Cognitive Science Online, Vol.3.1, pp.1-12, 2005 http://cogsci-online.ucsd.edu

Language Origins Without the Semantic Urge

Martin I. Sereno
Department of Cognitive Science
University of California San Diego
sereno@cogsci.ucsd.edu

Abstract

Despite the paucity of direct evidence about the origin of human language, the great intrinsic interest in this question has made it difficult for writers to resist speculating about it (Harnad et al., eds., 1976; Merlin, 1991; Deacon, 1997; Jablonski & Aiello, eds., 1998; King, ed., 1999; Knight et al., eds., 2000). The following attempts to bring a fresh perspective on this old question, using an analogy with the origin of cellular coding systems and applying it to what we know about the evolution of vocal behavior in animals. In other places (Sereno, 1991b), I have argued that DNA and protein based life and language based human thought may have enough in common as the only two naturally occurring examples of a code-using system to make it useful to take an analogical look at one system in order to make predictions about the other. Rather than rehearsing those arguments, I will only visit two jumping off points reached while developing that analogy: the difference between origin and evolution, and the foundational role of an intermediate string of "symbol representation" segments with properties partway between symbol and meaning.

Origins versus evolution

Discussions of how life came into existence (Wills & Bada, 2000) often distinguish *origin* of life from the *Darwinian evolution* of life. The core of every living cell is a system for converting genes into proteins; that is, DNA sequences into amino acid chains that spontaneously fold into the 3D molecular machinery of the cell including enzymes, receptors, force producing strands, and so on.

The problem of the *origin* of this system is not really an evolutionary problem in the usual Darwinian sense of the word. If we shrink ourselves down to molecular size and look at what cells are doing, it becomes clearer that cells have invented a new kind of molecular level intentionality as a way to partly overcome the deterministic thermodynamic buffetings to which all matter is subject that went far beyond the chemical dynamics of the landscape before there was life. This does not imply that cells create mysterious, irreducible holistic forces; in fact, we know quite a lot about how they work. But it is a natural way of characterizing what goes on in cells that distinguishes them from the prebiotic chemical cycles in clouds, rock piles, streams, beaches, and ocean floors. The prebiotic soup was already a complex, energy-dissipating system containing many different types of dynamically stable subunits. Cells, however, invented a way to encode, use, and reproduce information about how to cause thousands of different chemical reactions in this soup to happen. The tricky part is that the information, as well as all of the interpreting apparatus has to be in the soup where everything is still subject to the soup's deterministic buffetings. The cellular system speeds many chemical reactions, slows or prevents others, invents many new ones that never used to happen at all before, and above all, orders and organizes the reactions. In short, code-using cells have taken over forceful control of chemical phenomena in local regions of the otherwise still prebiotic soup. But the 'evolution' from prebiotic to biotic systems was not *Darwinian* evolution. Until the DNA code-using system was in place, Darwinian evolution as it is usually defined, "heritable variations in fitness" (Lewontin, 1970), was not possible. The central problem of the origin of the coding system in life is to try to imagine how such an intentional system could have arisen out of prebiotic situations lacking intentionality.

In thinking about the origin of language and cultural evolution, the situation is more complex since human language was built upon a pre-existing genetic system already capable of Darwinian evolution. There has been a recent revival of interest in 'evolutionary psychology' or human sociobiology, which attempts to come up with plausible scenarios for how biological evolution might have directly driven the origin of many human behaviors. One difficulty with this approach is the huge increase in the pace of behavioral change supported by the origin of language, which is what most distinguishes humans from other animals. Cultural evolution is so fast that it makes biological evolution look effectively stationary; this great difference in velocity makes it difficult for culturally transmitted memes to be fixed in much more slowly evolving genes.

It is certainly true that there is one great point of interaction between the DNA based genetic system and the language based human cultural system, which is the genetic basis of the peculiar human ability to readily learn a language. However, I think we may be able to make more progress by considering the *origin* of language as essentially a *pre*-evolutionary problem -- that is, as the second origin of a symbol using evolution supporting system, one that partly relies on DNA based symbols for its persistence, but that is largely decoupled from biological evolution in its content. In fact, human language might best be thought of as a brain operating system that allowed us to partly overcome the constraints on the biological evolution of behavior much in the way that cells have partly overcome the deterministic constraints on the 'evolution', now in the physicist's sense of 'the evolution of a dynamical system', of prebiotic soups.

The semantic urge and the 'RNA world'

I want to take issue with an assumption that lies behind almost every language origin scenario, something that could be called the 'semantic urge'. This widespread intuition grows out of the fact, just mentioned above, that the human linguistic coding system was built on top of a lower level biological coding system already capable of constructing sophisticated, nonlinguistic cognitive systems such as those in parrots and primates. The sustained goal directedness of animals makes it very hard to avoid the notion that human language must have somehow grown out of an insistent craving of inarticulate hominins to communicate complex meanings to each other ("hominins" because of a demotion of one level in the hierarchy from hominids, since chimpanzees are now thought to be more closely related to humans than to gorillas).

This is at heart a Baldwinian picture, where behavior provides a selective context that drives standard Darwinian evolution (Baldwin, 1902). As noted above, there is little tendency to fall back on intuitions like this in thinking about the origin of cellular life because cellular life had no analogous pre-existing code using system capable of goal directed behavior beneath it. Thus, one influential picture about the origin of life is that protosymbol chains emerged first without standing for anything and then only later were taken over as a code for other chains that could fold up and control chemical reactions; there is no Baldwinian urge of prebiotic soups to control their surroundings that drives the emergence of cellular symbol chains. The corresponding picture of protolanguage that we arrive at by analogy is somewhat peculiar, but it fits better with what we know about the evolution of vocal behavior in other animals. Let's first review some ideas about protosymbols at the cellular level.

The idea of an 'RNA world' (Gilbert, 1986) as a predecessor of modern DNA and protein based life was put forward soon after demonstrations in the early 1980's that RNA could act as a bona fide, enzyme-like catalyst (specifically, a spliced out segment of RNA in the single-celled organism *Tetrahymena* was discovered to fold up and catalyze RNA splicing). Three main arguments for the foundational role of RNA come from observing its current position in cells:

- RNA can act either as a 1D symbol string (mRNA) *or* a 3D controller of chemical reactions (structural RNA's, based on the protein-like ability of RNA to form precisely shaped surface cavities with high specificity for particular substrates)
- Modern cells mostly use proteins, not RNA's, to control and catalyze reactions
- The instances where RNA *is* used as a protein-like structure stand at the very center of code use in cells – splicing of code-like RNA (nucleolus, spliceosome), recognizing words in code-like mRNA (tRNA), and assembling amino acids into proteins (rRNA, SRP RNA)

This idea that the dual roles of RNA as code and catalyst might have bootstrapped life have gained support in recent years as additional catalytic RNA's were discovered, as large scale RNA's (ribosomes) were finally crystallized showing that RNA itself, not proteins, catalyzed the attachment of each amino acid onto the growing protein chain (Ban et al., 2000; Nissen et al., 2000; Yusupov et al., 2001), and most recently, when a small RNA was created (in a laboratory selection experiment) that catalyzed the attachment of an amino acid onto the small RNA itself. This last observation was particularly evocative since RNA-amino acid bonds are made and broken during each chain lengthening step in modern cellular protein synthesis (Illangasekare et al., 1997; Zhang and Cech, 1997; Jenne and Famulok, 1998).

Despite the intuitive attractiveness of the 'RNA world', however, it has turned out to extremely difficult to find plausible prebiotic synthesis pathways for nucleotides, the subunits of RNA, which stands in sharp contrast to the easy prebiotic availability of amino acids (Miller and Orgel, 1974; Schwartz, 1998). This led many origin of life researchers to search for prebiotic precursors of RNA constructed from other more easily obtainable subunits (Joyce et al., 1987). A key feature of this search, so obvious to those within the field that it is rarely explicitly stated, is to find reasons *other* than the ability to code for proteins (or to catalyze chemical reactions in a protein-like way) as to why a pre-RNA-like molecule might have come into existence. Once RNA or something like it existed, its dual role as a catalyst and a code chain could then be discovered, as it were, leading perhaps to something like an RNA world, and then finally DNA/RNA/protein life.

In turning back to language, many language origins scenarios start with a repertoire of already meaningful vocalizations like those used by many different animal species (Hauser, 1995) and then attempt to come up with a reason, typically, the semantic urge – for why they might have multiplied. Several people (Zahavi, 1993; Knight, 2002) have pointed out that there is a major problem with this standard scenario. Animal calls, such as the well studied set of vervet monkey alarm calls, are laden with emotional meaning (Cheney and Seyfarth, 1990). The usual explanation for this is that alarm calls are emitted in life and death situations, which generates strong selection pressure to maintain call reliability, and that this reliability seems to have been ensured across many species by tightly tying calls to the emotional state of the sender and the receiver. This linkage, however, presents a problem for the scenario of calls as a proto-language; the number of different emotional states is rather small, and emotional states don't follow each other in a quick, regular succession like words do.

The origin of language required the development of a large inventory of thousands of words whose meanings are made more specific by assembling them into strings that are deployed in a rapid, regular sequence at a rate of several words per second. Individual words, especially the high frequency polysemous words that are central to every language such as "over", "put", "give", "line", "big", or "hand", are freed from emotion and bleached of emotion when compared to animal calls. Certainly, some single words such as epithets can be intrinsically emotive, but these are in a small minority. Perhaps the difficulty of imagining a path from a handful of emotive calls to the five thousand word core of emotionally neutral words in human language stems from the fact that the two are phylogenetically unrelated. The analogy with the RNA world and the pre-RNA world suggests that perhaps we should instead try to find a way by which a large pool of pre-RNA-like prewords might have been generated as units that are like words in some sensory and motor respects, but that *don't* yet stand for anything.

Birdsong and language preadaptations

The relation between birdsong and speech was noted early on. Darwin, who in *Origin of Species* (1859) often discussed the relation between biological and linguistic evolution (interestingly, to argue that biological evolution might be like language evolution, not *vice versa*!), turned briefly, in the *Descent of Man* (1871), to language origins. Darwin was especially fond of scenarios in which a structure had initially evolved for one purpose only to become a 'preadaptation' for another. Breaking with a common view that language arose from a gestural substrate, he suggested that language developed out of a form of "rudimentary song," a kind of purely prosodic pre-language that conveyed emotions and other broad, unitary meanings in much the same way that pitch modulation and emphasis are used in modern speech. Darwin mentions the flashy hooting vocalizations of gibbons, which are generated during territorial and courtship displays, as something like what he had in mind, but pointed out that birdsong provided "in several respects the nearest analogy to language" (p. 55), citing the work of Daines Barrington (a century before) on the extended learning period for birdsong, the initial "babbling" stage, and the development of birdsong dialects. I think Darwin's idea should be revived but also revised in light of newer work on birdsong and primate vocalizations.

Modern research on birdsong has provided a neurobiological foundation for these earlier hunches, but has also revealed a system that looks a good deal more like human-style, ‘left-hemisphere’ speech than like the call systems of other animals, including New and Old World monkeys and gibbons (Konishi, 1985; Nelson and Marler, 1989; Doupe & Kuhl, 1999), but also the call systems of songbirds themselves, who have retained their limited set of emotional calls alongside song. There is a powerful perennial tendency outside fields explicitly focussed on evolutionary processes to think of evolution in terms of a “Great Chain of Being” and to ignore the mosaic nature of evolution. Thus, birdsong has often been dismissed as a model of human language for the reason that monkeys seem much smarter than some birds, or that monkey calls seem to have more semantic content than birdsong. In fact, the importance of birdsong in the present context (only dimly glimpsed by Darwin) is exactly the fact that a set of language-like features have evolved *in the absence of* a semantic function.

Birdsong requires a significant learning period, during the early parts of which the bird is silent. If a bird is not exposed to a tutor song within a certain early “critical period”, it will produce only a crude version of its species’ song. Normally exposed young birds initially produce sounds called “subsong” which resemble the progression of types of “babbling” in baby humans: Initially a broad range of sounds are produced, followed by an unorganized recombination of species specific song fragments, and then finally, adult song. Within a species, there are regional “dialects” which are learned from a bird's regional peers; artificial rearing experiments show that birds learn the dialect of their tutors, regardless of their genetic background. Adult song repertoires can be quite considerable; some wrens produce hundreds of distinct songs each containing 5-20 ‘syllables’, while mockingbirds produce virtually endless sequences of different syllables in variable orders. Good singers may have a thousand or more distinct ‘syllables’ (a syllable consists of a particular figure sometimes repeated once or twice; in this respect, it is unlike a phonetic syllable which consists of one or more consonants and a vowel). If a songbird is deafened before learning to sing, it will fail to produce songlike sounds as an adult. By contrast, nonsong birds and many other animals including nonhuman primates (including gibbons) that do not learn complex serial vocal patterns from their peers, still come to produce their species specific sound repertoire when deafened at birth (Merker, 2000). In many respects, it might be more accurate to call it ‘birdspeech’, since birdsong differs from human singing and musical performance in many ways, birdsong lacks a regular meter, musical tonality, and harmony.

The parallel evolution of fine grained vocal control in singing birds affords a crucial comparative perspective on the anatomical and neural constraints on auditory-motor learning and performance. Birdsong is initiated in a structure called the syrinx, which is evolutionarily related to (and controlled by the same nerve as) the tongue, but which corresponds functionally to the human larynx. Human speech sounds are generated by filtering and modulating the higher harmonics of the fundamental frequency of the vocal cords (by controlling the position of the tongue in the pharyngeal and oral cavities), making the higher frequency parts of sounds independent of fundamental frequency (voice pitch). Birdsong, by contract, is generated primarily by directly controlling of the fundamental frequency produced by the syrinx. Nonetheless, in other respects, birdsong is much more like human speech than are the vocalizations of other animals, some of which can (like monkeys and male deer) modulate laryngeal harmonics (Owren et al., 1997; Fitch & Reby, 2001) in a human speech-like fashion.

There are intriguing clues about the evolution of fine vocal control from neuroanatomy of the song system. For example, motor output neurons in the forebrain (in nucleus RA) of songbirds have gained direct access to motoneurons controlling the syrinx vocalization musculature. Projections from RA also bypass the brainstem pattern generator circuitry for calls through which all forebrain outputs must pass in nonsong birds like ducks (Nottebohm et al., 1976; Arends & Dubbeldam, 1982) but also in squirrel monkeys (Ploog, 1981; Kirzinger & Jurgens, 1991) and macaque monkeys (Simonyan & Jurgens, 2003). There is a striking parallel here to the evolution of fine finger control in primates (but also finger control in raccoons, as a yet another reminder that evolution is a bush, not a linear Great Chain of Being), where motor cortex neurons have also come to contact finger motoneurons directly, bypassing pattern generators for coordinated limb movement in the spinal cord; hand motor cortex in cats, by contrast, contacts primarily the spinal pattern generators, which then have the only private access to motoneurons. The more direct access afforded the forebrain in the case of the songbird syrinx and the primate and raccoon hand presumably underlies more complex, differentiated, learned control observed in these systems. Note that this means that the relevant forebrain output areas have essentially come to assume a *lower* level in the motor control hierarchy, allowing the development of higher forebrain pattern generating centers that can operate somewhat independently of the brainstem and spinal pattern generating circuitry that is still needed for locomotion and nonsong vocalization.

Speech-like birdsong carries *less* meaning than vocal call systems do

The most striking characteristic of birdsong, however, in light of its prodigious complexity, is its essential lack of semantic content. Individual syllables or song fragments do not seem to have any specific meaning outside of being part of a particular song; and particular songs do not seem to convey specific content. Nor do birds appear to produce anything like 'words' by recombining their 'syllables' in order to signify concepts. Despite having vocal and auditory equipment ideally suited to support the recombinable speech symbol half of a language-like meaning conveying system, birdsong seems to communicate only very general meanings. Songs serve to mark territories, identify the singer's species, attract mates, often all at once. The messages communicated by birdsong are, in fact, *less* content filled than the messages communicated by, say, vervet monkey calls, which have been shown to signify rather elaborate distinctions among predators and conspecifics (Cheney & Seyfarth, 1990), despite the comparative simplicity of those unlearned monkey calls. This difference in referential content is particularly obvious when we consider the 'meaning' of a single syllable of a songbird's song; though emotion is keenly involved in motivating the bird to begin singing, the identity and order of syllables carry no specific emotional baggage.

One plausible theory about birdsong is that it was a product of runaway sexual selection, like the male peacock tail or outsize antlers in males, or huge sexual swellings in female baboons and chimpanzees. Elaborate singing abilities seem to have been preferred by mates, despite making little contribution to fitness beyond the fact that they were preferred. Sexual selection stands in contrast to natural selection, which rewards improved function like a stronger beak or more efficient wings. Certainly, a complex song could initially have been a sign of a mate fitter in other nonsong respects. But it is harder to explain the maintenance of extreme examples this way, especially when sexually selected features impede other functions (huge antlers) or attract predators (elaborate vocal displays). Sexual selection is not confined to female choice affecting male characters. And in some songbird species, both the male and female sing. Bay wren male-female monogamous pairs, for example, execute precisely coordinated unison “duets”, and the song control nuclei are large and hormone sensitive in females as well as males (Brenowitz & Arnold, 1985). The generally accepted explanation for this behavior is that the attractiveness of the male's song to listening females is reduced when a duetting female is singing along.

Several whale species have evolved a vocal learning system that resembles birdsong in a large number of respects, and provides a key additional example of how a speech-like vocal learning system can evolve without a ‘semantic urge’ (Tyack & Sayigh, 1997). Humpback whales learn to precisely reproduce long sequences of sounds and culturally transmit them to animals that are genetically unrelated. The main difference is that whale songs are lower in pitch, and individual songs unfold over several minutes instead of several seconds. The underwater acoustic environment of the ocean is quite reverberant, due to the faster and more efficient propagation of sound as well as the air-water boundary, which may be one reason for whale's more leisurely tempi. As with birdsong, whale song has social and sexual functions.

A birdsong-like ‘RNA world’ for pre-language

With the context provided above, we can see our way to a surprising extension of Darwin's language origins theory. On the evidence of the avian case, it seems possible that early hominids might have initially evolved an elaborate system of essentially phonetic vocalizations -- a kind of “talking song” with no component semantics -- as a result of sexual selection. In this view, a number of the specializations for auditory-vocal control evolved for entirely nonsemantic reasons. Perhaps early hominid pairs initially duetted like bay wrens, innocent of reference for a million years. Turning standard language origins scenarios on their heads, the preadapted ‘symbols’-without-meaning system might have only later been taken over for use as a semantic vehicle, that is, speech-sound ‘song’ before referential speech. An odd scenario, but then birdsong is quite odd itself.

This scenario contrasts with Fitch's idea that laryngeal descent in hominids (which occurs early in development) might have occurred as strategy to indicate large size in males, by analogy with the realtime laryngeal descent that occurs during calls made by rutting male deer (Fitch & Reby, 2001). A functionally similar sort of call, though using air sacs instead of laryngeal descent, is already well known in gibbons and orangs, and it closely resembles emotional meaning laden signals in standard animal call systems; and unlike birdsong, these primate calls develop even in deafened animals, indicating that learning is not required. The birdsong model suggests instead that there was runaway selection for complex *sequences* of essentially meaningless segments -- each untied from particular emotions -- as opposed to selection for a large sounding roar; it's the elaborateness of the sequences that the mates found attractive, not their throatiness.

As mentioned above, RNA serves both as a code (mRNA), but also as a noncode-like, folded word recognition device (tRNA) and chain assembly device (rRNA). By analogy, the internal representations of speech sound sequences that a primate neurobiologist would expect to find in the human lateral temporal cortex may have some other function besides merely serving as internal copies of the speech stream; these uninterpreted speech sound representations could also be involved in word recognition and assembly of primarily visual meaning units into coherent discourse structures. By this account, what distinguishes humans is the ability to use a sequence of symbol patterns from another modality to cause the assembly of meaning patterns in higher visual cortex. But the product of that assembly may be very similar to patterns assembled from direct visual inputs arriving from earlier visual areas during scene comprehension, which also involves rapid serial assembly (of successive glances). The implication is that the trick of language was not to have invented the basic meaningful units but to have found a way of making standardized connections between them (see Sereno, 1984; 1991b).

One longstanding problem in the origin of human language is its sudden appearance. Most commentators agree that modern-style human language is less than 100,000 years old and probably less than 50,000 years old, based on the appearance in the cultural record of modern appearing artifacts together with *Homo sapiens*. Set against this is the much longer time that it must have taken for the anatomical and neural structures that control human language production to have evolved from their primitive condition in all other anthropoid primates. The birdsong/RNA world picture presented above provides one way out of this problem. Perhaps there was a long period of a million years or more in which the neural and anatomical basis of language production evolved for essentially *nonsemantic* reasons in early *Homo* or even *Australopithecus* species. This would have set the stage for the emergence of a linguistic ‘RNA world’, where the word recognition and chain assembly properties of meaningless speech-sound representations could be discovered, and then eventually grafted on to a productive meaning construction system that began to use visual representations to do most of the work (Sereno, 1991a), the analogue of the mostly protein based world at the cellular level.

***Homo floresiensis* and the ebu gogo**

After this paper was finished, a new dwarf hominin species, *Homo floresiensis*, was discovered on the island of Flores in Liang Bua cave in eastern Indonesia (Brown et al., 2004). Dating by radiocarbon, luminescence, uranium series, and electron spin resonance indicates that the species existed from before 38,000 years ago until at least 18,000 years ago. It was associated with a large number of small stone tools (Morwood et al., 2004). Dwarfing and giantism is a common evolutionary response in species that are confined to islands, and this dwarf species of *Homo* coexisted on Flores island with dwarf *Stegodon* elephants (elephants are good long distance swimmers), giant rats, giant tortoises, Komodo dragons (Komodo is a small island just west of the much larger Flores island), and an even larger extinct varanid lizard. *Homo floresiensis* is thought to be most closely related to *Homo erectus* on the basis of the dentition (lightly built jaw containing small canines and small premolars and molars), the skeleton (indicating it was an obligate biped, though with somewhat long arms and fingers), and the extremely small brain case (slightly smaller than a chimpanzee). Stone tools on the island previously attributed to *Homo erectus* date back to 800,000 years ago. There is archaelogical evidence that anatomically modern *Homo sapiens* had already arrived in the area (East Timor, Australia) by 40,000 years ago, but no *Homo sapiens* fossils have yet been found on Flores. No specimens of the dwarfed *Stegodon* were found above a 12,000 year old tuffaceous deposit that resulted from a large eruption of the Flores volcano. Given the small number of hominin fossils found, however, it is less clear that they went extinct, too. Austronesian speaking immigrants arrived on the island about 2,000 years ago to find indigenous Melanesians (the current inhabitants speak Austronesian languages). There was contact with India and China; and then the Portuguese first arrived in 1520, and a century later, the Dutch.

A decade before the *Homo floresiensis* find, Gerd van den Bergh, a paleontologist working on the faunal remains who speaks Indonesian, had heard stories from villagers living in several different towns near the foot of the volcano about a race of hairy, three foot tall people, the “ebu gogo” (literally, ‘the grandmother who eats anything’). The ebu gogo were long-haired, potbellied cave dwellers with protruding ears, and long arms and fingers, and they walked with a slightly awkward gait and would climb small trees. The villagers said that the last ebu gogo was seen in the 19th century, when the Dutch settled in central Flores (Roberts, 2004). Although the folklore of many groups around the world mention small people (leprechauns in Ireland, menehune on Hawai'i), the ebu gogo stories are unique among them in matching several specific physical aspects of local subfossil remains.

In the context of the present paper, the most poignant aspect of these stories concerns the putative vocal abilities of the ebu gogo (and by implication, of *Homo erectus*!) that were observed by the villagers as they tolerated the ebu gogo raiding their crops, and during closer encounters when the villagers provisioned them with grains, vegetables, fruits, and meat, all of which the ebu gogo ate raw. The ebu gogo “murmured at each other and could repeat words verbatim” in a parrot-like fashion; for example, “in response to ‘here's some food’ [in Indonesian], they would respond ‘here's some food’” (Roberts, 2004). Although this evidence is incomplete and indirect, there is an uncanny fit to the scenario introduced above in which modern *Homo sapiens* style language emerges ‘at the last minute’ from an initial set of auditory and motor system modifications of much greater antiquity that had originally evolved to support *nonsemantic*, birdsong-like vocalizations. The remote but exciting possibility that the ebu gogo still exist might someday make it possible to test these ideas directly.

References

Arends, J.J. & Dubbeldam, J.L. (1982). Exteroreceptive and proprioceptive afferents of the trigeminal and facial motor nuclei in the mallard (Anas platyrhynchos L.). *Journal of Comparative Neurology, 209*, 313-329.

Baldwin, J.M. (1902). *Development and Evolution*. New York: Macmillan.

Ban, N., Nissen, P., Hansen, J., Moore, P.B. & Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science, 289*, 905-920.

Brenowitz, E.A. & Arnold, A.P. (1985). Lack of sexual dimorphism in steroid accumulation in vocal control brain regions of duetting song birds. *Brain Research 344,* 172-175.

Brown, P., Sutikna, T., Morwood, M.J., Soejono, R.P., Jatmiko, Saptomo, I.E. & Due, R.A (2004). A new small-bodied hominin from the Late Pleistocene of Flores, Indonesia. *Nature, 431*, 1055-1061.

Cheney, D.L. & Seyfarth, R.M. (1990). *How Monkeys See the World*. Chicago: University of Chicago Press.

Darwin, C. (1859). *On the Origin of Species.* Cambridge, MA: Harvard University Press.

Darwin, C. (1971). *The Descent of Man, and Selection in Relation to Sex* (2nd ed. Rev). New York: Appleton.

Doupe, A.J. & Kuhl, P.K. (1999) Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience, 22*, 567-631.

Deacon, T.W. (1997). *The Symbolic Species*. New York: Norton.

Fitch, W.T. & Reby, D. (2001). The descended larynx is not uniquely human. *Proceedings Royal Society of London B, 268*, 1669-1675

Gilbert, W. (1986). The RNA world. *Nature, 319*, 618.

Harnad, S.R., Steklis, H.D. & Lancaster, J. eds. (1976) *Origins and Evolution of Language and Speech*. Annals of the New York Academy of Sciences, v. 280: New York Academy of Sciences.

Hauser, M. (1996). *The Evolution of Communication*. Cambridge, MA: MIT Press.

Illangasekare, M., Kovalchuke, O. & Yarus, M. (1997). Essential structure of a self-aminoacylating RNA. *Journal of Molecular Biology 274*, 519-529.

Jablonski, N.G., & Aiello, L.C. eds. (1998). *The Origin and Diversification of Language.* San Francisco: University of California Press.

Joyce, G.F., Schwartz, A.W., Miller, S.L. & Orgel, L.E. (1987). The case for an ancestral genetic system involving simple analogues of the nucleotides. *Proceedings of the National Academy of Sciences 84*, 4398-4402.

Jenne, A. & Famulok, M. (1998). A novel ribozyme with ester transferase activity. *Chemical Biology, 5*, 23-24.

King, B.J., ed. (1999). T*he Origins of Language: What Nonhuman Primates Can Tell Us*. Santa Fe, N.M.: School of American Research Press.

Kirzinger, A. & Jurgens, U. (1991). Vocalization-correlated single-unit activity in the brain stem of the squirrel monkey. *Experimental Brain Research, 84*, 545-560.

Knight, C., Studdert-Kennedy, M. & Hurford, J.R. eds. (2000). *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form.* New York: Cambridge University Press.

Knight, C. (2002). Language and revolutionary consciousness. In A. Wray (ed.), *The Transition to Language* (pp. 138-160). Oxford: Oxford University Press.

Konishi, M. (1985). Birdsong: From behavior to neuron. *Annual Review of Neuroscience, 8*, 125-170.

Lewontin, R.C. (1970). The units of selection. *Annual Review of Ecology and Systematics, 1*, 1-18.

Merker, B. (2000). Gibbon songs and human music from an evolutionary perspective, In N.L. Wallin, B. Merker, and S. Brown, eds., *The origins of music* (pp. 103-124). MIT Press.

Merlin, D. (1991). *Origins of the Modern Mind*. Harvard University Press.

Miller, S.L. & Orgel, L.E. (1974). *The Origins of Life on Earth*. Englewood Cliffs, NJ: Prentice-Hall.

Morwood, M.J., Soejono, R.P., Roberts, R.G., Sutikna, T., Turney, C.S.M., Westaway, K.E., Rink, W.J., Zhao, J.-X., van den Bergh, G.D., Due, R.A., Hobbs, D.R., Moore, M.W., Bird, M.I. & Fifield, L.K. (2004). Archaeology and age of a new hominin from Flores in eastern Indonesia. *Nature*, *431*, 1087-1091.

Ploog, D. (1981). Neurobiology of primate audio-vocal behavior. *Brain Research, 228,* 35-61.

Nelson, D.A. & Marler, P. (1989). Categorical perception of a natural stimulus continuum: birdsong. *Science, 244*, 976-978.

Nissen, P., Hansen, J., Ban, N., Moore, P.B. & Steitz, T.A. (2000). The structural basis of ribosome activity in peptide bond synthesis. *Science, 289*, 920-930.

Nottebohm, F., Stokes, T.M. & Leonard, C.M. (1976). Central control of song in the canary, Serinus canarius. *Journal of Comparative Neurology, 165*, 457-486.

Owren, M.J., Seyfarth, R.M. & Cheney, D.L. (1997). The acoustic features of vowel-like grunt calls in chacma baboons (Papio cynocephalus ursinus): implications for production processes and functions. *Journal of the Acoustic Society of America, 101*, 2951-2963.

Roberts, R. (2004, October 28). Villagers speak of the small, hairy Ebu Gogo. *Daily Telegraph (UK).*

Schwartz, A.W. (1998). Origins of the RNA world. In A. Brack (ed.), *The Molecular Origins of Life* (pp. 237-254). Cambridge: Cambridge University Press.

Sereno, M.I. (1984). *'DNA' and Language: the Nature of the Symbolic-representational System in Cellular Protein Synthesis and Human Language Comprehension*. Unpublished doctoral dissertation, University of Chicago, Illinois.

Sereno, M.I. (1991a). Language and the primate brain. *Proceedings, Thirteenth Annual Conference of the Cognitive Science Society* (pp. 79-84). Lawrence Erlbaum Associates.

Sereno, M.I. (1991b). Four analogies between biological and cultural/linguistic evolution. *Journal of Theoretical Biology, 151*, 467-507.

Simonyan, K. & Jurgens, U. (2003). Efferent subcortical projections of the laryngeal motorcortex in the rhesus monkey. B*rain Research, 974*, 43-59.

Tyack, P.L. & Sayigh, L.S. (1997) Vocal learning in cetaceans. In C.T. Snowdon, M.

Hausberger et al. (eds.), *Social Influences on Vocal Development* (pp. 208-233). Cambridge, England: Cambridge University Press.

Wills, C. & Bada, J. (2000). *The spark of life: Darwin and the primeval soup*. Cambridge, MA: Perseus Publishing.

Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H.D. & Noller, H.F. (2001). Crystal structure of the ribosome at 5.5 A resolution. *Science, 292*, 868-299.

Zahavi, A. (1993). The fallacy of conventional signaling. *Philosophical Transactions of the Royal Society of London B, 340*, 227-230.

Zhang, B. & Cech, T.R. (1997). Peptide bond formation by in vitro selected ribozymes. *Nature, 390*, 96-100.