THE UNIVERSITY OF AUCKLAND
LIBRARIES AND LEARNING SERVICES

ResearchSpace@Auckland

Journal Article Version

This is the publisher's version. This version is defined in the NISO recommended practice RP-8-2008 http://www.niso.org/publications/rp/

Suggested Reference

Corballis, M. C. (2014). The gradual evolution of language. *Humana.Mente Journal of Philosophical Studies*, 27, 39-60. Retrieved from http://www.humanamente.eu/PDF/Issue27_Papers_Corballis.pdf

Copyright

Items in ResearchSpace are protected by copyright, with all rights reserved, unless otherwise indicated. Previously published items are made available in accordance with the copyright policy of the publisher.

This is an open-access article distributed under the terms of the Creative Commons Attribution License

http://www.humanamente.eu/callforpapers.html

https://researchspace.auckland.ac.nz/docs/uoa-docs/rights.htm

The Gradual Evolution of Language

Michael C. Corballis [†]
m.corballis@auckland.ac.nz

ABSTRACT

Language is commonly held to be unique to humans, and to have emerged suddenly in a single "great leap forward" within the past 100,000 years. The view is profoundly anti-Darwinian, and I propose instead a framework for understanding how language might have evolved incrementally from our primate heritage. One major proposition is that language evolved from manual action, with vocalization emerging as the dominant mode late in hominin evolution. The second proposition has to do with the role of language as a means of communicating about events displaced in space and time from the present. Some have argued that mental time travel itself is unique to human, which might explain why language itself is uniquely human. I argue instead that mental time travel has ancient evolutionary origins, and gradually assumed narrative-like properties during the Pleistocene, when language itself began to take shape.

Keywords: evolution, gesture, hippocampus, language, mental time travel, mirror system, speech.

1. Introduction

> He thought he saw a Rattlesnake
> That questioned him in Greek.
> He looked again, and found it was
> The Middle of Next Week.
> "The one thing I regret," he said,
> "Is that it cannot speak!"
> —from *The Gardner's Song*, by Lewis Carroll

Language poses a substantial problem for the theory of evolution. It is a complex faculty, yet seemingly unique to our species. Members of other species

[†] School of Psychology, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand.

communicate with one another, but their communications seem to lack both the open-ended quality and propositional structure of human language, and their vocal communications also lack the purposeful, intentional character of human discourse. Nonhuman species do not seem to exchange information about events, or about the nature of the world; they do not tell stories. According Darwin's theory of evolution through natural selection, complex traits emerge incrementally, in piecemeal fashion, yet language seems distinctive in that appears to be fully formed as a complex, integrated whole. The task of explaining how language might have emerged through Darwinian evolution is therefore a challenging one.

Some have responded to the problem by proposing that language did not evolve in Darwinian fashion, but instead emerged in a single step. In the view of Noam Chomsky, the most prominent of contemporary linguists, this was indeed a profound change, and one that could not have occurred incrementally. It was, moreover, merely a by-product of a more fundamental restructuring of the human mind, involving two "paths," one intellectual and one moral:

> One path seeks to understand more about language and mind. The other is guided by concerns for freedom and justice. There should be some shared elements, in particular, what the co-founder of modern evolutionary theory, Alfred Russel Wallace, called 'man's intellectual and moral nature': the human capacities for creative imagination, language and symbolism generally, interpretation and recording of natural phenomena, intricate social practices and the like, a complex of capacities that seem to have crystallized fairly recently among a small group in East Africa of which we are all descendants, sometimes called simply 'the human capacity'. The archaeological record suggests that the crystallization was sudden in evolutionary time. Some eminent scientists call the event 'the great leap forward', which distinguished contemporary humans sharply from other animals (Chomsky, 2007, p. 3).

Despite Chomsky's reference to Wallace, the notion of the "great leap forward" is profoundly at odds with evolutionary theory, which holds that evolution occurs in small increments. But Chomsky's objection to an evolutionary account runs deeper. He suggests that language cannot have evolved through natural selection because the symbols and concepts we use have no external reality, and therefore could not have been shaped by environmental contingencies. The emergence of what he calls I-language–the internal language of thought–must have occurred entirely within the brain, without reference to the external environment. In this respect, «natural language diverges sharply ...from animal communication, which appears to rely on a one–one relation between mind/brain processes and 'an aspect of the environment to which these processes adapt the animal's behavior'» (Chomsky, 2007, p. 10). Language as a form of communication is in this view simply a by-product of a fundamental restructuring of human thought. It depends on the invention of E-languages–external means of communication–that allow one person's thoughts to map onto the thoughts of another. Language depends on so-called "theory of mind," the understanding of what others are thinking, rather than on reference to the external world.

This notion of a restructuring–the "great leap forward"–has been supported by a number of archaeologists, based on evidence from hominin artifacts of a profound transformation at some point within the past 100,000 years, perhaps even as recently as 50,000 years ago (Klein, 2008), with striking advances in technology and evidence of symbolic representation. The archaeologist Ian Tattersall (2012) writes:

> Our ancestors made an almost unimaginable transition from a non-symbolic, nonlinguistic way of processing information and communicating information about the world to the symbolic and linguistic condition we enjoy today. It is a qualitative leap in cognitive state unparalleled in history. Indeed, as I've said, the only reason we have for believing that such a leap could ever have been made, is that it *was* made. And it seems to have been made well *after* the acquisition by our species of its distinctive modern form (Tattersall, 2012, p. 199).

Such proclamations have an almost Biblical sweep, perhaps owing more to wishful thinking than to a critical appraisal of the evidence. Indeed, not all are agreed that a sudden transformation took place within the past 100,000 years. Some have proposed a more gradual development of tools and other artifacts from the Middle Pleistocene, which dates from around 750,000 years ago (McBrearty, 2007; Shea, 2011). The "great leap forward" also seems to deny Neanderthals human-like cognition and language, yet the Neanderthals had brains as large as those of humans, perhaps slightly larger, and we know that early humans did interbreed to some extent with the Neanderthals before they died out some 30,000 years ago (Green et al., 2010). A recent review suggests that Neanderthal language and culture may not have differed substantially from those of *Homo sapiens*, even raising some question as to they were actually distinct species (Johansson, 2013). Yet if we ignore the limited amount of interbreeding, we must go back some 500,000 years to find the common ancestor in the Middle Pleistocene. It seems unlikely that a sudden genetic transformation within the past 100,000 years so transformed *Homo sapiens* as to create a species cognitively and morally unique, and sharply differentiated from the equally large-brained Neanderthals.

The notion of the great leap forward, then, is counter to evolution by natural selection, as proposed by Darwin. Indeed, if true, it might even herald the demise of the Darwinian theory itself, for Darwin (1859) himself wrote in *Origin of Species* as follows:

> If it could be demonstrated that any complex organ existed, which could not possibly have been formed by numerous, successive, slight modifications, my theory would absolutely break down. But I can find no such case (Darwin, 1859, p. 158).

Is language, then, the case that Darwin feared? But as Pinker and Bloom (1990) pointed out, we should not give up an evolutionary account lightly; to them the idea that language evolved through natural selection is inescapable:

> The only successful account of the origin of complex biological structure [such as language] is the theory of natural selection, the view that the differential reproductive success associated with heritable variation is the primary organizing force in the evolution of organisms (Pinker & Bloom, 1990, p. 708).

2. Toward an evolutionary account

2.1 Nonhuman animals don't lack concepts

As noted above, Chomsky argues that human concepts and symbols have no direct reference to the external world, and that this distinguishes them from various sounds and movements that underlie animal communication. Evidence from animal communication suggests otherwise. Even Darwin (1872) noted that animals can effortlessly learn to associate myriad sounds and signals with 'general ideas or concepts' (p. 83). These sounds and signals, moreover, typically bear no direct relation to the aspects of the world they signal, and in that respects are like spoken words. For instance Cheney and Seyfarth (1990) famously observed that vervet monkeys produce a variety of different calls to indicate different predators, such as a snake, leopard, or eagle, suggesting that the calls have different meanings. It is now known that many other primate species, including chimpanzees, produce different calls to express different meanings (see Cheney & Seyfarth, 2005, for review).

Although primate calls suggest a degree of conceptual representation, they seem to lack the flexibility and intentionality of human language, where there seems no limit to the number of things–objects, qualities, actions, emotions–that we can name. But the conceptual repertoire of other animals is much more evident in their ability to comprehend than in their vocal productions. Savage-Rumbaugh et al. (1998) reported that Kanzi, a bonobo, was able to follow spoken instructions in English, made up of several words, at a level comparable to that of two-and-a-half-year-old child. Kanzi is now said to understand some 3,000 spoken words (Raffaele, 2006). Even domestic dogs can rapidly learn the meanings of spoken words, even though they cannot themselves articulate them. A border collie known as Rico responds accurately to spoken requests to fetch different objects from another room, and then either place the designated object in a box or bring it to a particular person. If Rico is asked to fetch an object with an unfamiliar name, he fetches an object he has not encountered before and thereafter knows that object's name–a phenomenon consistent with what has been termed "fast mapping" (Kaminski et al., 2004). Based on similar studies, another border collie called Chaser is said to respond meaningfully to the spoken names of 1022 objects (Pilley & Reid, 2011).

To Chomsky, the "great leap forward" was fundamentally a transformation of the conceptual capacity, with communication a secondary process. Animal research nevertheless suggests a continuity between humans and other animals, and therefore an evolutionary continuity, in the conceptual repertoire. The primary limitation evident in nonhuman animals' vocal calls is therefore not so much one of representing the world, but rather one of communicating about it. Cheney and Seyfarth (2005) remark that «Animals' limited vocal repertoires are particularly puzzling because they appear to have so many concepts that could, in principle, be articulated» (p. 142).

2.2 A vocal limitation

The inability of animals, with the exception of some birds, to produce anything resembling speech, is therefore primarily a vocal limitation rather than one of conceptual understanding. Primate vocal calls probably evolved as "honest signals," for the most part innately specified, with relatively little scope for intentional control or even vocal learning. This may mean that vocalization was not a natural platform for language itself, which is heavily dependent on the capacity to learn new signals and produce them intentionally. Jane Goodall once wrote that «(t)he production of sound in the absence of the appropriate emotional state seems to be an almost impossible task for a chimpanzee»(Goodall, 1986, p.125). David Premack, another pioneer in the study of chimpanzee behavior, suggests that even chimpanzees, our closest nonhuman relatives «lack voluntary control of their voice» (Premack, 2007, p.13866).

These conclusions do need some qualification. Cheney and Seyfarth (2005) draw attention to examples of primates modifying their vocalizations, sometimes even suppressing them, depending on the audience. Vervet monkeys seldom give alarm calls when they are alone, and are more likely to do so in the presence of kin than of non-kin. Chimpanzees modify their screams when under attack, depending on the severity of the attack and their status relative to that of nearby chimps (Slocombe et al., 2010), and when encountering food chimps emit different kinds of grunts depending on the type of food (Slocombe & Zuberbühler, 2005). Such examples, though, suggest subtle changes within call types rather than the generation of new call types (Egnor & Hauser, 2004). Some modifications involve the face and mouth rather than voicing itself. For instance chimpanzees can modify emitted sounds to attract attention by vibrating their lips, as in the "raspberry" sound (Hopkins *et al,* 2007), and this call can be imitated by naïve animals in captivity (Marshall *et al.* 1999)–although these sounds depend on movement of the lips rather than of the larynx. Reviewing these and other examples, Petkov and Jarvis (2012, p. 5) write that

> [...] we would interpret the evidence for vocal plasticity and flexibility in some non-human primates as limited-vocal learning, albeit with greater flexibility via non-laryngeal than laryngeal control. But they do not have the considerable levels of laryngeal (mammalian) or syringeal (avian) control as seen in complex vocal learners.

Complex vocal learning, and therefore speech itself, appears to have evolved late in primate evolution, and was possibly restricted to the hominins–and perhaps even to our own species. I discuss this in more detail below, but first it is useful to point out that an inability to speak need not mean that language itself is ruled out, as the signed languages of the deaf remind us.

2.3 Manual communication

A more solid basis for intentional communication of learned signals may come from the hands rather than the voice. Great apes have not learned anything approaching speech, but attempts to teach them simplified forms of sign language have been moderately successful. The bonobo Kanzi communicates by pointing to arbitrary signs on a keyboard, representing objects and actions (Savage-Rumbaugh et al., 1998), and his keyboard now has over 300 signs, which he supplements these by inventing gestures of his own. The gorilla Koko is said to use and understand over 1000 signs (Patterson & Gordon, 2001). These examples demonstrate little in the way of grammatical competence, but at least show intentional use of gesture to represent objects and actions, and some limited competence at combining a few gestures to create simple requests. The productive vocabularies of Kanzi and Koko probably still fail to match their conceptual understanding–although this is also true of humans. We often find ourselves at a loss for words.

The evolutionary origins of intentional communication are probably to be found in the primate mirror system, a network in the brain that is active both when the animal observes a particular action and when it performs that action (Rizzolatti & Sinigaglia, 2010). The motor component of this system in primates lies in Area F5, which is homologous to Broca's area, the cortical area critically involved in the production of speech. In nonhuman primates, though, this area appears unresponsive to the vocalizations of conspecifics, but it is responsive to the sounds produced by manual actions, such as the noise of a stick being dropped, or the cracking of peanut shells (Kohler et al., 2002). In this respect it does not merely "mirror" an action, as often supposed, but is involved more generally in perceptuo-motor learning. The relative deafness of the mirror system to vocal production may again reflect the non-intentional aspect of primate vocalization, necessary to maintain vocal calls as honest signals that can't be faked.

2.4 Interaction between hand and mouth

Although the primate mirror system does not appear to accommodate vocalization, the perception and control of hand and mouth are tightly integrated. This is true even of the primary motor cortex, located in the precentral gyrus. Since the pioneering work of Penfield and Rasmussen (1950), it has long been held that the primary motor cortex represents simple movements of the body, from which more complex movements are constructed. It is now known that some integrated movements, especially of hand and face, are organized within the precentral gyrus. In monkeys, stimulation of the rostral part of the precentral gyrus elicits coordinated hand-to-mouth movements (Graziano & Aflalo, 2007)–movements already evident in primate newborns (Allman, 2000). In a study of humans aged from 2 to 60 undergoing operations in which the precentral gyrus was exposed, Desmurget et al. (2014) found that stimulation of some sites elicited independent movements of mouth and arm (including hand and wrist), some sites elicited coordinated movements. These included gradual opening of the mouth while the closing hand moved toward the face, as though wanting to bring food to the mouth. This movement seems to be innately programmed; ultrasound recordings show that human fetuses suck their thumbs from as early as the 11th week of gestation–and more often such the right than the left thumb (Hepper et al., 2005). Bringing the hand to the mouth seems to be the first coordinated movement to appear in development.

More complex movements, which require learning, depend on premotor areas, of which area F5 is an example. In the monkey, some neurons within this area respond to grasping with the mouth as well as to grasping with the hand (Rizzolatti et al., 1988). Petrides and Panya (2009) also identify neurons in the homologue of Broca's area in monkeys which control orofacial muscles, and identify connections from the parietal and temporal lobes that terminate in that area. They write that their findings "are consistent with suggestions that control of action and gesture may have preceded specialization for language" (Petrides & Panya, 2009, p. 13). Gentilucci has documented close correspondences between hand and mouth movements during speech itself; for instance, when uttering the syllable "ba" the mouth opens wider when the speaker grasps a larger object than a smaller one, or even when the speaker watches another person making these movements (Gentilucci & Corballis, 2006; Gentilucci et al., 2012).

Facial gestures play an important role in sign languages (Emmorey, 2002; Sutton-Spence & Boyes-Braem, 2001), and even normal speech retains a visible component. This is illustrated by the McGurk effect: A syllable (such as *da*) is dubbed onto a mouth saying another syllable (such as *ba*), and people tend to "hear" what they see rather than what was actually voiced (McGurk & MacDonald, 1976). Other studies show the parts of the brain involved in producing speech are activated when people simply watch silent videos of people speaking (Calvert & Campbell, 2003; Watkins et al., 2003). Ventriloquists know the power of vision over what they hear when they project their own voices onto the face of a dummy by synchronizing the mouth movements of the dummy with their own pursed-lipped utterances.

The visible accompaniments of speech also include expressive movements of the hands and arms. Indeed the distinction between speech and signed languages is not absolute, since speech is universally accompanied by manual gestures, and the tight synchrony between the two suggests that they are controlled by a single integrated system (McNeill, 1985). Experiments show that gestures influence the understanding of speech, just as speech influences the understanding of gestures, so the interaction is mutual and obligatory, implying that speech and gesture are «two sides of the same coin» (Kelly et al., 2010, p. 260). Modern language may actually range from pure speech, as on radio or telephone, to pure manual gesture, as in signed languages. Moreover, if prevented from speaking, people naturally invent gestural communication, which can take on grammatical properties (Goldin-Meadow et al., 1996).

These considerations support the view that language itself evolved from manual communication, and may even have evolved to a level comparable to that of modern signed languages (Corballis, 2002). Language, then, may well have emerged before speech itself became the dominant mode.

2.5 Finding voice

If language can be traced to the intentional movements of the body, with emphasis on the hands and face, we need still to explain how the voice was incorporated. This was probably not a major step, since speech involves movements of the mouth that are not themselves vocal, and indeed overlap with movements involved in eating (MacNeilage, 2008). One view of speech is that it arose from the gradual shift of expressive facial movements into the mouth itself, where they are at least partly invisible. The addition of voicing was a device to render these movements accessible to the receiver–not through sight, but through sound.

Except for humans, primates are poor vocal learners—that is, as noted earlier, movements of the larynx are for the most part involuntary and not susceptible to learning (Petkov & Jarvis, 2012). Nevertheless it may have required a relatively small step in hominin evolution to bring the larynx under the control of the intentional motor system, and enable the learning of vocal sounds. One possibility is that the capacity for vocal learning was achieved through an extension of pre-existing motor pathways involved in movement—a possibility that arises from considerations of why some birds are vocal learners and some are not (Feenders et al., 2008). As in vocally articulate birds, so in humans. The analogue of Broca's area in the macaque modulates movements of the mouth and face, but not of the larynx itself (Petrides et al., 2005), and it may have been a straightforward step to incorporate voicing into the system.

This might have been accomplished according to a parsimonious evolutionary principle whereby new cortical representations occur through the enlargement of older areas, with part of the enlarged area allocated to the new function and the remainder retaining the original function (Finlay et al., 2005). In endorsing such a mechanism for how some birds and humans evolved the capacity for vocal learning, Feenders et al. (2008, p. 21) write:

> Our results are ... concordant with the gestural origin of spoken language hypothesis whereby the motor learning ability of gestures in humans and non-human primates has been argued to be the precursor behavior of the motor learning of speech/language.

Language, then, may not have depended on a "great leap forward" that transformed the manner in which concepts are represented and manipulated in the brain, but may rather have arisen gradually through the emergence of gestural communication systems tapping into representations of the world, and that themselves go far back in evolution. Late in primate evolution, and probably after the hominins separated from their common ancestry with great apes, this gestural system incorporated vocalization into a motor control matrix that already included manual and facial movements. Nonhuman primates are indeed manipulative creatures, using the hands for grooming and extracting food, and both hand and mouth for grasping, eating, and fighting. These activities provided a natural template for the evolution of intentional communication systems, going beyond the fixed systems of vocal calls.

But there is still one ingredient that is missing from this account.

3. Toward grammar

Nonhuman primates have the capacity to generate intentional gestures, and great apes have also demonstrated some capacity to create sequential signs with some semblance of grammar. But this capacity still seems to fall far short of the human ability to create grammatically complex, novel utterances, whether gestural or vocal. Of course grammatical language may well have preceded speech itself, and its origins may well lie in bodily gesture.

3.1 Displacement and mental time travel

The grammatical, generative aspect of language may have been reliant on displacement, and the advantages to be gained from reference to events that are not-present (Bickerton, 2010; Corballis, 2011; Gärdenfors & Osvath, 2010); this in turn may have depended on the evolution of "mental time travel," a term first used by Tulving (1985) and elaborated by Suddendorf and Corballis (1997, 2007). We humans, at least, carry the ability to consciously relive past events and imagine future ones, and indeed to construct entirely imaginary scenes with no reference to specific points in time ("Once upon a time"). The construction or reconstruction of events in the mind requires the arrangement of internal representations in a sequential fashion.

It has been argued that mental time travel itself is uniquely human (e.g., Köhler, 1925; Premack, 2007; Suddendorf & Corballis, 1997, 2007; Tulving, 1985), which itself might be taken as evidence that language itself emerged only in our species, or in the now extinct hominins that preceded us. Recent evidence suggests, though, that precursors of mental time travel may go far back in mammalian evolution. Mental time travel in humans depends critically on the hippocampus, a brain structure that is activated both when people "relive" past events or imagine future ones (Addis et al., 2011; Martin et al., 2011), and destruction of the hippocampus results in an inability to recover memories of personal events or the conjuring up of possible future ones (Corkin, 2013; Tulving, 2002; Wearing, 2005). But the hippocampus itself is not unique to humans, and plays similar roles in other species.

Individual neurons in the rat hippocampus fire when the animal is in particular locations in an environment, such as a maze, suggesting that the hippocampus is involved in the construction and activation of cognitive maps of the environment

(O'Keefe & Nadel, 1978). This activity occurs while the animal is exploring a maze, but also occurs in sharp-wave ripples (SWRs) sometime after the animal has actually been in the maze, either during slow-wave sleep (Wilson & McNaughton, 1984) or when the animal is awake but immobile (Karlsson & Frank, 2009). The paths indicated by the replay need not correspond to the actual paths taken in the maze, sometimes corresponding to the reverse of those actually taken (Foster & Wilson, 2006). Sometimes they correspond to paths the animal will take takes in the future (Pfeiffer & Foster, 2013), and sometimes to paths the animal never takes at all (Gupta *et al.*, 2010). The ripples are typically compressed in time, just as our own mental travels are generally much shorter than actual ones. Summarizing these and other findings, Dragoi and Tonegawa (2013, p. 6) write:

> The existence of temporally compressed neuronal sequences that are independent of the recent experience of the animal could support the existence of the ability to travel mentally into its past as well as into the future, a sophisticated process that may underlie higher cognitive functions like memory recollection, navigational planning, imagining, cognitive map formation, and schema-based rapid learning.

And that is in the rat. This suggests, contrary to my earlier view (Corballis, 2011; Suddendorf & Corballis, 2007), that the basic ability to travel mentally in time and space goes far back in evolution (Corballis, 2013).

Of course remembered or imagined trajectories in a spatial environment do not have the complexity of human episodes, which include objects, people, animals, emotions, happenings—as well as locations in space and time—and their combinations. Nevertheless it seems likely that mental time travel evolved gradually over the tens of millions of years that separate us from our common ancestry with the other mammals, and perhaps goes further back to common ancestry with birds, which also possess a hippocampal analogue (Macphail, 2002). In creatures that move freely over the surface of the earth—and beyond—a mechanism for spatial memory and mental exploration in time and space may have been a very early requisites for successful adaptation.

3.2 The emergence of narrative

Perhaps it was the capacity for narrative, in which imagined events are woven into coherent sequences, that separated human mental time travel from that of other extant species, and led to the capacity to communicate our travels to others. The literary scholar John Niles (2010) suggests that our species should be renamed *Homo narrans*–the story tellers. Cosentino and Ferretti (2014) suggest that language itself, in the form of connected discourse, depended on an advanced capacity for navigation in space and time. The elaboration of mental time travels into narratives probably dates from the Pleistocene, dating from some 1.6 million years ago, when our hunger-gatherer forebears foraged over increasingly wide terrain for food. It would have been increasingly adaptive to relay the foraging experiences to others. Children, too, would have benefited from the tales told by the adult hunters and foragers, gaining knowledge about food sources and hunting techniques before themselves graduating to the hunt. Sugiyama (2011) suggests that the importance of stories may explain the prolonged juvenile phase in humans. The juvenile phase began to increase with the earliest members of our genus, *Homo habilis,* early in the Pleistocene, reaching a peak with *Homo sapiens* (Locke & Bogin, 2006).

The Pleistocene also saw the emergence of our genus *Homo* as intensely social creatures, occupying what has been terms the "cognitive niche" as an adaptation to a scattered and dangerous environment (Tooby & De Vore, 1987). This created safety in numbers, and also more effective methods of foraging and problem solving. The evolution of language as a way to exchange complex information was perhaps the most critical accomplishment in the establishment of the cognitive niche (Pinker, 2003). It was not simply a matter of exchanging practical information. Stories are the product of imaginary mental travels, not only into past or future and into the minds of others, but also into realms of fantasy and play. Stories continue to dominate our lives, whether in the form of plays, novels, operas, television soaps, or bedtime tales told to children.

Story telling may have originated in pantomime, as our forebears acted out their experiences, and perhaps even their fantasies. Some degree of pantomime is evident in the natural communications of apes, including orangutans (e.g., Russon & Andrews, 2011), although it is unlikely that gestural communication evolved a story-like structure until the emergence of bipedal hominins, and perhaps not even until the emergence of the genus *Homo* in the Pleistocene (Thompson, 2010). A critical development may have been the development of stone tool-making industries, perhaps at the transition from the earlier Oldowan industry to the later, more complex Acheulian industry dating from around 1.7 million years ago. Stout et al. (2008) asked three archeologists experienced in early tool technologies to make Oldowan tools or Acheulian tools, or simply strike cobbles together, and in each case measured the their brain activity using a PET scanner with a slow-decaying radiological trace. Only in the case of Acheulian tool-making did the brain activity overlap with language circuits, including Broca's area. The active areas also corresponded to the known mirror system. Unlike Oldowan tools, Acheulian tool-making involves planned sequential action. Whether pantomime evolved from the telling of stories to sequential action, or sequential action came first, is a moot point. Perhaps it was a matter of co-evolution, as the brain itself evolved to accommodate to increased cognitive and social demands.

3.3 From pantomime to language

Pantomime, though, is inefficient, and can be streamlined by a process of conventionalization (Burling, 1999), whereby pictorial or iconic representations are gradually replaced by arbitrary signals, shaped to convey information with maximum efficiency and minimum effort. In losing their pictorial quality, though, these arbitrary signals must be sustained and taught by the linguistic community, and transmitted between generations. The process of conventionalization is evident in the development of sign languages. For example, in American Sign language the sign for *home* was once a combination of the sign for *eat*, which is a bunched hand touching the mouth, and the sign for *sleep*, which is a flat hand on the cheek. Now it consists of two quick touches on the cheek, both with a bunched handshape, so the original iconic components are effectively lost (Frishberg, 1975). Albeit with a lack of modern cultural sensitivity, Darwin (1872) remarked

> on the practice of the deaf and dumb and of savages to contract their signs as much as possible for the sake of rapidity. Hence their natural source or origin often becomes doubtful or is completely lost; as is likewise the case with articulate language (Darwin, 1872, p. 62).

In this view, speech itself can be regarded as the product of conventionalization, in which the element of pantomime is effectively lost, and spoken words are sustained through cultural transmission, albeit subject to mutations than have resulted in the some 7,000 different languages in today's world. Language, then, is not due to the sudden emergence of symbolic representation, as maintained by Chomsky and others, but is rather the end process of a gradual process by which the elements of discourse have lost their iconic connection with external reality. Grammar, too, might be a result of this process. Pantomimic sequences might well have retained the sequential structure of the events they represent, but new conventions would then have been introduced to render the sequence of symbols coherent. Conventions are needed to differentiate different elements—as Pinker (2003, p. 27) put it, "who did what to whom, when, where and why". This gives rise to such grammatical contrivances as case, tense, mood, and the like. The emergence of grammar, then, can also be seen as an incremental process, building up as the means of communication became less pantomimic and more arbitrary in format.

This process in turn might well have allowed language to move from the telling of stories involving concrete objects and physical actions to more abstract accounts, of which this very article is an example. But our ability to deal in abstract concepts depends heavily on metaphor, treating the abstract as we treat real-world objects and events (Lakoff & Johnson, 1980). We *grasp ideas*, talk of being *at a crossroad in a relationship*, or *stumble toward a solution*. Gallese and Lakoff (2005) suggest how metaphorical language might have arisen from elaboration of the mirror system itself.

Grammatical language, then, need not have arisen from a great leap forward that bestowed a procedure such as Chomsky's unbounded Merge, but may have emerged from increasingly complex representations of real-world events in the form of internal narratives, and the establishment of gestural systems to communicate them. As these systems became less pantomimic and more conventionalized, rules were needed to establish convey properties of the narratives—who did what to whom, when, where and why. Such an account is similar to Chomsky's in that it is the structure of our thoughts that fundamentally controls how we communicate them, but that structure evolved gradually through primate evolution, probably accelerating during the Pleistocene, rather than as a single leap within the timespan of our own species.

Conclusions

Many of those who have speculated about the evolution of language have been misled by the assumption that language is identified with speech. The grammatical sophistication of signed languages should remind us that this is not the case. Speech bears misleading witness to language evolution because most animals do not have the intentional vocal control or capacity for vocal learning necessary for articulate speech. Animals capable of vocal learning include some birds, cetaceans, pinnipeds, and at least one elephant (Petkov & Jarvis, 2012), but conspicuous among those with little if any capacity for vocal learning are the primates. Yet great apes and even dogs appear able to learn to comprehend human vocalizations, even if unable to produce them. The limitation imposed on nonhuman animals is therefore in large part one of vocal production, and not of understanding, nor even one of conceptual representation.

The more likely platform for language evolution is the sensorimotor cortex, which mediates the learning of motor sequences. In primate evolution, this was dominated by bodily movement, and especially manual and facial action, suggesting that language itself evolved through modifications of sensorimotor cortex. The mirror system, in particular, is the most likely candidate, since it maps the perception of bodily movement onto its production—a natural basis for communication. In our hominin forebears, the incorporation of vocalization into this system seems to have been late, and possibly restricted to *Homo sapiens*, although one might surmise that it was gradual rather than sudden.

Perhaps the most important feature of language, as distinct from other forms of animal communication, is that it permits communication about the non-present, such as past events, imagined future events, or hypothetical and even impossible events. The ability to mentally represent such events includes mental time travel, whereby we form internal scenarios. Although some have suggested that mental time travel is itself uniquely human, evidence from hippocampal recordings in the rat suggest that the ability to replay, pre-play, or even construct purely imaginary trajectories in spatial environments goes far back in evolution. The complexity of these imagined events no doubt increased through evolutionary time, perhaps gaining a narrative-like character in our hominin forebears after they split from the great apes.

The adaptive advantages of communicating internal narratives perhaps surfaced during the Pleistocene, when our forebears were forced away from a forested environment to a foraging existence on the open savanna. This created an imperative for increased social bonding and the exchange of information about food sources and other components of foraging expeditions. At first, this exchange might have depended on pantomime, a gestural replay of events that occurred or of future plans. Through time, pantomime would have been gradually replaced by more efficient gestures, and sustained by convention rather than by pictorial or iconic resemblance to what is communicated. Speech was the end product of this process of conventionalization, creating a medium that was energy-efficient, and that freed the hands and rest of the body for other activities, such as tool manufacture. Removal of the iconic component also allowed different languages to mutate, increasing bonding within groups but denying access of information to other groups. There are now some 7,000 different languages.

This article is not of course the final word on language evolution, and many details need to be worked out. My main purpose, though, is to suggest that an evolutionary account is at least plausible, and we do not need to postulate any kind of miraculous "great leap forward."

REFERENCES

Addis, D.R., Cheng, T., Roberts, R.P. & Schacter, D.L. (2011). Hippocampal contributions to the episodic simulation of specific and general future events. *Hippocampus, 21*, 1045–1052.

Allman, J.M. (2000). *Evolving brains.* New York: Scientific American Library.

Bickerton, D. (2010). On two incompatible theories of language evolution In R.K. Larson, V. Déprez, & H. Yamakido (Eds.), *The evolution of human language.* Cambridge: Cambridge University Press, 199-210.

Burling, R. (1999). Motivation, conventionalization, and arbitrariness in the origin of language. In B.J. King (Ed.), *The origins of language: What nonhuman primates can tell us.* Santa Fe, NM, School of American Research Press.

Calvert, G.A., & Campbell, R. (2003). Reading speech from still and moving faces: The neural substrates of visible speech. *Journal of Cognitive Neuroscience, 15*, 57-70.

Cheney, D.L., & Seyfarth, R.M. (1990). *How monkeys see the world.* Chicago: University of Chicago Press.

Cheney, D.L., & Seyfarth, R.M. (2005). Constraints and preadaptations in the earliest stages of language evolution. *The Linguistic Review, 22*, 135-159.

Chomsky, N. (2007). Biolinguistic explorations: Design, development, evolution. *International Journal of Philosophical Studies, 15*, 1-21.

Corballis, M.C. (2002). *From hand to mouth: The origins of language.* Princeton, NJ: Princeton University Press.

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

Corballis, M.C. (2013). Mental time travel: The case for evolutionary continuity. *Trends in Cognitive Sciences, 17*, 5-6.

Corkin, S. (2013). *Permanent present tense: The man with no memory, and what he taught the world.* London: Allen Lane.

Cosentino, E., & Ferretti, F. (2014). Communication as navigation: A new role for consciousness in language. *Topoi – An International Review of Philosophy, 33*, 263-274.

Darwin, C. (1859). *The origin of species*. London: John Murray.

Darwin, C. (1872). *The descent of man, and selection in relation to sex*. London: John Murray.

Desmurget, M., Richard, N., Harquel, S., Baraduc, P., Szathmari, A., Mottolesi, C., & Sirigu, A. (2014). Neural representations of ethologically relevant hand/mouth synergies in the human precentral gyrus. *Proceedings of the National Academy of Sciences, 111*, 5718-5722.

Dragoi, G., & Tonegawa, S. (2014). Selection of preconfigured cell assemblies for representation of novel spatial experiences. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369,* 20120522. http://dx.doi.org/10.1098/rstb.2012.0522

Egnor, S.E.R. & Hauser, M.D. (2004). A paradox in the evolution of primate vocal learning. *Trends in Neurosciences, 27*, 649-654.

Emmorey, K. (2002). *Language, cognition, and brain: Insights from sign language research*. Hillsdale, NJ: Erlbaum

Feenders, G., Liedvogel, M., Rivas, M., Zapka, M. Horita, H., Hara, E., Wada, K., Mouritsen, H., & Jarvis, E. D. (2008). Molecular mapping of movement-associated areas in the avian brain: a motor theory for vocal learning origin. *PLoS ONE, 3*, e1768. doi:10.1371/journal.pone. 0001768

Finlay, B.L., Cheung, D., & Darlington, R.B. (2005). Developmental constraints on or developmental or developmental structure in brain evolution. In Y., Munakata & M., Johnson (Eds.), *Attention and performance XXI: Process of change in brain and cognitive development*. Oxford: Oxford University Press, 131-162.

Foster, D.J., & Wilson, M.A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature, 440*, 680-683.

Frishberg, N. (1975). Arbitrariness and iconicity in American Sign Language. *Language, 51*, 696-719.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensorimotor system in conceptual knowledge. *Cognitive Neuropsychology, 22*, 455-479.

Gärdenfors, P., & Osvath, M. (2010). Prospection as a cognitive precursor to symbolic communication. In R.K., Larson, V. Déprez, & H. Yamakido (Eds.), *The evolution of human language*. Cambridge: Cambridge University Press, 103-114.

Gentilucci, M., Campione, G.C., De Stefani, E. & Innocenti, A. (2012). Is the coupled control of hand and mouth postures precursor of reciprocal relations between gestures and words? *Behavioural Brain Research, 233*, 130-140.

Gentilucci, M., & Corballis, M.C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience & Biobehavioral Reviews*, *30*, 949-960.

Goldin-Meadow, S., McNeill, D., & Singleton, J. (1996). Silence is liberating: Removing the handcuffs on grammatical expression in the manual modality. *Psychological Review, 103*, 34-55.

Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behaviour*. Cambridge, MA: Harvard University Press.

Graziano, M.S., & Aflalo, T.N. (2007). Mapping behavioral repertoire onto the cortex. *Neuron*, *56*, 239-251.

Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M. et al. (2010). A draft sequence of the Neanderthal genome. *Science, 328*, 710-722.

Gupta, A.S., van der Meer, M.A.A., Touretzky, D.S., & Redich, A.D. (2010) Hippocampal replay is not a simple function of experience. *Neuron, 65,* 695-705.

Hepper, P.G., Wells, D.L., & Lynch, C. (2005). Prenatal thumb sucking is related to postnatal handedness. *Neuropsychologia*, *43*, 313-315.

Hopkins, W.D., Taglialatela, J.P. & Leavens, D.A. (2007). Chimpanzees differentially produce novel vocalizations to capture the attention of a human. *Animal Behaviour, 73*, 281–286.

Johansson, S. (2013). The talking Neanderthals: What do fossils, genetics, and archaeology say? *Biolinguistics, 7*, 35-74.

Kaminski, J., Call, J. & Fischer, J. (2004). Word learning in a domestic dog: Evidence for "fast mapping." *Science*, *304*, 1682–1683.

Karlsson, M.P. & Frank, L.M. (2009). Awake replay of remote experiences in the hippocampus. *Nature Neuroscience, 12*, 913–918.

Kelly, S.D., Özyürek, A. & Maris, E. (2010). Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological Science, 21*, 260-267.

Klein, R.G. (2008). Out of Africa and the evolution of human behavior. *Evolutionary Anthropology, 17*, 267–281

Köhler, W. (1925). *The mentality of apes*. New York: Routledge & Kegan Paul. (Originally published in German in 1917).

Kohler, E., Keysers, C., Umiltà, M.A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science, 297*, 846–848.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Locke, J.L., & Bogin, B. (2006). Language and life history: A new perspective on the development and evolution of human language. *Behavioral & Brain Sciences, 29*, 259-325.

MacNeilage, P.F. (2008). *The origin of speech*. Oxford: Oxford University Press.

Macphail, E.M. (2002). The role of the avian hippocampus in spatial memory. *Psicologica, 23*, 93-108.

Marshall, A.J., Wrangham, R.W. & Arcadi, A.C. (1999). Does learning affect the structure of vocalizations in chimpanzees? *Animal Behaviour, 58*, 825–830.

Martin, V.C., Schacter, D.L., Corballis, M.C., & Addis, D.R. (2011). A role for the hippocampus in encoding simulations of future events. *Proceedings of the National Academy of Sciences, 108*, 13858-13863

McBrearty, S. (2007). Down with the revolution. In Mellars, P., Boyle, K., Bar-Yosef, O., & Stringer, C. (Eds.), *Rethinking the human revolution*. McDonald Institute Monographs. Cambridge: McDonald Institute for Archaeological Research, 133–152.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746-748.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review, 92*, 350-371.

Niles, J.D. (2010). *Homo narrans: The poetics and anthropology of oral literature*. Philadelphia: University of Pennsylvania Press.

O'Keefe, J. & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Clarendon Press.

Patterson, F.G.P., & Gordon, W. (2001). Twenty-seven years of project Koko and Michael. In B.M.F. Galdikas, N.E. Briggs, L K. Sheeran, G. L. Shapiro, & J. Goodall (Eds.) *All Apes Great and Small, Vol. 1: African Apes*. New York: Kluver, 165-176.

Penfield, W., & Rasmussen T. (1950). *The cerebral cortex of man*. New York: Macmillan.

Petkov, C. I., & Jarvis, E.D. (2012). Birds, primates, and spoken language origins: Behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience, 4*, article 12.

Petrides, M., Cadoret, G., & Mackey, S. (2005). Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. *Nature, 435*, 1235-1238.

Petrides, M., & Panya, D.N. (2009). Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. *PLoS Biology, 7*, e1000170, http://www.doi:10.1371/journal.pbio.1000170.

Pfeiffer, B.E., & Foster, D.J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature, 497*, 74-80.

Pilley, J. W., & Reid, A.K. (2011). Border collie comprehends object names as verbal referents. *Behavioural Processes, 86*, 184-195.

Pinker, S. (2003). Language as an adaptation to the cognitive niche. In M.H. Christiansen & S. Kirby (Eds), *Language evolution.* Oxford: Oxford University Press, 16-37.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral & Brain Sciences, 13.* 707-784.

Premack, D. (2007). Human and animal cognition: Continuity and discontinuity. *Proceedings of the National Academy of Sciences (USA), 104*, 13861-13867.

Raffaele, P. (2006). Speaking bonobo. *Smithsonian Magazine,* November 2006. Online at: http://www.smithsonianmag.com/science-nature/speakingbonobo.html (accessed 25 August 2014).

Rizzolatti, G., Camardi, R., Fogassi, L., Gentilucci, M., Luppino, G. & Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Experimental Brain Research, 71*, 491-507.

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience, 11*, 264–274.

Russon, A., & Andrews, K. (2011). Orangutan pantomime: Elaborating the message. *Biology Letters, 7*, 727-630.

Savage-Rumbaugh, S., Shanker, S.G. & Taylor, T.J. (1998). *Apes, language, and the human mind.* New York: Oxford University Press.

Shea, J.J. (2011). *Homo sapiens* is as *Homo sapiens* was. *Current Anthropology, 52*, 1-35.

Slocombe, K.E., Kaller, T., Call, J., & Zuberbühler, K. (2010). Chimpanzees extract social information from agonistic screams. *PLoS ONE, 5*, e11473.

Slocombe, K.E., & Zuberbühler, K. (2005). Functionally referential communication in a chimpanzee. *Current Biology, 15*, 1779–1784.

Stout, D., Toth, N., Schick, K., & Chaminade, T. (2008). Neural correlates of early Stone Age toolmaking: Technology, language and cognition in human evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*, 1939-1949.

Suddendorf, T., & Corballis, M.C. (1997). Mental time travel and the evolution of the human mind. *Genetic, Social, and General Psychology Monographs, 123*, 133-167.

Suddendorf, T., & Corballis, M.C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences, 30*, 299-351.

Sugiyama, M.S. (2011). The forager oral tradition and the evolution of prolonged juvenility. *Frontiers in Psychology*, *2*, Article 133.

Sutton-Spence, R., & Boyes-Braem P. (Eds.) (2001). *The hands are the head of the mouth: The mouth as articulator in sign language.* Hamburg: Signum-Verlag.

Tattersall, I. (2012). *Masters of the Planet: The Search for Human Origins*. New York: Palgrave Macmillan.

Thompson, T. (2010). The ape that captured time: Folklore, narrative, and the human-animal divide. *Western Folklore, 69*, 395-420.

Tooby, J. & DeVore, I. (1987). The reconstruction of hominid evolution through strategic modeling. In W.G. Kinzey (Ed.), *The evolution of human behavior: Primate models.* Albany, NY: SUNY Press, 183-238.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *26*, 1-12.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*, 1-25.

Watkins, K.E., Strafella, A.P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia, 41*, 989-994.

Wearing, D. (2005). *Forever today.* New York: Doubleday.

Wilson, M.A., & McNaughton, B.L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science, 265*, 676-679.