, perhaps around 50 kya. This sudden burst of innovation was due, the thought went, to some genetic change that provided a cognitive upgrade, though opinions varied about the character of that upgrade (Klein 2008; Henshilwood and d'Errico 2011; Wynn and Coolidge 2011; Mithen 2013. As I noted in the previous section, there is now close to a consensus that there was no Upper Palaeolithic Revolution; there is no archaeological evidence for a sudden upwards shift in human cognitive power, for the historical record does not show a threshold-like pattern. Signature traits of behavioral modernity appear, then disappear, in the African record long before the presumptive cognitive innovation (typically dated to somewhere between 100 kya and 50 kya). Moreover, they often disappear after the supposed date of that innovation (McBrearty and Brooks 2000; Hiscock and O'Conner 2006; McBrearty 2007). So, while obviously, behaviorally modern culture depends on individual cognitive capacity, the *difference* between Middle Stone Age cultures and behaviorally modern cultures is probably not due to a change in intrinsic individual cognitive capacities.

The alternative to a genetic forcing model is that behavioral modernity is the reliable capacity for cumulative culture, and cumulative culture depends on features of social life (Sterelny 2011, 2012a, b). But which features? The size of the community—both the size of the core foraging band, and the other bands with which there is regular, friendly interaction—really matters. Both size and regular and friendly interaction with neighboring groups support redundancy. If a particular skill is difficult to acquire, it helps to have more models, and more occasions in which a naive subject can see a skill exercised. If a skill is rarely deployed (on a per capita basis), in larger groups, a naive subject will see it deployed more often. Size buffers a group against the loss of cognitive capital through unlucky accident. If there is only one woman in the band with a good knowledge of how to find, recognize, and use medicinal herbs, the group is very vulnerable.

In addition, size also supports specialization (Ofek 2001). A group of ten probably cannot allow a particularly good arrowhead maker to concentrate on arrowhead making; a group of fifty may well be able to do so. Specialization makes it economically possible to expand the range and quality of technology. It cannot pay a forager to invest in making or improving (say) specialist fishing gear, if that gear is used rarely. That is especially so given that foragers are mobile, and hence pay transport as well as production costs for any gear that is too expensive to make, use, and discard. On the plausible assumption that those who develop a special expertise in a practice are the ones most likely to find improvements in it, specialization will also increase the innovation rate. Since specialization reduces redundancy, these factors trade off against one another. Even so, both modeling and some ethnographic examples support the idea that smaller groups find it difficult to retain or expand cognitive capital (Henrich 2004, 2016; Powell et al. 2009; but see Henrich 2006; Read 2006).

Informational capital is, then, vulnerable to demographic attrition. But not all forms of information are equally vulnerable. Vulnerability is increased:

(1) to the extent that information is in few heads rather than many.
(2) if it is difficult to reverse engineer the information from physical products and traces. Transformative technologies like pottery are more vulnerable than more readily reverse engineered techniques like spear-making.
(3) if models do not manifest a skill repeatedly, in daily interaction. There are fewer opportunities to learn about the skill, and fewer occasions in which those with a skill reinforce it through its use.

(4)   if the transmission of skills or information packages requires repeated exposure and/or intensive teaching and/or practice.

(5)   if retaining, not just acquiring, a skill requires regular practice.

Heidelbergensians and their immediate descendants were subject to these demographic constraints. While it is almost impossible to find direct evidence of ancient hominin population sizes, indirect evidence suggests small, scattered populations. There is little or no evidence of technical specialization, or of depleting the local supply of favored resources (we see such indirect evidence of population growth in the last 100,000 years). These demographic constraints would not prevent the stable transmission of a basic protolanguage: of terms for everyday activities, for the objects of daily life, for specific individuals. Such signs would (I conjecture) be in daily use by many members of the local band, thus maintaining capacity. The younger members of that band would have many opportunities to learn through observation and linguistic experiment, as items of basic vocabulary are often used in face-to-face interaction with their target (for example, in using names in greetings), and this aids their acquisition.

However, these constraints would impede the cultural evolution of a richer system. First, they would make it difficult to build and transmit specialist technical vocabularies. Forager herbals, for instance, can be very rich indeed, with thousands of plants identified and named (Berlin 1992). For example, according to one very recent study, the peoples of Nepal use (or have used) about a thousand plant species in their (regionally and ethnically distinct) herbal medicines (Saslis-Lagoudakis et al. 2014). Vocabulary sets of this size and nature are not in daily use. Many plants are encountered only occasionally. In seasonal environments, many are only visible or recognizable at specific times of the year. Names for ubiquitous plants, or those of great resource value, might be in regular use. But that is only a small fraction of these specialist vocabularies.[15] Acquiring such expertise is challenging, probably requiring intensive effort by the less knowledgeable, and explicit teaching by the more knowledgeable. If there were reasonably complete Heidelbergensian herbals, it is unlikely that they were mastered by all in the group. If such

herbals were built by some mix of individual and collective learning, their transmission to the next generation would always be fragile. If my own birding experience is any guide, specialist vocabularies must be practiced to be maintained. I have now lived away from New Zealand for more than five years, and my memory for the names and the field marks of New Zealand's rather modest avifauna has faded badly.

On some conceptions of the emergence of grammar, that process too would be subject to a demographic constraint. Michael Tomasello envisages a process of protolanguage grammaticalization by stages, as lexically expressed information becomes contracted into grammatical particles (Tomasello 2008). For example, information about the time of an action, the number of agents involved, and perhaps their roles (as agent or patient) that is initially expressed with freestanding vocabulary items becomes abbreviated, fixed in a particular place in a term sequence, and becomes attached to, modifying, other vocabulary items. For this grammaticalization machine to work, information about time, number, and role must be needed regularly, and expressed lexically, in Heidelbergensian protolanguage. Tomasello's crank will not turn if information about number, time, and role is inferred from physical context and common knowledge, rather than being explicitly expressed. In these face-to-face microworlds, such information may well have been typically implicit rather than expressed. There seems to be reasonable evidence that if these contextual features are lexically expressed, there are unconscious processing of imitation and mutual adjustment that will result in standardized patterns (Tamariz and Kirby 2015), which are likely then to become abbreviated and attached to other items in the ways that Tomasello has in mind. But in intimate microworlds, common context may well inhibit this initial step.

In brief, the social microworlds of the Heidelbergensians constrained their intergenerational social learning possibilities, and that in turn constrained the richness and complexity of their communicative possibilities.

The Changing Communicative Landscape

It is likely then that demography was important, and the emergence of behaviorally modern humans was in part due to the relaxation of demographic constraints on cumulative cultural evolution. But that was not the only factor. Demographic constraints do not explain the late emergence of material signs of an ideological life. Some symbolic technologies—crafted vulture bone flutes, late Pleistocene cave paintings—depend on very complex technical skills. But many do not. The structured, ritual disposal of the dead probably did not come late to hominin evolution because it

---

[15] Perhaps supporting this analysis, there is evidence that the languages of small-scale societies typically have (much) smaller vocabulary sizes than those of large groups (see Henrich 2016, pp. 239–243). On the other hand, there is evidence that small, closed language communities have languages with the greatest morphological complexity. That may be an effect of such communities having very few adults learning the language as a second language, rather than a direct effect of size (Trudgill 2011), but even so, it is reason to be cautious here.

was too difficult to remember where to take dead bodies. In Sterelny (2014), I argue that the later Pleistocene also saw an economic revolution: a shift from an economy based on face-to-face immediate return mutualism in which the adults of a band foraged together as a team, and divided the spoils on the spot, to an economy based more on direct and indirect reciprocation, in which one agent's contribution might be returned significantly later, in a different form, and perhaps by an indirect beneficiary of the initial prosocial act (see also Tomasello et al. 2012). Cooperation in a reciprocating world can be stable and mutually beneficial, but the cognitive and motivational challenges of managing cooperation are much greater.

I have argued that managing these challenges fuelled the expansion of hominin ideological life, and in turn imposed new demands on forager communicative capacities. A simple protolanguage was no longer enough. One important consequence was the expansion in conversational range. Most simply, the technical toolkit became more diverse, and the resource base was broader. New tools, new targets, new skills, so new terms. Perhaps more fundamentally, in a world of direct and indirect reciprocation, agents need to be able to track and describe their own contributions and those of others, and locate those contributions in time and space. Agents needed to avoid being taken to be free riders, and they had to guard against free riding. Unless ancient hominins were of an implausibly saintly disposition, these would be matters of negotiation and dispute. Chris Boehm claims that historically known foragers are very volatile and voluble about who gets what, though disputes about food rarely escalate into real strife (Boehm 2012). Saints aside, these agents needed the linguistic resources to unambiguously express claims about past contributions and future expectations. Moreover, accurate reputation plays a very important role in stabilizing cooperative practices based on indirect reciprocation, so agents need to be able to specify to third parties the actions of other agents and the contexts of those actions (Binmore 2005).

In Sterelny (2014), I argue for a direct link between reciprocation-based cooperation and a much expanded role for norms in the lives of these agents. Even disregarding the temptations to overvalue one's own contribution, it is difficult to specify a fair return in these more complex situations. How many fish next week is today's duck worth? Perfectly fair-minded agents could disagree. Norms reduce conflict costs by making mutual expectations unambiguous, and by reinforcing prosocial motivations. These economic challenges of managing reciprocation over time might well have been exacerbated by more fraught sexual politics. If the foraging pattern changed so that the adults of a band, or the adult males of a band, no longer foraged as a single group, but split into smaller parties

scattered over substantial territory, and as we will see below, this may well have happened, sexual partners would be less able to directly monitor one another's fidelity.[16] This is a potential amplifier of conflict, and would select for a larger role for norms and for a cultural apparatus that supports them.[17] These agents needed the linguistic tools to express, debate, and teach norms; to negotiate their place in their social network. In short, a shift to an economy of reciprocation made it essential for foragers (1) to master an expanded vocabulary of tools, targets, and skills; (2) to be able to specify the time and value of their contributions and those of others; (3) to report to third parties the actions of others, and the circumstances and effects of those acts, i.e., to gossip; and (4) to express normative claims; to use a normative vocabulary.

This later Pleistocene economic revolution affected the communicative landscape in a second way. Clive Gamble has argued that the later Pleistocene (from perhaps 100 kya) saw a "release from propinquity" (Gamble 1998). The spatial scale of social networks increased, so network links could no longer depend on daily interaction. Gamble had in mind the relations between bands, in ethnolinguistic groups. He interprets the out-of-Africa movements as deliberate migrations, involving planned there-and-back travel, rather than accidental and aimless drifting. As a consequence, Gamble thinks that these humans possessed cultural tools that stabilized cooperative social relations over time and space. Without such stabilized relations, returning parties would have to renegotiate and reestablish their place in their social world. He suggests that elaborate kinship systems were in part solutions to this problem: another aspect of the expanded technical vocabulary of the later Pleistocene.

There was also a release from proximity within the band. Though the dates remain controversial (Sisk and Shea 2011), the later Pleistocene saw a projectile revolution. Hunting with high-velocity weapons (bows, woomera-thrown javelins) selects for smaller hunting parties, as one or a few projectiles can kill. The advantages of quiet movement in ambush and stalking outweigh the larger throw weight of larger parties. Bow-and-arrow hunters typically hunt in groups of two or three (sometimes even alone) (Layton et al. 2012). When such a team size is effective, by splitting into a number of hunting parties, the band will search territory more effectively, and the group as a whole will reduce variation in success. Fracturing the band also results from an expansion in resource breadth, perhaps initially through a sexual division of labor

---

[16] This may help explain why women typically forage in mid-size groups; rarely alone (Layton et al. 2012).

[17] Deacon makes this idea a centerpiece of his whole theory of language evolution in his (1997) book.

(O'Connell 2006). Different resources are found in different places, and they often must be harvested with different skills and equipment. Sometimes a party can hunt game and fish at the same time and place, but often these targets will be incompatible, and it will make sense for different teams to chase different targets. Moreover, once foragers begin to target a broader range of resources, mobility decisions become more fraught, for resources deplete at different rates. Women will often prefer to stay when men would choose to shift base camp. One solution is to adopt a different mobility pattern (known as "logistic" mobility), in which the base camp of the group as a whole moves less often, but work parties targeting specific resources in specific locations travel widely, sometimes staying at work camps for days or weeks (Binford 1980). The band is less often together.

The release from proximity and the shift to reciprocation imposed new demands on communication. These foragers needed both a much richer vocabulary and just about the full illocutionary menu of the modern world. They needed language not just to inform and coordinate, but to argue, barter, gossip; to talk about the possible and the forbidden; the esoteric as well as the mundane. The extent to which Neandertals were experiencing similar social and technical changes remains very controversial. My best guess is that there was some parallel cultural evolution in that lineage too: they used ochre, there is a little Neanderthal jewelry, some funeral practices (Zilhão 2007, 2011; Zilhão et al. 2010).

The expansion of communicative demands just sketched is compatible with late Pleistocene humans communicating with a lexically rich protolanguage. After all, pidgins and trading lingua franca support a diverse array of speech acts. That said, these changes do also select for more regular, systematic, and conventionalized ways of talking. The later Pleistocene economic and technical revolution, and the social changes that accompanied it, led to a more fractured group; and to a spatially and temporally expanded fission–fusion cycle. As a consequence, these changes also increased the "information gradient" in the band. Different agents will typically be exposed to different samples of the ephemeral information about their local world. As Dan Dennett pointed out long ago, in an otherwise cooperative world, steeper informational gradients select for communication and information sharing (Dennett 1983). But these steeper gradients also select for tweaking the communicative format. Interpreting idiosyncratic and enthymematic utterances depends heavily on common knowledge. Jochen says "the window" and I look through a side window to see a yellow-tailed black cockatoo in a tree. My ability to understand his advice depends heavily on our mutually rich understanding of the context and one another: we both heard the distinctive call, Jochen knows the pleasure I take

from seeing these parrots, and so on. Much can be left implicit when these rich common knowledge conditions are satisfied. Protolanguage-like systems depend heavily on such shared and mutually recognized contexts. The later Pleistocene economic revolution eroded this foundation of pragmatic interpretation, this rich mutual knowledge. Not entirely of course; these foragers knew a lot about one another and their world. The release from proximity gave them more to talk about; they had more information to trade. But as a consequence, they are less well poised to use non-linguistic context to guide interpretation. The steeper informational gradient selects for more explicit, conventional, regularized communication. It selects for something like grammaticalization.

Back to Methodology

Let me finish by returning to the methodological theme of this article. The claim developed is not, of course, just that the evolution of language has been shaped by, and is an instance of, cumulative cultural evolution (probably involving gene-culture coevolution). Nor is it the claim that proposals about the timing and shape of the evolution of language should be tested against the material record of hominin evolution. Both of those ideas are common ground to the broad family of views of which this article is an instance. Rather, it is a proposal for, and an example of, the integration of evidential streams from the historical record. Attempts to tie the evolution of language to the paleoanthropological record have standardly looked for a specific behavioral or technical signal of the arrival of language: *a language signature*. For example, the regular use of material symbols is often seen as the signature of language (see, for instance, Tattersall 2016); so too are long-distance trade networks (Marwick 2003). This article does not look for a specific signature of language, or of the various versions of protolanguage. Rather, it integrates information about (1) different foraging economies and the communication and coordination demands those economies impose; (2) the cognitive capacities implied by the manufacture, use, and social transmission of different technological suites; (3) the social and demographic conditions on high volume, high fidelity social transmission; and (4) the complexity of hominin social worlds at different times, as a function of (a) territory size and movement patterns; (b) group size (for which we very rarely have direct evidence); and (c) economic complexity—the division of labor, the organization of collective action, sexual politics, the distribution of resources.

Collectively, these evidential streams enable us to form, in an admittedly fragmentary and fallible way, pictures of the differing social worlds of long-vanished hominins, and

of the ways those worlds require and constrain communicative capacities ancestral to language. I have used this stance to argue that any communication system approximating, or even approaching, the scope of known languages presupposes a demanding form of cultural transmission, even if supported by genetic changes that made transmission more reliable. While it is difficult to get direct evidence about ancient communication systems, we have better evidence about ancient groups' more general capacities to accumulate and transmit information. We can use this to probe the communicative demands on ancient groups, and their capacity to meet these demands by transmitting large, complex, and arbitrary systems to the next generation. I have exploited this methodology to suggest a relatively late, gradual emergence of lexically rich protolanguage (or full language), perhaps in the last 200,000 years. That argument depends on the claim that we see then, but not before then, an expansion in ecological, technical, and social complexity; an expansion that signals more reliable capacities to keep and transmit information, and an expansion indicating a heavier load on communicative skills. New discoveries could easily change those dates, and our views of the complexity of the social lives of ancient hominins. But while that would undermine the timing of language evolution suggested in this article, it would not undermine the methodology of seeing the evolution of language as a special case of a general process, one whose operations we can more directly identify.

References

Anderson M (2014) After phrenology: neural reuse and the interactive brain. MIT Press, Cambridge

Attwell L, Kovarovic K, Kendal JR (2015) Fire in the Plio-Pleistocene: the functions of hominin fire use, and the mechanistic, developmental and evolutionary consequences. J Anthropol Sci 93:1–20

Avital E, Jablonka E (2000) Animal traditions: behavioural inheritance in evolution. Cambridge University Press, Cambridge

Barham L (2013) From hand to handle: the first industrial revolution. Oxford University Press, Oxford

Berlin B (1992) Ethnobiological classification: principles of categorization of plants and animals in traditional societies. Princeton University Press, Princeton

Berwick RC (2011) All you need is merge. In: Di Sciullo A, Boeck C (eds) The biolinguistic enterprise. Oxford University Press, Oxford, pp 461–491

Berwick RC, Chomsky N (2016) Why only us: language and its evolution. MIT Press, Cambridge

Berwick RC, Friederici A, Chomsky N, Bolhuis JJ (2013) Evolution, brain, and the nature of language. Nat Rev Neurosci 17(2):89–98

Bickerton D (2002) From protolanguage to language: the speciation of modern Homo sapiens. In: Crow TJ (ed) The speciation of modern *Homo sapiens*. Oxford University Press, Oxford, pp 193-120

Bickerton D (2009) Adam's tongue: how humans made language, how language made humans. Hill and Wang, New York

Binford L (1980) Willow smoke and dogs' tails: hunter-gatherer settlement systems and archaeological site formation. Am Antiq 45(1):4–20

Binmore K (2005) Natural justice. Oxford University Press, Oxford

Boehm C (2012) Moral origins: the evolution of virtue, altruism and shame. Basic Books, New York

Bolhuis JJ, Tattersall I, Chomsky N, Berwick RC (2014) How could language have evolved? PLoS Biol 12(8):e1001934

Bunn H (2007) Meat made us human. In: Ungar P (ed) Evolution of the human diet: the known, the unknown, and the unknowable. Oxford University Press, Oxford, pp 191–211

Bunn H, Gurtov A (2014) Prey mortality profiles indicate that Early Pleistocene Homo at Olduvai was an ambush predator. Quat Int 322:44–53

Bunn H, Pickering TR (2010) Bovid mortality profiles in paleoecological context falsify hypotheses of endurance running–hunting and passive scavenging by early Pleistocene hominins. Quat Res 74(3):395–404

Chater N, Reali F, Christiansen M (2009) Restrictions on biological adaptation in language evolution. Proc Natl Acad Sci USA 106(4):1015–1020

Chomsky N (2016) Minimal computation and the architecture of language. Chin Semiot Stud 12(1):13–24

Christiansen MH, Chater N (2015) The now-or-never bottleneck: a fundamental constraint on language. Behav Brain Sci 39:e62

Clark A (1992) The presence of a symbol. Connect Sci 4(3–4):193–204

Cloud D (2015) The domestication of language. Columbia University Press, New York

Corballis M (2009) The evolution of language. Ann N Y Acad Sci 1156(March):19–43

Corballis M (2011) The recursive mind: the origins of human language, thought, and civilization. Princeton University Press, Princeton

Crockford C, Wittig RM, Mundry R, Zuberbühler K (2012) Wild chimpanzees inform ignorant group members of danger. Curr Biol 22(2):142–146

Deacon T (1997) The symbolic species: the co-evolution of language and the brain. Norton, New York

Dediu D, Levinson S (2013) On the antiquity of language: the reinterpretation of Neandertal linguistic capacities and its consequences. Front Psychol 4:397

Dennett DC (1983) Intentional systems in cognitive ethology: the "Panglossian paradigm" defended. Behav Brain Sci 6:343–390

Driskell JE, Copper C, Moran A (1994) Does mental practice enhance performance? J Appl Psychol 79(4):481–492

Fitch WT (2010) The evolution of language. Cambridge University Press, Cambridge

Foley R, Lahr MM (2003) On stony ground: lithic technology, human evolution and the emergence of culture. Evol Anthropol 12:109–122

Gamble C (1998) Palaeolithic society and the release from proximity: a network approach to intimate relations. World Archaeology 29(3):426–449

Genty E, Zuberbuhler K (2014) Spatial reference in a bonobo gesture. Curr Biol 24:1601–1604

Genty E, Breuer T, Hobaiter C, Byrne R (2009) Gestural communication of the gorilla (Gorilla gorilla): repertoire, intentionality and possible origins. Anim Cogn 12(3):527–546

Gowlett J, Wrangham R (2013) Earliest fire in Africa: towards the convergence of archaeological evidence and the cooking hypothesis. Azania Archaeol Res Africa 48(1):5–30

Grice HP (1957) Meaning. Philos Rev 66:377–388

Henrich J (2004) Demography and cultural evolution: why adaptive cultural processes produced maladaptive losses in Tasmania. Am Antiq 69(2):197–221

Henrich J (2006) Understanding cultural evolutionary models: a reply to Read's critique. Am Antiq 71(4):771–778

Henrich J (2016) The secret of our success: how culture is driving human evolution, domesticating our species and making us smarter. Princeton University Press, Princeton

Henshilwood C, d'Errico F (eds) (2011) Homo symbolicus: the dawn of language, imagination and spirituality. John Benjamins, Amsterdam

Hiscock P (2014) Learning in lithic landscapes: a reconsideration of the hominid "tool-using" niche. Biol Theory 9:27–41

Hiscock P, O'Conner S (2006) An Australian perspective on modern behaviour and artefact assemblages. Before Farming 2:4

Hobaiter C, Byrne R (2014) The meanings of chimpanzee gestures. Curr Biol 24:1596–1600

Hrdy SB (2009) Mothers and others: the evolutionary origins of mutual understanding. Harvard University Press, Cambridge

Hurford J (2004) Human uniqueness, learned symbols and recursive thought. Eur Rev 12(4):551–565

Isler K, van Schaik C (2012) How our ancestors broke through the gray ceiling: comparative evidence for cooperative breeding in early Homo. Curr Anthropol 53(S6):S453–S465

Jackendoff R (1999) Possible stages in the evolution of the language capacity. Trends Cogn Sci 3(7):272–279

Jeffares B (2014) Back to Australopithecus: utilizing new theories of cognition to understand the Pliocene hominins. Biol Theory 9:4–15

Klein R (2008) Out of Africa and the evolution of human behavior. Evol Anthropol 17:267–281

Klein R, Steele T (2013) Archaeological shellfish size and later human evolution in Africa. Proc Natl Acad Sci USA 110(27):10910–10915

Kohn M, Mithen S (1999) Handaxes: products of sexual selection? Antiquity 73:518–526

Layton R, O'Hara S, Bilsborough A (2012) Antiquity and social function of multilevel social organization among human hunter-gatherers. Int J Primatol 33:1215–1245

Lieberman P (1998) Eve spoke: human language and human evolution. Norton, New York

Malafouris L (2010) Metaplasticity and the human becoming: principles of neuroarchaeology. J Anthropol Sci 88:49–72

Manzi G (2011) Before the emergence of Homo sapiens: overview on the early-to-middle Pleistocene fossil record (with a proposal about Homoheidelbergensis at the subspecific level). Int J Evol Biol. doi:10.4061/2011/582678

Marwick B (2003) Pleistocene exchange networks as evidence for the evolution of language. Camb Archaeol J 13(1):67–81

McBrearty S (2007) Down with the revolution. In: Mellars P, Boyle K, Bar-Yosef O, Stringer C (eds) Rethinking the human revolution: new behavioural and biological perspectives on the origin and dispersal of modern humans. McDonald Institute for Archaeological Research, Cambridge, pp 133–151

McBrearty S, Brooks A (2000) The Revolution that wasn't: a new interpretation of the origin of modern human behavior. J Hum Evol 39(5):453–563

Mithen S (2005) The singing Neanderthals: the origins of music, language, mind and body. Weidenfeld & Nicholson, London

Mithen S (2009) Holistic communication and the coevolution of language and music: resurrecting an old idea. In: Botha R, Knight C (eds) The prehistory of language. Oxford University Press, Oxford, pp 58–76

Mithen S (2013) The cathedral model for the evolution of human cognition. In: Hatfield G, Pittman H (eds) Evolution of mind, brain and culture. University of Pennsylvania Press, Philadelphia, pp 217–234

Moore R (2015) Meaning and ostension in great ape gestural communication. Anim Cogn 19:233–238

O'Connell JF (2006) How did modern humans displace Neanderthals? Insights from hunter-gatherer ethnography and archaeology. In: Conard N (ed) When Neanderthals and modern humans met. Kerns Verlag, Tübingen, pp 43–64

O'Connell JF, Allen J (2007) Pre-LGM Sahul (Pleistocene Australia-New Guinea) and the archaeology of early modern humans. In: Mellars P, Boyle K, Bar-Yosef O, Stringer C (eds) Rethinking the human revolution. McDonald Institute for Archaeological Research, Cambridge, pp 395–410

Ofek H (2001) Second nature: economic origins of human evolution. Cambridge University Press, Cambridge

Pickering TR (2013) Rough and tumble: aggression, hunting, and human evolution. University of California Press, Los Angeles

Powell A, Shennan S, Thomas M (2009) Late Pleistocene demography and the appearance of modern human behavior. Science 324:298–1301

Read D (2006) Tasmanian knowledge and skill: maladaptive imitation or adequate technology. Am Antiq 71(1):164–184

Richerson PJ, Boyd R (2005) Not by genes alone: how culture transformed human evolution. University of Chicago Press, Chicago

Roberts P (2015) "We have never been behaviourally modern": the implications of material engagement theory and metaplasticity for understanding the Late Pleistocene record of human behaviour. Quart Int. doi:10.1016/j.quaint.2015.03.011

Saslis-Lagoudakis H, Hawkins J, Greenhill S et al (2014) The evolution of traditional knowledge: environment shapes medicinal plant use in Nepal. Proc R Soc Lond Ser B 281:20132768

Scott-Phillips T (2015a) Nonhuman primate communication, pragmatics, and the origins of language. Curr Anthropol 56(1):56–80

Scott-Phillips T (2015b) Speaking our minds. Palgrave-Macmillan, London

Sisk M, Shea J (2011) The African origin of complex projectile technology: an analysis using tip cross-sectional area and perimeter. Int J Evol Biol 2011:968012

Slocombe K, Zuberbühler K (2007) Chimpanzees modify recruitment screams as a function of audience composition. Proc Natl Acad Sci USA 104:17228–17233

Smith G (2012) Hominin-carnivore interaction at the Lower Palaeolithic site of Boxgrove, UK. J Taphon 10(3/4):373–394

Sperber D, Wilson D (1986) Relevance, communication and cognition. Blackwell, Oxford

Sterelny K (2007) Social intelligence, human intelligence and niche construction. Proc Natl Acad Sci USA 362(1480):719–730

Sterelny K (2011) From hominins to humans: how sapiens became behaviourally modern. Philos Trans R Soc Lond B 366(1566):809–822

Sterelny K (2012a) The evolved apprentice. MIT Press, Cambridge

Sterelny K (2012b) Language, gesture, skill: the coevolutionary foundations of language. Philos Trans R Soc Ser B 367(1599):2141–2151

Sterelny K (2014) A Paleolithic reciprocation crisis: symbols, signals, and norms. Biol Theory 9:65–77

Sterelny K (2017) Language: from how-possibly to how-probably? In: Joyce R (ed) Routledge Handbook of Evolution and Philosophy. Routledge, London

Stout D (2010) The evolution of cognitive control. Top Cogn Sci 2(4):614–630

Stout D (2011) Stone toolmaking and the evolution of human culture and cognition. Philos Trans R Soc Ser B 366:1050–1059

Stout D, Chaminade T (2012) Stone tools, language and the brain in human evolution. Philos Trans R Soc Lond B 367:75–87

Tamariz M, Kirby S (2015) The cultural evolution of language. Curr Opin Psychol 8:37–43

Tattersall I (2016) Language origins: an evolutionary framework. Topoi. doi:10.1007/s11245-016-9368-1

Thompson B, Kirby S, Smith K (2016) Culture shapes the evolution of cognition. Proc Natl Acad Sci USA 113(16):4530–4535

Thornton A, Raihani NJ (2008) The evolution of teaching. Anim Behav 75(6):1823–1836

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

Tomasello M (2008) Origins of human communication. MIT Press, Cambridge

Tomasello M, Melis A, Tennie C et al (2012) Two key steps in the evolution of human cooperation: the interdependence hypothesis. Curr Anthropol 53(6):673–692

Trudgill P (2011) Sociolinguistic typology: social determinants of linguistic complexity. Oxford University Press, Oxford

Twomey T (2013) The cognitive implications of controlled fire use by early humans. Cambridge Archaeological Journal 23(1):113–128

West-Eberhard MJ (2003) Developmental plasticity and evolution. Oxford University Press, Oxford

Wynn T, Coolidge F (2011) The implications of the working memory model for the evolution of modern cognition. Int J Evol Biol 20:741357

Zilhão J (2007) The emergence of ornaments and art: an archaeological perspective on the origins of "behavioural modernity." J Archaeol Res 15:1–54

Zilhão J (2011) The emergence of language, art and symbolic thinking: a Neandertal test of competing hypotheses. In: Henshilwood C, d'Errico F (eds) Homo symbolicus: the dawn of language, imagination and spirituality. John Benjamins, Amsterdam, pp 111–132

Zilhão J, Angelucci D, Badal-García E et al (2010) Symbolic use of marine shells and mineral pigments by Iberian Neandertals. Proc Natl Acad Sci USA 107:1023–1028

Zollman K, Smed R (2010) Plasticity and language: an example of the Baldwin effect? Philos Stud 147(1):7–21