Wild Voices
Mimicry, Reversal, Metaphor, and the Emergence of Language
by Chris Knight and Jerome Lewis

Why Do Only Humans Talk?

Anthropology is the study of what it means to be human. So it must be at least part of our job to explain why it is that, out of 220 primate species, only humans talk. Speculative theories abound, but little agreement has been reached so far. In our view, a viable hypothesis should invoke well-understood evolutionary mechanisms; respect core aspects of hunter-gatherer ethnography, archaeology, and the fossil record of human evolution; and yield testable predictions.

A good scientific theory should also be conceptually elegant. Here, we explore an entirely new explanation based on two closely linked principles—reversal and metaphor.

A word of warning. The way we have constructed this article is novel, and we ask the reader not to be surprised that we can join a wide range of previously unconnected fields. Our basic idea is simple: using language is so closely bound up with everything else humans do—singing, ritual, kinship, economics, and religion—that no separate, isolable theory of its origins is likely to work.

Our basic assumption is that words and grammar are means of navigating within a shared virtual world. Singing, dancing, and other forms of communal ritual are necessary to join people together in such ideal or imagined worlds. Since language is not a system for navigating within the physical or biological world, it follows that nonhuman primates—creatures whose existence is confined to the realm of brute facts, not institutional ones (Searle 1996)—will have no need for either words or grammar. In an evolving hominin species, we argue, language will not even begin to evolve unless ritual action has already begun to establish intensified levels of community-wide trust in association with a shared virtual domain.

Paradigms Apart

The theoretical paradigms used to study animal communication are incommensurable with those used by linguists to study language. Although speech consists of vocal signals, Darwinian theory faces the difficulty that it does not apply to language. Costs or handicaps (Zahavi and Zahavi 1997), while central to the theory of animal communication, have no place in linguistics. According to Maynard Smith and Harper (2003), the costs to an animal of producing a signal may be divided into two parts—the “efficacy cost” (the investment needed to clearly transmit the signal) and the “strategic cost” (the amount needed to convince an audience of its reliability). The classic example here is the peacock, which must invest massively more time and energy in proving that it can afford that extravagant tail than in clarifying that its signal is a courtship display. To philosophers of language, meanwhile, none of this is relevant, because communicative intentions cost nothing: the listener needs merely to infer what the speaker intends (Grice 1969; Sperber and Wilson 1987).

Not all symbolic signaling is cost free. Collective ritual can be seen as a specific form of costly signaling that underpins the entire human symbolic domain (Durkheim 1976 [1915];
Rappaport 1999). Yet, because animals also perform rituals, we need a robust criterion for distinguishing between symbolic and nonsymbolic displays.

In what follows, we adopt Sperber’s (1975:94) rule of thumb: “That’s symbolic. Why? Because it is false.” From this theoretical standpoint (Sperber and Wilson 1987), symbolic communication rests on the ability of listeners to infer relevant communicative intentions from expressions that, interpreted literally, are inadequate or untrue.

The Pivotal Role of Metaphor

Since the function of language is to communicate thoughts and ideas, we need to understand how speakers succeed in this before asking how the system evolved. Language works through the complementary processes of ostension and inference, ostension being the production of cues to communicative intentions and inference being the interpretation of these cues. Viewed in this light, language takes its place as a particular type of ostensive-inferential communication. What distinguishes language from other such systems is that the cues provided by speakers are vastly more precise (Sperber and Wilson 1987).

Far from being a rare and exotic deviation from the norm, figurative usage underlies all linguistic communication. It was once assumed that interpreting a metaphor involved literal translation, but nowadays, this is a minority view. To translate “John is a pig” into, say, “John is greedy” would be to lose much of the metaphor’s point. Often, there is no literal translation. Abstract concepts such as “time” in fact require metaphorical representation, as when we say “he has a great future in front of him” or “the summer is flying by” (Evans 2004). As figurative expressions become increasingly familiar, conversationalists resort to shortcuts, abbreviations, and conventionalizations in a complex process that, in principle, is entirely sufficient to explain how complex lexical and grammatical structures arise (Smith and Höfler 2014, 2016).

The Language Evolution Conundrum

The fictional status of metaphors poses an evolutionary conundrum. In the absence of very high levels of mutual trust and perceived common ground, we would expect listeners to reject all such fictions as attempts at deceit. Apes do not even attempt metaphor, insisting on hard-to-fake vocalizations that just cannot lie. While their manual gestures may be more flexible, there is nothing metaphorical about these.

Despite their intelligence, apes not only do not talk—they will not even point things out for one another using their hands (Tomasello 2006). Tomasello (2008:5) comments that when a whimpering chimpanzee child is searching for her mother, it is almost certain that all of the other chimpanzees in the immediate area know this. But if some nearby female knows where the mother is, she will not tell the searching child, even though she is perfectly capable of extending her arm in a kind of pointing gesture. She will not tell the child because her communicative motives simply do not include informing others of things helpfully.

Tomasello’s argument about the arm applies equally to an ape’s lips, tongue, soft palate, and mandible, all of which closely resemble the human speech articulators (Duchin 1990). Despite their sophisticated cognition, apes restrict these features to basic functions, such as chewing and breathing (MacNeilage 2008). While emitting a bark or cry, the tongue, for example, plays little or no role. “These expressive limitations,” notes Zuberbühler (2003:299), “seem to be rooted in at least two deficiencies: a lack of sophisticated control over the articulators in the supralaryngeal vocal tract and a remarkable shortcoming in social cognition.”

Although references to ape shortcomings, deficiencies, and lack of control permeate the language-origins literature, we prefer Tomasello’s motivational account. Ape have many more capacities than they are normally inclined to use. Any hominin ancestor must have been able to control its tongue—otherwise, it would have been unable to taste, masticate, or safely swallow food. No ape or monkey has an inflexible tongue. When the animal needs to communicate a thought, however, it leaves the tongue out of it. It is this that needs to be explained.

Signal evolution theory (Maynard Smith and Harper 2003) immediately suggests an explanation. Among the advantages of sound are that—unlike visible gesture—it carries over distances, goes around corners, and works in the dark. But insofar as a sound emanates from an invisible or distant source, the listener is deprived of contextual evidence of its reliability. Keeping vocalizations tied to bodily states may seem inexplicable to linguists, but it is a good way to give nonhuman listeners confidence in what they hear. “Who are you gonna believe, me or your own eyes?” joked Groucho Marx, reminding us that humans often acknowledge the same need. Mistrusting one another’s scheming, Machiavellian minds, primates ignore the all-too-flexible tongue, preferring to rely on the evidence of their own eyes and ears.

So here is the conundrum of language evolution. We need to explain how and why natural selection, in the human case, switched from quarantining the primate tongue—excluding it from all but a marginal communicative role—to developing and fine-tuning that same tongue’s role as the most important speech articulator of all. Since this development was biologically unprecedented, something quite specific and remarkable must have happened. The challenge is to narrow down what it was.

Song First: Vocal Flexibility and Deceptive Signaling

First, consider the ecological niche of Early Pleistocene hominins. It included a formidable community of at least 12 species of saber-toothed cats, eight species of other felines, and nine hyena species (Foley 1984; Marean 1989). What is extraordinary about our ancestors’ successful occupation of this savanna habitat is not just that they avoided being eaten by these predators (Hart and Sussman 2002) but that they went beyond
such dangers to compete directly with these predators "at their own carnivorous game" (Whiten 1999:175).

When our ancestors were vulnerable hominins living in the open with limited weaponry, they may have survived by increasing the range and diversity of their vocal calls. Lions prowling in the dark may have been more wary of approaching a noisy bunch of females and infants if unexpected pitch variations made it difficult to estimate group size and risk. Ethnography from Central African (Lewis 2009) and Indian forest people (Thin 1991:102–103) describes how forest dwellers use rhythmic clapping, drumming, chanting, and choral singing explicitly to keep wild animals away. Marshall Thomas (2006: 271–272) suggests that San trance dancing once served a similar purpose. Our suggestion is that, over time, enhanced vocal range and control—capacities initially adapted to prevent us from becoming prey—eventually allowed us to reverse the situation and become effective predators ourselves.

To this day, Pygmies in the Congo Basin, like many human hunters across the world, imitate the cry of, say, a young monkey who has just fallen from a tree to lure the larger males to come down into range. Similarly, imitating the "good food" grunt of a wild pig draws unsuspecting animals close enough to be speared (Lewis 2009). The animal victims of human vocal deception are honest signalers, hence correspondingly hardwired to expect honesty in return. Since, from the standpoint of any duiker, the frequency of incoming calls from conspecifics far exceeded that of the occasional human fake, resistance to deception evolved much more slowly than human capacities to deceive. Human deceptive signaling was, in this sense, one among several "evolutionary surprise attacks" (Whiten and Erdal 2012) directed by humans against other species, escalating the arms race between predator and prey so rapidly that victims cannot keep up.

It might be wondered why early humans adapted in this way, whereas other savanna–dwelling hominins—perhaps equally threatened by ferocious carnivores—apparently did not. We have no way of reconstructing the vocal repertoire of Pananthropus. Yet, for strategies of complex vocal imitation to become evolutionarily stable, deceptions appropriate to a wide range of recurrent environmental challenges would have to be learned, shared, and intergenerationally transmitted. There is a close coevolutionary relationship between coalitional resistance to being dominated and intergenerational cultural transmission (Whiten and Erdal 2012). The marked sexual dimorphism characteristic of Pananthropus robustus—with the male sporting a sagittal crest and apparently under pressure to grow larger more quickly—suggests conditions of intensified primate-style dominance (Lockwood et al. 2007). In the absence of cooperative childcare in association with increasingly stable counterdominance coalitions, cumulative cultural evolution based on the "ratchet effect" ( Tomasello et al. 1993) would not have come into play.

Tongue gymnastics, being voluntarily controlled, divulge little about emotional states. It is safe to assume that even our most distant primate ancestors could silently and dispassionately manipulate the tongue. But our distinctively human ability to produce pitch variations is a more recent development (Fitch and Zuberbühler 2013:33). If music so powerfully wrenches our emotions, it may be because we still retain a naive costly signaler faith in the honesty of those alterations of pitch representing genuinely changing arousal states. To alternate between high notes and low, we still need to work ourselves up, experiencing real changes in bodily and emotional state.

The Reversal Principle

From explaining how humans began deceptively signaling while respecting costly signaling constraints, we now turn to the social and political conditions necessary for language to emerge. Humans are a hypersocial species, adapted to life in groups consisting of multilayered coalitions. From a signal evolution perspective, this has profound theoretical implications. Dyadic communication differs fundamentally from signaling as part of a coalition. The critical point is that a coalition differentiates insiders from outsiders, each vocalization varying in significance from threatening to reassuring, negative to positive, according to the standpoint from which it is heard.

This brings us to the concept of reversal. A display of resistance against some external threat, while sounding aggressive to outsiders, may be heard as comforting and supportive by members of the signaler’s own group. This contrast is a logical opposition or reversal, but one that, in principle, might have evolutionary implications. Pursuing this thread, we suggest that the first vocally expressed metaphors may have been fake versions of animal cries—literal falsehoods whose significance was reversed by the fact that they were now uttered and interpreted within the signaler’s own group.

Smiling and laughter beautifully illustrate the reversal principle. It is widely believed that the distinctively human smile has its evolutionary origin in the nonhuman primate “fear grin”—a gesture of tense, nervous submission (Van Hooff and Preuschoft 2003). The relaxed human version of this primate facial expression—the good-humored smile—is a fear grin under reversed social conditions. Eibl-Eibesfeldt (1989:138) likewise argued that human laughter can be traced back to the rhythmic cries of group-living primates mobbing a common enemy. In conformity with the reversal principle, we can see laughter as aggressive vocal mobbing, except that, in our case, it culminates as the apparent threat dissolves, allowing the rhythmic chorusing to be enjoyed for its own sake.

Human laughter consists of repeated segments of sound that are all emitted during the same prolonged exhalation, unlike chimpanzee “tickle play” vocalizations (glossed by primatologists as laughter), which strike us as little more than heavy breathing. Although Provine (2000:96–97) sees the evolutionary roots of both types of laughter in primate tickle play, this is not convincing. In humans, extreme laughter may involve tears, revealing its close neurological links with crying (Provine 2000:187). The connection with mobbing is also clear. “Laughter,” as Provine (2000:2) points out, “is a harlequin that shows two faces—one smiling and friendly, the other dark and
ominous... Laughter can serve as a bond to bring people together or as a weapon to humiliate and ostracize its victims.” Chimpanzee laughter lacks these ambivalent in-group/out-group dimensions, indicating that the two types of rhythmic vocalization may have different evolutionary roots.

Eibl-Eibesfeldt’s (1989:138) mobbing hypothesis brings to mind the chimpanzee “waa-bark.” Goodall (1986:130) describes this typically collective, choral vocalization as a kind of “running commentary” provided by bystanders during a conflict between others, usually indicating sympathy for the victim. De Waal (1996:91–92) describes an incident involving “Jimoh,” the alpha male holding sway over a large group of chimpanzees in a spacious enclosure. Having just discovered one of his favorite females secretly copulating with an adolescent male, Jimoh rampaged around the enclosure intent on catching the male culprit. Before he could accomplish his aim, however, several females close to the scene began to waa-bark louder and louder, until literally everyone’s voice was part of a deafening chorus. As the rebellion reached its climax, Jimoh retreated with a nervous grin on his face; he had got the message.

Mixed with aggressive “waas” and various hoots, similar mobbing noises are heard when neighboring chimpanzee patrols encounter one another and when hunting parties fend off threats posed by enraged animals that they are trying to hunt. What is interesting for our purposes is that a coalition of females will, at one moment, be mobbing, say, a python, only to redirect the very same barks to defy an over-aggressive male within their own social group. Not merely communicative, these barks perform a normative function. “A handful of scattered subordinate protests up in trees can be ignored by a superior as he displays,” writes Boehm (1999:169), “but an entire group waaing in a context that suggests imminent physical intervention will get his attention. In this sense, waa-barks provide a signal by which individuals in various roles can read the political dynamics that are taking place in their group. The subordinates, if they sense enough support, may be emboldened to rebel in deed, rather than by voice alone.”

Although this is not yet laughter, we see it as prefiguring the kind of laughter that, among hunter-gatherers especially, cuts over-assertive individuals down to size while bonding the in-group together (see Knight 2000; Lee 1988; Lewis 2009; Woodburn 1982). Laughter is an important force preserving those distinctively human levels of in-group trust and mutuality on which linguistic creativity in turn depends. Humiliation for breaking social norms reverses into solidarity through recognizing them—laughing at becomes laughing with.

Cooperative Eyes

The contrast between hunter-gatherer egalitarianism and primate-style dominance is reflected in the reversed design of the eyes. Those of adult apes appear dark against an equally dark background, making it difficult to detect direction of gaze. Only in humans is the location of the iris visible from a distance, showing dark against a white sclera background (Kobayashi and Kohshima 2001). Reflecting face-to-face interpersonal dynamics quite unlike those of chimpanzees, such eyes appear designed for turn-taking and two-way or inter-subjective mind reading (Tomasello et al. 2007). This can be observed as human mothers with infants watch each other’s eyes while playing with objects, progressively establishing the “we” intentionality essential for language skills (Bruner 1983; Tomasello 2003). If the “cooperative eye hypothesis” is correct, “it would be especially useful to know when in evolution humans’ highly visible eyes originated, as this would suggest a possible date for the origins of uniquely human forms of co-operation and communication” (Tomasello et al. 2007).

Whichever date is proposed, this was no mere modification but marks a systematic reversal of previously prevailing political relationships. Long before the emergence of language, evolution in our species accomplished a cascade of profound reversals, from being prey to becoming predator, turning the fear grin into the smile, aggressive vocal mobbing into relaxed laughter, and eyes for looking out into eyes enabling others to look in.

Encephalization and Life-History Constraints: The “Gray Ceiling”

Beginning some two million years ago, as our relatively slow-moving and defenseless ancestors became exposed in more open environments, it was the fear of predators that began driving these reversals, leading to larger group sizes, increased social complexity, and selection for larger brains. Faced with heavier pregnancy and childcare costs, females began responding with novel foraging, sexual, and alloparenting strategies, modifying male behavior accordingly.

Why did our ape relatives not similarly evolve? Ape mothers cannot give birth to offspring with brain sizes approaching human levels, because they get no help with their babies. In their case, sustaining brains exceeding a “gray ceiling” of around 600–700 cc would involve levels of infant mortality and maternal stress likely to compromise the long-term viability of populations. Cooperative breeding allowed Homo erectus to increase population sizes even when greatly exceeding this ceiling, producing brains twice as large as those of chimpanzees (Isler and van Schaik 2012:S463–S464).

It is precisely when an evolving hominin mother lets others hold her baby that selection pressures for two-way mind reading and triadic structures of joint attention are set up (Burkart et al. 2014; Hrdy 2009). The mother must elicit support and accurately judge others’ intentions toward her offspring; likewise, her baby must interact with its new caregiver while monitoring “where’s mum gone?” The alloparent—necessarily a female relative in the original scenario—adopts a quasi-maternal role. In the course of such activities as mutual gazing, cooing, babbling, and kissfeeding, an array of cognitive skills and dispositions develops to help mother, baby, and alloparent stay in mutually reassuring contact with one another. On this basis, we suggest
that our uniquely cooperative eyes began evolving around two million years ago.

Great ape mothers never needed human-like elaborate mechanisms for bonding with their offspring, because they rarely if ever let their babies go. They cannot fully trust those around them, because on reaching maturity, they must usually leave their natal group, severing their bonds with familiar female kin. By implication, Homo mothers actively resisted such isolation, reversing the primate tradition by remaining in their natal groups, close to kin that they could trust with their babies. In line with the “grandmother” hypothesis (e.g., Hawkes et al. 1998; O’Connell, Hawkes, and Blurton Jones 1999), we take it that a mother’s first alloparenting recourse would have been her own mother.

For a mother to assist with her daughter’s children, she must live close by. While this idea is incompatible with traditional patrilocal band assumptions (e.g., Foley and Gamble 2009; Gavrilets 2012), the growing consensus around cooperative breeding has lent new credence to Dunbar’s (1996:150) suggestion that “female bonding may have been a more powerful force in human evolution than is sometimes supposed.” The importance of female coalitions (cf. Knight 1991, 2008; Opie and Power 2009) is confirmed by population geneticists reporting a deep-time bias to matrilocality among African hunter-gatherers (for Khoisan, see Schlebusch 2010; for Central African Pygmies, see Verdu et al. 2013).

The Egalitarian Revolution

The “basic thing about linguistic symbols is that they are intersubjective . . . meaning that they are comprehended and understood in the context of self-other equivalence” (Tomasello and Rakoczy 2003:128). Since self-other equivalence is unlikely to be fostered under conditions of dominance and subordination, we follow Boehm (1999), Whiten and Erdal (2012), and Gavrilets et al. (2008) in highlighting political egalitarianism as the critical element in the transition to symbolic culture and modern mind.

Against this, Tomasello and colleagues (2012) argue that the critical element must have been warfare:

Creating cultural conventions, norms, and institutions at the level of the social group as a whole requires a new way of thinking in which there is a “we” that constitutes not just my current partners in a collaborative enterprise but all of us in this society. This new way of thinking—that we are a “we”—very likely evolved in response to group competition, as each group had to “circle its wagons.” (Tomasello et al. 2012).

There are two problems with this scenario. One is that, even if it occurred in certain populations, it is difficult to see how it could possibly have led to language. Neighboring groups of male chimpanzees frequently engage in violent conflict without evolving toward language. The cultural ratchet effect is unlikely to be fostered by such strategies, because preserving innovations requires continuity and stability—and in war, no one can expect to win all the time.

The other difficulty with Tomasello’s scenario is that warfare is not what egalitarian, immediate-return hunter-gatherers do. Without leaders, warfare is not possible. Far from being warlike, the ethnographic literature describes people who actively welcome, include, and even marry their neighbors; when conflicts do arise, the first choice is typically to retreat from hostility. The theory of primitive warfare does have its supporters (e.g., Alexander 1987; Pinker 2011; Shackelford and Weekes-Shackelford 2012), but unless the idea is ethnographically supported, we can legitimately dismiss it as an ethnocentric assumption.

Immediate-return hunter-gatherers do not erect fences or defend borders and are systematically disengaged from property; hence, they are disengaged from its potential to create dependency (Woodburn 1982). Demand sharing is imposed on anyone with more than they can immediately consume, preventing saving and accumulation. Everyone can move freely and has direct individual access to the resources necessary for survival. Individuals cannot coerce others to do their will, and people who brag or try to impose themselves on others are mercilessly teased and avoided. Such societies exhibit greater gender, age, and interpersonal egalitarianism than any other human societies and exist in Central Africa (Pygmies, such as BaYaka, Efe, Mbuti, and so on), Tanzania (Hadza), Namibia and Botswana (Khoisan groups), India (Andaman Islanders, Hill Pandaram, Nayaka, and so on), and Southeast Asia (Agta, Batek, Maniq, Penan, and so on). The global distribution of these cultural traits suggests that such social systems are highly stable and successful adaptations whose key elements predate human migrations out of Africa. Any theorizing about early modern humans needs to take into account these core traits (Lewis 2014a). Given our origins in Africa, we focus particularly on the BaYaka as a highly resilient egalitarian group whose adaptation to a hunter-gatherer lifestyle, although contemporary, maintains key elements of immediate-return societies. Lewis has been conducting ethnographic research with BaYaka since 1994.

So what common threat can explain “we” intentionality and group-level morality? Here, then, is our alternative to the primitive warfare model. What most endangered cooperative childcare was the roving male strategy—always a Darwinian option—of impregnating one female after the next without stopping to invest. Strategies of this ancient primate kind would have seemed as threatening to would-be investor males as to mothers of large-brained, heavily dependent offspring. This gives us the prospect of a “coalition of everyone”—a “we”—against the spectre of the alpha male. Where conditions are favorable, such a grand coalition—unlike an all-male war party—might expect to win every contest it undertakes.

In this spirit, Boehm (1999:157) pictures hunter-gatherer egalitarianism as the product of “a large, well-united coalition of subordinates who assertively deny political power to the would-be alphas in their group.” Boehm (1999:193) observes that “col-
lectively creating and maintaining an egalitarian society requires a high degree of political intelligence and a systematic understanding of political dynamics and outcomes. It also requires a political capacity to operate in large coalitions and a cognitive capacity to arrive at a shared plan of action.” The cumulative result represented a sea-change—the “egalitarian revolution” (Gavrilets et al. 2008). Boehm (1999:255) terms it “the egalitarian surprise.”

Gender Dynamics in the Middle to Late Pleistocene

In accounting for this sea change, it is not possible to avoid sex. Tomasello and Boehm acknowledge sex when discussing chimpanzees yet manage to adopt a unisex, gender-blind approach when it comes to humans. Hrdy’s gendered model is a necessary corrective, but her persuasive account of cooperative breeding stops short of the crucial later stages of human evolution, when language and symbolic culture begin to evolve.

Our approach combines the insights of Boehm and Hrdy with those of Opie and Power (2009). The costs associated with renewed encephalization among Middle Pleistocene Homo heidelbergensis (c. 700 ka) prompt child-burdened females to pressure males into becoming increasingly reliable helpers alongside grandmothers and sisters. With males bringing nutritious resources, such as honey and meat, children and their carers can reside longer at a camp before relocating, a development reflected in the establishment of structured hearths from around 350 ka (Shimelmitz et al. 2014).

Males are induced to help through strategies of political and sexual countertopdominance. Carried to its extreme, this becomes reproductive leveling (Bowles 2006), with all males sharing roughly equal prospects of fertile sex and, correspondingly, every female having access to the mating effort of at least roughly equal prospects of fertile sex and, correspondingly, reproductive leveling (Bowles 2006), with all males sharing

reliable display of her commitment to the coalition. Males willing to commit should now be as hostile to philanderers as everyone else.

If males invest preferentially in females who resist noninvestors—that is, in females who have accepted initiation into a gendered ritual coalition—we would expect an explosive increase in cosmetics in the archaeological record in conjunction with modern human speciation (Power 2009). This theoretical expectation is empirically confirmed. Traces of predominantly blood-red pigments suggestive of group rituals can be discerned from around 500–300 ka, associated with H. heidelbergensis (Watts et al. 2016). This sparse pigment record precedes the final phase of encephalization, which, in Africa, sees the speciation of Homo sapiens. This is exactly when pigment use—with a marked preference for brilliant blood reds—becomes ubiquitous, being evident in virtually all southern African rockshelter occupations from ∼170 ka onward (Watts 2002, 2014; Watts et al. 2016).

For cooperative hunting to succeed, men’s capacities for violence must be directed outward, their weapons harmonized and used only against nonhumans. During the nights before an elephant hunt, BaYaka of both sexes engage in all-night singing and dancing, a communal activity that bonds everyone but also makes it difficult for couples to get intimate without being noticed. The underlying logic is that conceptualized by Knight (1991) in terms of “sex strike” action. A BaYaka informant remembers her father as a tuma—an elephant hunter: “When he hunted, he always brought back meat. He hunted the gazelle and the elephant. Before an elephant hunt, all the hunters gathered in a camp and danced all night. They take a big pot full of powerful medicine and put all their spears into this pot. Every night they dance and sing all night long. When they sleep, they sleep apart from their wives—no sex before the hunt” (Hewlett 2013:74).

In common with other African hunter-gatherers, such as the San (Marshall Thomas 2006:175) and Hadza (Marlowe 2003, 2005; Woodburn 1964), BaYaka women do not accompany men on dangerous hunting expeditions. Without segregation, the abstention required by a sex strike would be unenforceable. It is not male gender dominance but the dynamics of strike action that explain why hunter-gatherers so often divide gender roles categorically, insisting, for example, that no woman should shed blood in the hunt (Tesfaye 1986). These dynamics of female solidarity also underlie brideservice, in which a man leaves his natal group to work for a bride who continues to reside with her mother or other protective kin (for Hadza, see Wood and Marlowe 2011 and Woodburn 1972; for Pygmies, see Lewis 2002: 74, 127; for San, see Lee 1979:240–242 and Marshall 1959:352, 1976:169).

On Costly Signaling and Metaphor

With the core stable features of African hunter-gatherer ethnography in mind, we can now envisage the collaborative coinage of the world’s first word. “The first word only became a
word,” notes MacNeillage (2008:44), “when a receiver and a sender—a sociocultural dyad—came to treat a particular sound complex as standing for a particular concept. From that momentous occasion onward, every word of every language came into being in this way.”

In our model, however, this “momentous occasion” was not the matching of a concept with a sound but the more profound revolutionary transition just presented—the use of symbolic ritual to establish a way of life based on moral norms. In addition to setting up the conditions for verbal metaphor to be deployed, this ritual was itself a metaphorical performance of a particularly potent kind.

Metaphors need not be verbally expressed—even without words or grammar, a simple gesture might do the job (McNeill 1992). “Metaphor is primarily a matter of thought and action, and only derivatively a matter of language,” explain Lakoff and Johnson (1980:153). As with symbolism in general, the most basic social condition for metaphor is an audience prepared to tolerate apparent falsehood as a prompt to search for meaning. Shared understandings—abstract concepts—could never emerge in a world where everyone insisted that signals be literally true.

“Generally,” Davidson (1979:40) explains, “it is only when a sentence is taken to be false that we accept it as a metaphor and start to hunt out the hidden implication.” However, that implication—the point of the metaphor—may be lost on a listener who does not share the speaker’s relevant assumptions and experience of the world. A metaphor that seems self-evident to the Central African BaYaka—“Woman’s biggest husband is the Moon,” for example (Lewis 2008)—may strike Western scientists as incomprehensible and unworthy of serious study. The same applies to metaphors equating female menstruation with male bloodshed in the hunt (Knight 1991; Lewis 2008; Testart 1986). Although scientific conceptualization is itself fundamentally metaphorical (Lakoff and Núñez 2000), the core metaphors of hunter-gatherers appear so alien to Western assumptions that scientists from Western, educated, industrial, rich, and democratic (WEIRD) cultural backgrounds (Henrich et al. 2010) rarely feel able to grasp the logic or find common ground. The point here is that we should not expect indigenous categories to match the way Western science carves up the world (Lakoff 1987). If we find it difficult to comprehend the elaborate myths and cosmologies of other people, it is because their common ground is no longer ours.

Our Metaphors and Theirs

In the West, the moon and menstruation have largely faded as experiential sources of the metaphors we live by. For African hunter-gatherers, however, they remain central.

The language of hunting and romance across sub-Saharan Africa is saturated with references to the moon. Romantic liaisons are associated with the full moon, and menstrual seclusion is associated with the new moon (Watts 2005; cf. Knight 1987). When moonlight is absent and the night sky is dark, BaYaka women stay close together, keeping predators away from camp by singing as loudly as they can. Hadza women do the same, singing through the night during their major ritual performance—the epeme dance—timed to occur monthly on moonless nights (Woodburn 1982). One of the universals unearthed by Lévi-Strauss in his Mythologies is an association between darkness, the absence of cooking fire, and the production of loud noises. Since earliest times, ritual and narrative metaphors have elaborated prolifically on conceptual transitions between lunar darkness and light, danger and safety, death and resurrection (Eliaede 2005 [1949]; Lévi-Strauss 1970–1981; Marshack 1964).

Hunter-gatherers are well aware that menstrual bleeding is related to fertility. On a Darwinian “selfish gene” basis, we might expect would-be alpha males to respond by competing for the imminently fertile female. Yet this hardly occurs. Among African hunter-gatherers, the woman’s protective kin attribute supernatural potency to menstrual blood in ways that galvanize both sexes into gendered cooperation, successful hunting, childcare, sharing, conservation, and economic abundance (Knight 1991; Power 2001; Power and Aiello 1997; cf. Testart 1986). These effects—conceptualized by the BaYaka as so many manifestations of ekila (Lewis 2008)—are achieved by constructing women’s blood as magically dangerous to men’s blood-spilling activities, profoundly shaping belief, labor, and conduct down to many details of everyday life.

Given the facts of sexual difference, hunter-gatherer women might conceivably climb for honey and hunt big game animals, but in real life, they use their attractions and solidarity to get men to do this for them. Women’s success in this depends on their sexual autonomy—their freedom to link legitimate sex with male hunting success and the proper sharing out of meat. Wherever women are well-organized, sex becomes a negotiating resource for influencing and managing what males do. Among the techniques are songs, dances, and rituals in which women identify both with the moon and with the game animals they expect their menfolk to hunt (Knight, Power, and Watts 1995; Lewis 2002, 2008, 2014b).

The metaphors of BaYaka and other African hunter-gatherers are logical in the light of these peoples’ shared experiences of nature and its challenges.

The moon. On dark, moonless nights in exposed areas, lions exploit their excellent night vision to surprise their prey. Like other vulnerable creatures, humans are much more likely to be killed and eaten when the moon is not visible in the sky (Packer et al. 2011).

The menstrual cycle. At ~29.1 days, the human menstrual cycle quite closely approximates the moon’s 29.5-day synodic cycle, unlike the menstrual cycle of chimpanzees (~36 days) or bonobos (~40 days; Martin 1992; Saltzman, Tardif, and Rutherford 2010). A possible adaptive basis for the human cycle length could be reproductive leveling: among primates, synchronizing to any natural clock makes it difficult for alphas to monopolize fertile sex with multiple females (Carnes, Nunn, and Lewis 2011; Ostner, Nunn, and Schülke 2008; Power, Sommer, and Watts 2013; Turke 1988). An additional deep-
time evolutionary pressure may have been lions’ habit of eating people on moonless nights (Packer et al. 2011). When early Pleistocene hominins were attempting to survive by robbing big cats of their kills (Blumenschine 1986; O’Connell et al. 2002; Shipman 2011), it may have been logical to restrict overnight journeys—including sexual liaisons—to times when there was a bright moon in the sky. 

_Divergent sexual characteristics and reproductive strategies._ The female menstruates, gets pregnant, and nurses babies. Following impregnation, the male is, in principle, free to disappear.

Lunar periodicity, menstruation, and sexual difference are ever-present natural universals, to which African hunter-gatherers have responded with corresponding symbolic universals—“metaphors they live by.” The core metaphors that have seemed natural to successive generations of hunter-gatherers may be thought of, following Maynard Smith (1982), as optimal responses—evolutionarily stable cognitive strategies—shaped by challenges that are likely to vary only within a limited range.

From Singing to Vocal Symbolism

BaYaka living in the forests of northern Congo-Brazzaville regularly experience predation and attack by large dangerous animals, particularly leopards, elephants, buffalo, and gorilla. Yet they also hunt these and other animals. Evolving hominins must also have been predator at one moment and prey the next. Anthropologists have focused productively on this archetypal example of the reversal principle, aware of how universally it shapes ritual and cosmology (Bloch 1992; Descola 1993; Viveiros de Castro 1998).

For BaYaka, the forest is conscious and will offer abundant resources when pleased by the sounds of human laughter and song. The unhappy sounds of shouting, fighting, or children crying provoke it to withhold its bounty. The sounds most likely to enchant the forest are those of the forest itself, mimicked by animal mimicry on the other hand—choral singing on the one hand, complex vocal polyphony, their joy rises as their intermingled bodies and voices become one. Pinker (1997:534) views music as “auditory cheesecake” that confers no survival advantage. Most ethnomusicologists argue instead that it exists to promote group coordination and cohesion (e.g., Brown 2000; Lomax 1968; Merriam 1964). BaYaka women’s exquisitely synchronized choral singing deters predators by broadcasting to the forest that they are a large and well-organized group, much as synchronized roaring by coalitions of lionesses warns rivals of their numerical strength (McComb, Packer, and Pusey 1994). San all-night trance dancing is argued by Marshall Thomas (2006:271–272) to serve a similar function. This suggests that BaYaka women’s own interpretation of their singing—their insistence that they are singing for their lives—is probably right.

The descended larynx in the adult human male probably evolved initially not to enable language or song but as a size-exaggerating device adaptive in sexually competitive male-on-male roaring contests (Fitch 2002). The theory that music may have evolved initially thanks to sexual selection of this kind—with females choosing males for their vocal skills (Miller 2000)—is unlikely. Among Pygmies, San and Hadza women take the lead in singing, with men playing a secondary role. The ethnography fits the “women and children versus predator” idea better than the “sexual selection for male vocalizers” one. While women and young children have every reason to scare away predators by making noise, men’s priority while hunting is to give no audible sign of their presence, treading carefully and masking their signals by blending them with the forest’s own sounds.

It may seem that, to describe the earliest musical and linguistic sounds as essentially vocal “fakes,” is confusingly negative. But relative to primate vocalizations, that is what they are. Peirce (1940) distinguished between symbols (arbitrarily linked with their object), icons (similar to their object), and indices (physically connected with their object). Only indices are not fakes. By primate standards, both symbol and icon fall into the “fake” category, since they lack any built-in component of reliability.

But symbolism is more than just fakery. To qualify as a symbol, a fake or replica must meet two further conditions: (1) instead of being confused with the original, it must be acknowledged as distinct; and (2) instead of being rejected on those grounds, it must be accepted and socially circulated. Our point is that these conditions are met when vocal fakes originally aimed at deceiving animals—choral singing on the one hand, animal mimicry on the other—are reversed through being redeployed to share ideas within the group.

To sum up our evolutionary hypothesis, the vocal deceptions in which each sex specialized were originally directed outward, against animals unable to resist fake versions of their own species-specific calls. With defiant communal singing con-
structuring “we” intentionality and joint commitment, strategies of sexual and political counter dominance culminated eventually in a new normative order based on reverse dominance and egalitarianism. Internal trust intensified until something without precedent occurred. A hunter entering camp with a dead pig on his back made a “good food” pig vocalization even though no living pigs were nearby. The risk of confusion and misunderstanding might have seemed immense, but with good will, they were overcome. The key was that people began to place trust in one another’s communicative intentions, regardless of the literal truth or reliability of the signals themselves.

Reversing Sexual Violence into Play

Bateson (1973) first highlighted reversal as key to understanding animal play. He observed that, when young monkeys playfully “bite” and “chase” one another, their antics constitute violent aggression—but with all meanings reversed. A “nip” is an aggressive bite reversed by the message “this is play.” The animal’s preliminary play invitation—making a play face, for example—means, in effect, “the aggressive actions that follow are not to be mistaken for real.” Each partner can now afford to “lose” for the sake of the game. Turn-taking implies self-other equivalence and is as central to the logic of animal play as it is to reversing roles from speaker to listener in the dynamics of human gossip and conversation (Sacks et al. 1974; Stivers et al. 2009). Finally, when young animals play, they are at their most unpredictable and creative. While the vocal signals of young primates tend to be stimulus bound, inflexible, nonsymbolic, and liminally controlled, their playful bodily antics are strikingly imaginative, unpredictable, incipiently symbolic, and cognitively controlled—all suggesting a point of departure for the evolutionary emergence of language (Knight 1999, 2000; Lewis 2009).

Once monkeys and apes reach sexual maturity, play fighting turns irreversibly into serious conflict. Former primate playmates—two brothers, for example—become hostile contestants as sexual rivalries come to dominate their relationship. In a fight over sex, playfulness is ruled out, because neither side can afford to lose. This obliges us to ask how so deeply rooted a problem as primate sexual violence could have been contained and transcended during the course of human evolution.

A BaYaka creation myth illustrates how play can transcend sexual violence, reversing male-female conflict into laughter and mutual respect. In the beginning, say the BaYaka, men and women lived apart. Women produced babies without needing men. They did so by dancing with a forest spirit named Ejengi, a phallus-like being from whose raffia clothing little babies would fall. Ignorant of this, men copulated with mapombe—a large, hard fruit filled with white cream:

One day (according to the men’s version), men discovered for the first time that women existed in a distant part of the forest. When first spied upon, the women were dancing with Ejengi. Wanting these creatures for themselves, the men decided to ambush them as if they were wild pigs. Instead of using spears, however, the men beat them with honeycombs. On tasting the sweetness, the women agreed to yield to the men’s desire, whereupon everyone enjoyed sex for the first time. Delighted, one man shouted “I want three of these,” and another, “I want four!” But the elder woman, Beponga, refused. “Only one woman for each man!”

The climax of this myth is a play fight. The men fight the women by beating them with honeycombs. The women are happy to lose this particular game, but they insist on winning others, which the men are happy to lose in turn. Ritualized play pervades the very arena that, in other primates—chimpanzees, for example—leads recurrently to sexual violence. Among these and other African hunter-gatherers (Finnegan 2014), sex no longer shuts down play. Instead, play—now scaled up as adult playful ritual—succeeds in transforming and pervading the entire arena of sex (e.g., Keeney and Keeney 2013 on San; Power 2015 on Hadza).

The BaYaka creation myth ends with men seizing women’s baby-making dance, Ejengi, for themselves. To this day, men dance Ejengi to assert their muscular prowess and pride. Women’s answer is Ngoku—a defiant dance focused around women’s own sexual and reproductive secrets. Conceptualized as their communal spirit, Ngoku acts out the mythic theme of the time when women did not need men.

As the performance begins, little boys are frightened away, running to their fathers to hide. Adult men in huts close to the dance ground typically retreat, some trying to ignore the raucous proceedings by escaping into the forest. As men vacate the central communal space, the women seize control, subordinating the entire community to their authority.

The female community link arms, charging up and down the length of the camp singing “Ngoku! Ngoku!” Older women lead the songs, which consist largely of insults such as “Doto ba die be!” [Old men are no good!], or “Eneke mu ganye, mapindi ba mu pola!” [The vagina always wins, the testicles are empty!]. In one dance, the women lie on their backs rubbing their thighs together until they become frenzied and are lifted up from behind by an elder Ngoku initiate. In another, older female initiates undergo gender reversal to much laughter as they vividly mimic men attempting sex with the younger women. Ngoku reminds everyone that women have solidarity, that access to their bodies depends on consent, and that relations with men will be on women’s terms.

The BaYaka word massana encompasses “ritual” and “play.” Adopting a hunter-gatherer perspective on such things reveals how some of our intellectual difficulties are products of the way Western science carves up the world. From a BaYaka standpoint, ritual is not separated by a chasm from play, which in turn overlaps at many points with language (Lewis 2009, 2014a). A similar perspective is adopted on scientific grounds by Whitehead (2014) and also by Wyman (2014).

In seeking to explain why language evolved in our species but in no other, we have described how evolving human mothers
with increasingly large-brained babies faced progressively heavier childcare burdens, prompting them to resort to cooperative childcare (Hrdy 2009) and to sing to keep safe on dark nights. Once they had coalesced to share childcare and song, the potential existed for them to develop their solidarity for a new purpose—squaring up to threats posed by dominant males. Female-led ritual-cosmetic resistance to dominance by alpha males culminated eventually in “reverse dominance” (Boehm 1999; Knight 2014; Power 2014), outlawing violence or physical threat as a viable reproductive strategy for males. This liberated human creative potential in many ways. Up until this point, play had remained largely restricted to immaturity, because the transition to adulthood invariably caused sexual conflict to break out. Once sexual violence had been marginalized, imaginative play (as massana illustrates) was free to extend without a break into adult life, increasingly embracing it and structuring it, to the point where it becomes “a foundation for hunter-gatherer social existence” (Gray 2009).

The World’s First Metaphor?

Our model is testable in the light of evidence from a number of domains: from archaeology, a cosmetics industry focused on red pigments coinciding with evidence for male investment in cooperatively raised offspring and corresponding rapid encephalization; from genetics, a deep-time bias toward matrilocality among African hunter-gatherers; and from ethnohistory and rock art, “wrong plus red” as the core signature of ritual potency.

The first prediction, concerning cosmetics, arises as a result of the FCC strategy in response to female reproductive costs (see above). The record of red pigments in the African Middle Stone Age currently supports the hypothesis of a correlation with encephalization. A matrilocal bias is predicted, because only on that basis can females maintain kin-based cooperative childcare. As noted above, the genetic evidence reveals a distinctive signature of Central and Southern African hunter-gatherers confirming a deep-time matrilocal tendency.

The third prediction is derived by asking what reverse dominance should look like as a ritual display, given that its principle is sexual defiance. Stereotypically female roles must be systematically reversed. We might expect, in other words, roughly what we find among the BaYaka—women’s raucous singing, dancing, laughter, and rough sexual humor. Among other things, we would expect performers to proclaim a patent falsehood, playfully insisting that they are not what, in biological fact, they are. Since any alpha male would be seeking to impregnate someone of his own (human) species, of the opposite (female) sex, and currently in her fertile (ovulatory) period, it follows logically that female collective defiance should convey the opposite on each count, yielding “wrong species,” “wrong sex,” and “wrong time” (Knight, Power, and Watts 1995).

The universality and salience of these themes is striking to anyone familiar with the initiation rites of African hunters and gatherers, particularly those rites celebrating a young woman’s first menstruation (Power 2009; Power and Watts 1997, 1999; Watts 2005). A Ju’hoansi menarcheal girl is playfully constructed as an eland bull; the metaphor for her menstruation is that she has “shot her eland”; her !Xo counterpart has her face painted with gemsbok designs and ritually shoots a gemsbok mask (Lewis-Williams 1981; Power and Watts 1997). The matriarchal heroine of Hadza mythology is Mambedako, “the woman with the zebra’s penis” who originally owned epene meat (Woodburn 1964). Recalling this myth, the metaphor for any woman who has begun to menstruate is that, like Mambedako, she has “shot her zebra”—akakwa dongo (E. Mouriki, personal communication). Such metaphors make no logical sense until we grasp why, to express “wrong plus red,” women performatively become a gender-reversed, bleeding game animal.

Returning to the BaYaka creation myth, the battle of the sexes culminates in play. It could have ended in violence, but it does not. When a girl first menstruates, that is the moment of greatest risk. The danger is that, by signaling her fertility to the world, her body will incite rival males to fight over her. Here, as elsewhere, her relatives therefore choose this as the moment to act. Among the Ju’hoan Bushmen, details vary, but “what remains consistent in the girls’ puberty rite is that the first sign of bleeding signals a time for immediate community action.” The ensuing dancing counteracts violence by “soothing any existing tensions and giving new hope for the future” (Keeney and Keeney 2013:74). Among the Nharo of Botswana, the dancers—including older women and men—hilariously stage “mock male fighting” to make their point (Guenthner 1999: 166). As women’s humorous performance averts the danger of real fighting, typically to the accompaniment of peals of laughter, the founding principles of hunter-gatherer morality, kinship, ritual, and economics are restated and renewed.

From Mimesis to Grammar

So far, we have discussed how community-wide ritual established the normative conditions for language to evolve. It was Durkheim (1976 [1915]) who first realized that, for representations to be linguistically communicable, they had to form part of a collectively shared conceptual repertoire, ritual alone being capable of generating the necessary concepts in everyone’s head. In his major work, Rappaport (1999) points out that the function of ritual is not to differentiate between lexical meanings but to establish, for everyone, an overarching meaning—a metaperformative or Word—from whose subsequent fragmentation a limitless multiplicity of subsidiary meanings can be derived. If singing, drumming, and dancing last all through the night, it is not because so many different meanings need to be conveyed. The sounds do not need to mean anything and, like Pygmy vocal polyphonies, certainly do not need to include words. Yet, by showing willingness to expend so much time and energy, everyone is demonstrating commitment, leading to a powerful sense of belonging and corresponding mutual trust. For Rappaport, the evolutionary
emergence of such bonding mechanisms must have played a key role in encouraging public confidence in otherwise unreliable words. It is communal ritual that gives rise to what Searle (1996) terms “we-intentionality,” in turn the source of linguistic conventions and other institutional facts. The outcome of all this is a paradoxical insight. Rappaport explains how apparently irrational nonsense—perhaps the relentless repetition of just a few meaningless sounds—may “provide the ground, deeper than logic and beyond logic’s reach” upon which to establish sufficient collective authority and mutual trust to build up “the usages and rules of social life,” in turn enabling words to make sense. Building on this speech/ritual coevolution idea (Knight 1998), we have taken the risk of specifying, in some detail, the world’s first metaphor, identifying it as a gendered ritual performance in which the core principles of primate politics are overturned.

We now have a repertoire of shared, mutually recognizable symbolic gestures, dance steps, and snatches of song. This is not yet syntactically complex language, but we are well on the way. For grammaticization to get under way, according to Heine and Kuteva (2007), you need only one thing: freedom to innovate—freedom to “say” one thing while “meaning” another. Imagine an ancestral community that had recently converged on a few noun-like lexical items, such as “dance,” “spear,” or “fire.” What would stop speakers from using these conventionalized sounds or gestures, where necessary, as verbs? Only if you were worried about grammar—only if you had the “noun” concept already in your head—would this pose any difficulty. Heine and Kuteva insist that categories like “noun” and “verb” arise out of usage; they certainly do not need to be hardwired in the human brain from the start.

Over time, as the functions of words diversify, they become subject to subtle changes in the way they can be deployed. Preferences become habits, and habits become grammatical rules. Within a few generations, the community will have constructed for itself a fully grammaticalized language. “The speed of the emergence of the first grammar at the inception of language is astronomical in comparison to the speed of Darwinian evolution” (Li 2002:90).

This returns us to our opening question: what was the mysterious factor relentlessly blocking any hint of grammaticalization throughout the greater part of hominin evolution? The age-old obstacle, we have argued here, was the burden imposed on all signals to incorporate some costly component to demonstrate reliability. For as long as humans were restricted to such signals, there was no foundation on which grammaticalization could build. There is no fast, efficient, zero-cost way to overcome mistrust. On the other hand, as Steel (2014) points out, there would be no grammaticalization if efficiency did not come first. To demand reliability is to rule out efficiency and, by the same token, stop in its tracks any known process of grammaticalization. Roars, screams, pant-hoots, and comparably costly signals are just not the kind of entities that can be reduced, combined, or recursively structured in the manner that grammaticalization requires.

For grammar to evolve, speakers must first be liberated from primate-style worries about reliability. Listeners must be prepared to give speakers the benefit of the doubt, evaluating truth not signal by signal but holistically, postponing judgement until the entire utterance is complete, focusing at each point not on surface meanings but on underlying communicative intentions. The liberating freedom to “lie” not only depends on the speaker but also presupposes encouragement and trust on the audience’s part. Narrative cannot evolve without this precondition, and neither can grammar. Far from punishing imaginative creativity, sympathetic listeners must go out of their way to reward it, valuing fictions, deviations, and even apparent errors as cues to what speakers may have in mind.

Conclusion: The Breakthrough

As a constraint on speculation, we have tried to imagine an overarching metaphor—an overarching “falsehood”—of such value to a community that it somehow survived, providing a template from which other metaphors could then be derived. Above all, we explain why the falsehood was not immediately rejected, as signal evolution theory would predict. A winning ritually enacted falsehood, we have argued, was that hunter-gatherer women are game animals, their blood (fertility) in-separable from that of the hunt. This strange fiction succeeded because it was collective, essential to the survival of that collective—and aimed at an “enemy” who ultimately had good reason to collude.

These are tight constraints—so tight that, in the animal world, they exclude the very possibility of language. Since language exists, the solution must somehow have been found. Returning to the BaYaka, let us look again at Ngokir. Those women fresh from “singing for their lives” in the forest have now returned to camp. They redirect their singing internally, reversing its focus from outside enemies to ones inside—now their own menfolk, especially potential alphas. They do all they can to demonstrate erotic desire (saying “yes”) while defying male sexual desire (“not now”). By publicly controlling access to their bodies, they deny alpha males access to them through dominance.

At this point, something without evolutionary precedent occurs. The “enemy” reverses his position and joins in. There are good Darwinian reasons why men might accept “defeat” at the hands of women who are nursing their own genetic offspring. Any strategy likely to benefit a man’s offspring must, by definition, benefit his genes. It is men’s willingness to yield to this logic that distinguishes them as fully human for the first time. Just as the fear grin reverses its meaning into the smile and vocal mobbing reverses into socially inclusive laughter, so women’s defiant, boisterous singing and dancing—designed to make sexual violence unthinkable—collapses and reverses, yielding something else. That other thing, we suggest, is language-based human society.

In summary, we propose that a principle of reversal applied consistently down the generations explains the key steps
by which speech progressively emerged from a point of departure in costly signaling. Resistance to primate dominance drove the evolution of distinctively human eye morphology, social cognition, laughter, play, and music—all central to the most liberating development of all, the transition to language. Our hypothesis respects the constraints of Darwinian signal evolution theory while explaining how it can be brought into alignment with metaphor and grammaticalization theory. Finally, we have respected the ethnographic record while providing a set of predictions testable in the light of recurrent structures of egalitarian African hunter-gatherer ritual and belief in addition to data from archaeology, population genetics, and the fossil record of human evolution.

We have argued on the assumption that language is a system for navigating within a shared virtual world. We have brought together a range of apparently disconnected topics—most centrally singing, dancing, and other forms of communal ritual action—because such activities are necessary to conjure up that world of shared imagination in which alone language can thrive. In an evolving hominin species, language will not even begin to evolve unless and until intensified levels of community-wide trust and a shared virtual domain have already been put in place.

Acknowledgments

J. Lewis’s fieldwork was supported by Wenner-Gren grant 569S and Horniman and Swann Fund scholarships. Fieldwork began during 1994–1997 and has continued with annual or biannual return visits since then.

Comments

Wendy James
Emeritus Professor of Social Anthropology and Fellow of St. Cross College, University of Oxford, Oxford OX1 3LZ, United Kingdom (endy.james@anthro.ox.ac.uk). 10 XI 16

The field of anthropology as a whole has always needed hard evidence from a variety of scientific research methods but, at the same time, has been dependent on the creative, even poetic side of interdisciplinary interpretation. The rate at which astonishing new findings about the nature of our humanity are emerging these days from archaeology, genetics, neuroscience, and so on is without precedent (see journals such as Nature, Scientific American, or New Scientist). There is also plenty of collaboration between archaeologists and the behavioral sciences: a fine example is the British Academy’s seven-year project Lucy to Language: the Archaeology of the Social Brain, directed by Robin Dunbar. However, when it comes to the social and cultural side of anthropological studies, interdisciplinary crossover and mutual insights have been relatively rare.

This article is therefore especially welcome. It rests on collaboration between the radical thinker Chris Knight on the beginnings of human sociality, on the one hand, and the sustained field research of Jerome Lewis, on the other, in pursuing parallel themes over many years among the BaYaka hunter-gatherers of the Congo Basin. The key theoretical focus of the article is on the way in which language, along with other human signaling and communication systems, does not simply act to convey information in a straightforward way, in the way that a dog might bark in the night to warn you of a stranger. Language makes things up: it speaks clearly only to those who understand its conventions, while confusing or deceiving others who do not. Knight and Lewis argue that some of the first deceptive signals were intended by early hominin hunters to confuse potential prey, while collective choring, especially by females, could have developed as a way of giving the impression that the group was larger and more powerful than it actually was. Signals thus originally intended as deception and directed at sources of danger could be understood differently within the group, and from this foundation, we might see how further shared understandings of sound and gesture might lead to metaphor and grammar.

There is, of course, a reluctance among social anthropologists to return to old-fashioned assumptions about the primitive character of tribal ways that "we" have now put behind us. But the coherence of modern Homo sapiens (at least predating the latest emergence from Africa) is accepted today in ways that enable us to justify renewed comparisons with our ancestors on the social and cultural front. Perhaps we can begin regarding them, and even their immediate ancestors among Homo heidelbergensis (although I do not claim to speak with authority on dates), as being more like us than was once thought. Lewis’s detailed studies of central African song, dance, and ritual enactments bring to life an exercise in the making and meaning of such forms of expression that our authors here would regard as reflecting some of the basic, and possibly ancient, characteristics of human language and communication generally.

Central to this approach is the shared, collective, and coalitionary side to so much of our human communication, indeed approaching the "political" from the earliest times. Our authors assume that the well-documented challenges of power between alpha males in ape-style social lives would have led, in early human populations, to wider patterns of collective action among lower-status individuals, including by females, as population groups grew in size and complexity. One of the key shifts to greater female interaction, in particular, was the growing advantage of collaborative parenthood, especially with the growing need for assistance with the increased difficulty of giving birth to large-brained offspring and their extended need for care during infancy and youth. Such care would also have to include extra provisioning by the males in the way of food resources and protection. Our authors would trace back to such roots the elaborate myths and metaphors of songs and dances, still found today, that represent the fertile woman in the ritual role of game animal—something common among several groups of central...
African hunter-gatherers but also found in southern Africa and elsewhere.

One striking point where the archaeological evidence does seem to make live contact with our sense of what is special to the lives and experiences of humans, as distinct from nonhumans, even beyond stone tools and weapons, are the increasingly ancient findings of pieces or signs of red ochre. It is difficult to avoid the conclusion that red ochre, and perhaps other colors too, were used to mark objects or people and thus confer “identity” on an object or an individual within some created system of relationships between people who share an understanding of the whole. From the general perspective of social anthropologists, and certainly from the perspective of the authors of this article, this is a special type of discovery. The use of red pigment seems to demand that we have to assume possibilities of human-style sociability indicated by “symbolic” signs and ceremonial activity, including perhaps vocal sounds, gesture, and a game-like give-and-take in communication not only between individuals but on behalf of recognized groups. In addition to perhaps newly emergent groups based on gender, there might well emerge the kinds of reciprocity based on mutual exogamy and continuing relations between the generations, especially on the maternal side.

In the early days of anthropology, it was commonly assumed that documenting remote present-day tribes was a way of illustrating ways of life now virtually extinct, the peoples so distant from ourselves that they really were quite different beings. But with the increasingly widespread and personal character of field research from the early twentieth century onward, unsuspected depths of sophistication have been revealed in human language, imagination, and behavior across the world. Nowhere is this clearer than in the work of today’s field anthropologists, such as Lewis and other colleagues referred to in the paper; together with the theoretical insights of Knight and linked ideas of the linguists and philosophers also referred to, this collaborative piece suggests just how much we can learn from contemporary “ancient others” about ourselves.

Camilla Power
Anthropology Programme, School of Social Sciences, University of East London, London E16 2RD, United Kingdom (c.c.power@uel.ac.uk). 23 XI 16

This article has rare interdisciplinary range, proposing a synthesis of real originality and counterintuitive insight. The authors’ position, that we cannot explain the evolution of language by itself without understanding evolving human social and political contexts, is surely correct. If language is for navigating a shared virtual world, as Knight and Lewis claim, it could not even begin to evolve until communal ritual action—a kind of pretend play—had already begun constructing such imagined worlds.

It is refreshing to see social anthropologists grappling with evolutionary theory, the fossil record, life-history theory, and up-to-date cognitive linguistics to reconstruct the likely conditions under which language first evolved. The authors invoke central features of hunter-gatherer ethnography, especially ritual and play, to suggest how grammaticalization processes, now well understood, may have taken off from the “world’s first metaphor.”

To explain the establishment of distinctively human levels of group cooperation, too many theorists (e.g., Bowles 2009; Pinker 2011; Tomasello et al. 2012) have fallen back on the fashionable “war leads to in-group morality” argument. This is not a solution that fits with what we know of egalitarian African hunter-gatherers. By contrast, playful war between the sexes, as exemplified in this article, not only matches the ethnographic evidence but provides a plausible mechanism for generating moral norms both within and between groups. In particular, the gender conflict model, with sexual counter-dominance curbing the dominance strategies of alpha males, directly accounts for rules governing sex.

Warfare itself is not known to generate these. The specific FCC model argued by the authors predicts sexual morality of the kind found in initiation rituals and enforced through hunter-gatherer traditions of bridesservice.

The article makes good use of Lewis’ fieldwork among the BaYaka, although more might be done to justify the relevance of forest hunter-gatherer data for evolutionary scenarios. During fieldwork with the savannah-dwelling Hadza, I witnessed ritual gender contests strikingly similar to those described here as Ngoku, with specific gender-reversal or “wrong sex” being a key feature of the Hadza myth of epeme (Power 2015).

The article is valuable simply for its documentation of Ngoku and the presentation of gender ritual as a form of communal play (massana). Religion has often been cited as critical to the evolution of human hypersociality and symbolism (e.g., Deacon 1997; Rappaport 1999; Sosis and Alcorta 2003), but these scenarios tend to be colored by our late-historic experience of religion as solemn, patriarchal, hierarchical, and inflexible. The playfulness and creativity of African hunter-gatherer gender ritual surely give us greater insight into the nature of religion when language was beginning to evolve.

Central to the authors’ argument is what they term “the principle of reversal.” By this, they mean a consistent logic according to which relaxed social conditions transform anxiety into relief, the primate fear grin into the human smile, vocal mobbing into human relaxed laughter, and aggressive warfare into play fighting and “let’s pretend.” The authors make the crucial point that each of these reversals is revolutionary in that, once social conditions are transformed, negative reflexes can change into their complete opposite.

Against this background, the authors take traditional assumptions about the linguistic function of the tongue and turn them upside down. Where previous theorists have assumed that the ape or early hominin tongue lacked sufficient flexibility for speech, Knight and Lewis say that it was already too flexible, given the requirement for vocal signals to be reliable. Paradoxically, in other words, it is precisely owing to
its flexibility that the primate tongue plays no role in vocal communication.

Only the drastic lifting of reliability constraints could have liberated the tongue to take on its current function as the central speech articulator. Before this development, use of the tongue to modulate vocal signals must have been for purposes other than speech. Deceptive use of the voice as an “evolutionary surprise attack” directed against animal species sounds like a good idea. Are there possible phylogenetic comparisons with other vocal mimicking species—birds and cetaceans, for example? It would strengthen the argument if this could be shown. What marks out those species from their nonmimicking sister taxa?

To explain why apes do not even point, the authors accept Tomasello’s argument that they are simply too competitive. In their own social world, primates are forced to vocalize in ways that carry conviction, which means tying their signals to emotional and bodily states and, for that reason, minimizing flexibility. The authors offer a convincing account of how, during human evolution, vocal control crept into our ancestors’ repertoire. First came tuning the voice to others in choral singing, as female and juvenile hominins combined their voices to keep dangerous animals away. Then came deceptive mimickry, as hominin males became increasingly sophisticated hunters. The argument is that, taken together, these male and female vocal strategies preadapted the human vocal apparatus for speech.

The second part of the article offers a rich, ethnographically informed account of how metaphor and symbolism arise out of ritual processes. Drawing on philosophy, social anthropology, and evolutionary linguistics, Knight and Lewis provide an intellectually satisfying account. They conclude by asking what it was that first licensed the dynamics of grammaticalization familiar to historical linguists to get under way. The crucial new ingredient, they answer, was ritualized public trust. With listeners for the first time placing trust in one another’s communicative intentions, metaphorical and other figurative usage got under way.

Language does not fossilize—so how can we ever know how it evolved? Because this model embeds language origins within a wider theory encompassing the emergence of ritual, morality, kinship, and economics, it can be tested against evidence in a number of domains: archaeology, genetics, and ethnography. Have we got an alternative hypothesis of comparable ambition and rigor?

**Slawomir Wacewicz**

Center for Language Evolution Studies, Faculty of Languages, Nicolaus Copernicus University, C332, Bojańskiiego 1, 87-100, Toruń, Poland (wacewicz@umk.pl). 11 XI 16

Trust, Metaphor, and the Evolution of Language

So, why is it that “out of 220 primate species, we are the only one that talks”? Standard, intuitive approaches to explaining the absence of language in nonhuman primates have traditionally involved pointing to the anatomical factors, such as in the configuration of the vocal tract, or to cognitive-conceptual ones, such as the difficulty of mastering the arbitrary signifier-signified relation of linguistic reference. But very fine control over vocalization is exemplified in other mammalian taxa, and—perhaps more fundamentally—the auditory medium is not a sine qua non, as demonstrated by numerous deaf communities across the world, where fully linguistic communication works very well in the visuomotor channel. On the cognitive side, nonhuman animals are generally extremely proficient at matching arbitrary signals, vocal or otherwise, to their referents via a variety of simple conditioning mechanisms (which, admittedly, is still far from true symbolic reference; cf. Deacon 1997 in the language origins context). It is not that those differences are unimportant—rather, they only become relevant once something much more fundamental is in place: a platform of trust (Dor et al. 2014). Although it is both interesting and challenging to explain how the signifiers first came to be associated with their signifieds in a communicative act, the true challenge is to explain how such signals first came to be believed at face value, rather than promptly ignored as candidate lies.

Thus, Knight and Lewis undoubtedly begin from the right starting point. The main transition—the “missing link” that they identify—is not cognitive but sociocognitive (i.e., establishing a platform of trust and thus providing the motivation to exchange honest, rather than deceptive, signals). This connects the field of language origins research (although “language evolution” is the preferred term in the field; see Dediu and de Boer 2016) not just to cognitive evolution but to the enormous bodies of recent literature on cooperation, altruism, and evolutionary stability. Of course, this point is widely accepted by the insiders (e.g., Fitch 2010; Hurford 2007; Zlatev 2016). However, this point is still worth stressing repeatedly, because it is dramatically underappreciated in at least some closely relevant disciplines (as one example, such reflection is completely absent even from those linguistic textbooks that do take up the question of language origins; see Wacewicz et al. 2016).

The implications are ubiquitous and profound—and still not fully appreciated. It is particularly noteworthy that many of the key “design features” of language (see Hockett 1960, but also Wacewicz and Żywiczyński 2015 for partial criticism)—the key differences between language and animal communication—stem from this honesty constraint, in the sense that there is only room for them to emerge once that constraint is lifted. Examples include compositionality, arbitrariness, and displacement, which can only permeate a communication system if they are placed on top of a preexisting platform of trust. For example, combinatorial signals could easily convey compositional meanings, but when each of the meanings requires independent costly evidence to be accepted, the costs of complex messages quickly spiral out of control. Likewise, recent studies confirm displaced cognition in apes (e.g., Osvath and Osvath 2008), so the problem with displaced communication may lie not so much in the difficulty of conjuring up...
entities not present in the immediate here and now but rather in the difficulty of believing in them in the absence of any immediate evidence.

To this list, Knight and Lewis add metaphor. Here “metaphor” goes beyond the surface layer of nonliteral phrasing, because what they refer to is the cognitive-linguistic sense of Lakoff and Johnson (1980)—conceptual metaphor as a tool in structuring thought. Actually, Lakoff and Johnson’s concept of metaphor reaches even deeper, as it pervades all levels of linguistic expression, including, for example, coughing basic temporal and other abstract relations in terms of somewhat more tangible spatial relations. So “a long time,” “above average,” and so on are also metaphorical, and as such, they instantiate what Knight and Lewis call “literal falsehoods,” which can only be believed with aid from this platform of trust that we so often just take for granted (cf. Dahl and Adachi 2013 for some evidence that metaphorical mappings are not cognitively inaccessible to chimpanzees).

While the skeletal structure of Knight and Lewis’ account is robust, the complete story is an intricate one with many precarious points, rich in detail that, to a considerable degree, is speculative and probably nonessential to the main thread. Some of the connections appear tenuous and the leaps too large. For example, the fact that both play and language rely on “creativity” is hardly reason enough to link the two. As to the depigmentation of the sclera of the eye (a trait apparently unique to humans among primates; although see Perea 2016), there is little doubt about its evolutionary import for the genus Homo; however, it is far from clear that it indeed developed in the context of mother-infant interactions (which would seem to predict reverting to the primate baseline in adult males). Similarly, the hypothesis about the dual origins of human vocal prowess—mimicking the sounds of the forest in males but choral singing in females—would seem to predict differences in vocal capacities between the two sexes. For the claims about the risk of predation from large felines as a motivation for choral singing, quantifiable data, such as mortality rates, would help estimate the strength of this particular selection pressure. Finally, some of the “testable predictions” do not seem to have a true predictive nature; they are indeed valuable observations consistent with the account, but because they are made post factum, we cannot be sure whether this compatibility is systematic or merely contingent. In particular, I would welcome more evidence for the controversial FCC (Power 2014) before it is accepted.

Reply

We are grateful to the commentators for their care in detailing both the strong and weak points in our argument. We are naturally delighted with the positive tenor of all three responses, which summarize our central thesis in some ways better than we managed to do ourselves. We especially appreciate the consensus that language evolution presupposes biologically unprecedented levels of public trust. Given the overwhelming predominance of formalist perspectives within theoretical linguistics, it has proved difficult, until now, to get this point across. But as Slawomir Wacewicz points out, the presence or absence of trust has always been known among insiders as absolutely pivotal, despite the fact that language origins textbooks and popular debates routinely pass over this whole issue.

Wendy James makes the important point that misplaced political correctness has previously inhibited anthropologists from placing hunter-gatherers in any kind of evolutionary framework. The motivation was understandable: it seemed important to avoid collusion with social Darwinism’s treatment of farmers and city dwellers as civilized by comparison with the supposedly primitive mental and cultural level of hunter-gatherers. The result has been intellectually catastrophic as well as politically ironic. It meant that we stopped listening to the voices of real hunter-gatherers, failed to appreciate their deeply communal values, and blocked our ears to the crucial role of laughter, singing, playful ritual, and other expressions of joyful collectivity in achieving an egalitarian lifestyle every bit as cultural and civilized as our own. Although it was this lifestyle, rather than our current one, that shaped our emergence as modern humans, we have chosen to forget it. Unconsciously internalizing the assumptions of social Darwinism, politically aware theorists faced with evolutionary questions would feel deeply uncomfortable about hunter-gatherers, valuing instead what they might learn from informants deemed to be “like us”—people such as farmers, city dwellers, or students consulted in campus questionnaires. As a result, the horizons of evolutionary theorists became restricted to the prison of our own deepest cultural assumptions. We thank Wendy James for giving us the opportunity to elaborate this point.

In the same spirit, Camilla Power endorsing our argument that territorial warfare is unlikely to have generated the conditions for language to evolve. A more likely candidate, she agrees, is gender conflict of the kind characteristic of extant African hunter-gatherers—playful gendered resistance to dominance behavior in males. Power notes that our hunter-gatherer evolutionary ancestors occupied open savanna landscapes rather than deep forests, calling into question the relevance of the lifestyle of the BaYaka and other Congo Pygmy hunter-gatherers. We take this point, but as Power herself reminds us, the savanna-dwelling Hadza share many of the same structural features as the Bayaka, as do the Kalahari Bushmen and other southern African hunter-gatherer groups.

Slawomir Wacewicz is essentially supportive but feels that our argument has “many precarious points.” We take each one in turn.

Concerning play, we would point out that, unlike chimpanzees, children are born with an appetite to read and express communicative intentions and, in the process, to acquire the...
words and grammar of their natal tongue. If we regard this as genetically evolved, we have no choice but to assume some precursor. Primate social play seems to us the best candidate for a number of reasons. In our article, we build on Gregory Bateson’s insight that a playful “nip” between two young monkey playmates in some ways anticipates symbolism. Although this is not yet language, we see in the primate play instinct one of the necessary precursors of human symbolic communication.

Wacewicz pictures our discussion of primate play as if it were a separate claim, unconnected with reverse dominance, cooperative eyes, predation pressure, and the other topics that we cover in our article. Clearly, we could have done more to convey how our arguments interconnect.

For example, Wacewicz agrees with us that trust is vital but does not see the relevance of primate play to the rest of our argument, as if we were simply asserting that both play and language rely on “creativity.” In fact, we were arguing that primate play has the potential to evolve in language-like directions but fails to do so because sex gets in the way. As they mature into adults, nonhuman primates become less and less trustful and playful with each other and more and more sexually suspicious, competitive, and demanding of reliability in their social signals. In a conflict-ridden world, the incentive to rely on any innate symbolic capacities diminishes. In humans, evidently, this entire trajectory must have been reversed, the playful instincts and capacities of the young increasingly finding scope for expression in an adult life based on the use of symbols. Only a profound transformation of sexual dynamics could have led to this uniquely human development.

Slawomir Wacewicz interprets us as claiming that our species’ uniquely transparent eyes evolved exclusively in the context of mother-infant interactions. But our whole argument is that, during human evolution, the trusting, playful relationships that nonhuman primates experience only as youngsters extended increasingly into adult life. The solidarity achieved by females in forming childcare alliances was subsequently extended as their primary means of securing male provisioning and outlawing primate-style dominance as a means of gaining reproductive success. So the cooperative parenting strategies we document did not simply affect infants but led to an entire social system based on counterdominance, reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and corresponding relationships of mutuality and trust. It is against this background that we explain reverse dominance, and correspond

Wacewicz highlights the absence of mortality data proving lions’ habit of eating people on moonless nights. In fact, our main reference (Packer et al. 2011) provides detailed statistics for Tanzania, where lions attacked 1,000 people between 1988 and 2009. Over two-thirds of these attacks were fatal, and the victims were eaten. The vast majority of victims were attacked after dark. The hourly distribution of attacks each night across the lunar cycle showed victims to be significantly more vulnerable during the darkest part of each month and night. Because lions prefer attacking victims in total darkness, attack rates varied strikingly with the phase of the moon. Hourly attack rates were two to four times higher in the first 10 days after the full moon than in the 10-day period before the full moon.

But Wacewicz’s overarching point is that our core argument—that language depends on trust—might better have been made without all the speculative detail we provide. To this, we can only answer that speculative details—predictions derived by following the internal logic of an argument—are what make a scientific hypothesis interesting and, above all, testable.

—Chris Knight and Jerome Lewis

References Cited


