To appear in: D. Cummins and C. Allen (eds.) *The Evolution of Mind* Oxford University Press

The evolution of reference

Colin Allen and Eric Saidel

The unrefined, untutored mind
Of *Homo javanensis*
Could only treat of things concrete
And present to the senses.
(Quine, 1953, p. 77, fn. 5)

1. Birds do it, bees do it, and specially-educated chimpanzees do it. But how do nonhuman capacities for doing it compare to the seemingly infinite variety of human ways to do it? When the capacity in question is communication, there is little agreement on what we can learn about the human capacity for linguistic communication through studying animal communication. There is great skepticism among many linguists, psychologists, and philosophers about what the grunts, growls, whistles and howls of our furry and feathered relatives can tell us about human language, and there is even greater controversy about the significance of attempting to get members of other species to learn human languages.

2. Some scientists contend that human language is the output of a special brain module whose evolutionary development lies buried with our hominid ancestors. This, they believe, makes human language so utterly unlike any other form of animal communication that there is no point trying to compare what we do with what other animals do. But not everyone agrees that the origins of language are as recent as the hominid line, and there are many scientists who believe that much can be learned about human language by studying nonhuman communication both in the laboratory and in the field.

3. The authors of this chapter are philosophers interested in exploring the continuity between human minds and animal minds from a Darwinian perspective. The analysis of language is central to this project because it has often been claimed that the human capacity for language marks a difference in kind between human mentality and the mentality of nonhuman animals. Descartes, the founder of modern philosophy, thought that the human capacity for language was unexplainable in terms of the mechanical operations of a material body, and that this capacity could only be explained by the operations of an immaterial mind. The apparent inability of nonhuman animals to use language provided evidence for Descartes that they lacked minds (Descartes 1637). Even contemporary philosophers who do not believe in Descartes' distinction between material bodies and immaterial minds assert that language is the basis of a unique mentality (Dennett 1995).

4. To assess these claims of radical differences it is important to know who is right: is language a uniquely hominid innovation, or does it have a much older phylogenetic history? Our answer is neither a simple "yes" nor a simple "no". The human capacity for language is a complex mosaic of traits, each of which can be investigated independently. Such investigations provide insight into the ontogeny and phylogeny of language (Locke 1995 makes a similar point). In this chapter we will focus on one of those traits -- the capacity for symbolic reference -- and show what insights are available by taking a fully comparative, evolutionary approach.

Three views of language

5. The "language-is-recent" perspective is recently exemplified by psycholinguist Steven Pinker (1994) who has done a fine job of making the case that human language is an evolutionary adaptation. He repeatedly emphasizes the differences between human language and other forms of animal communication as part of his general argument that language is a recent, hominid adaptation. He lampoons attempts to teach human languages to members of other species with an analogy: trying to teach languages to nonhumans is, he thinks, like trying to teach the nearest living relatives of elephants---the shrew-like hyraxes---to pick things up with their short, unremarkable snouts (pp. 332-333). This parody could easily be extended to naturalistic studies of animal communication: ethologists who study animal communication in the field are like camouflaged researchers stalking wild hyraxes with field glasses and video recorders hoping to catch them using their snouts in a prehensile way.

6. Many comparative psychologists disagree with the language-is-recent view that restricts its evolution to the hominid line. They see human language as built on top of more general cognitive and communicative abilities with a much longer evolutionary history. Their hunch is that much can be learned about the evolution of human language by studying both the natural communication capacities of nonhuman species and their ability to use artificial languages. This perspective is exemplified by Patricia Greenfield (1991). Greenfield argues that language and tool use share an underlying cognitive basis in the capacity for completing object manipulation tasks that are hierarchically structured, for instance tasks that involve the completion of subassemblies for combination into larger objects. Following Chomsky (1957), linguists agree that the syntax of human languages is best understood in terms of hierarchical relations between subassemblies (such as noun phrases and verb phrases) that can be combined into higher-level structures and a potentially infinite number of grammatical sentences. Greenfield suggests that the capacity for a hierarchically structured grammar has its origins in primate abilities to engage in structured manipulation of tools. She argues that the cognitive capacity for tasks with hierarchical structure has a neural basis that evolved well before the hominids and the pongids went their separate ways. Many of Greenfield's conclusions are based on laboratory studies that directly compare the cognitive skills and development of humans and other primates, particularly chimpanzees. These conclusions are controversial, but for our purposes it suffices that they represent a serious attempt to understand the phylogenetic origin of an important human linguistic capacity---the capacity to parse hierarchical syntactic structures---in terms of more general hierarchical capacities.

7. A third perspective is provided by biologists who are concerned with the evolution of behavior--the subdiscipline usually called "ethology". From an ethological perspective, human language is an adaptation to a specific regimen of natural selection but it is nonetheless evolutionarily related to other forms of animal communication. Human language skills may share common origins with other forms of animal communication (human language and animal communication may be homologous) or there may been convergent evolution for similar functions (they may be examples of homoplasy or analogy). Although attempts to teach human languages to members of other species may provide interesting suggestions about the evolution of human language, ethologists prefer to compare the natural communication systems of different species. This perspective is exemplified by attempts to show that the communicative signals of the members of various species are what ethologists call "referential" signals. Cheney and Seyfarth (1990) use data collected from their field research on vervet monkeys to argue that some vervet vocalizations convey information about specific predators such as leopards, snakes, and eagles. Peter Marler and Christopher Evans have pressed a similar point about the alarm vocalizations of chickens that are different for aerial and terrestrial predators (Evans and Marler 1995; Marler and Evans 1995). These and other ethologists are interested in the evolution of reference.

8. This third perspective is a response to an early but still common view of animal communication that takes signals to be involuntarily caused by immediate stimuli (things concrete and present to the senses, to borrow the line from Quine). According to this view, a monkey's alarm call (or a

chicken's) is simply an emotion-caused response to the appearance of a predator---a direct consequence, for example, of the fear that a predator causes in the monkey. Proponents of this "emotion" account of communication are not pressing any strong claims about the consciousness of the animal. Rather, they claim that it is not necessary to attribute complex information processing to explain nonhuman animal communication. They suppose that relatively simple mechanisms can explain the range of behaviors observed in instances of nonhuman animal communication, whereas, they argue, human communicative skills require much more complicated cognitive skills. Prominent in these arguments are claims about the amount of cortical involvement in processing signals and the degree of voluntary control that organisms have over their signalling behavior. Those who wish to emphasize the uniqueness and recency of human language evolution embrace the view that animal communication is involuntary, perhaps under the control of the limbic system (the subcortical parts of the brain that are also responsible for emotions) while human communication is highly voluntary and relatively detached from emotional responses (Pinker 1994, p. 334; see also Lieberman 1995 for discussion).

9. Contemporary ethological research suggests, however, that animals have a much higher degree of voluntary control over their vocalizations than was originally supposed (Marler et al 1991; Macedonia and Evans 1993; Hauser and Marler 1993a,b; Marler and Evans 1995). Additionally, neurological studies of nonhuman primates provide evidence that both comprehension and production of their vocalizations involve higher cortical parts of the brain, lending further credence to the view that there are closer parallels between human language and the vocalizations of other animals than the received view would suggest. For example, MacNeilage et al. (1991) present evidence that babbling in young human children is controlled by a part of the motor cortex that also controls the vocalizations of some primates. Hauser and Andersson (1994) established that adult rhesus macaques (but not infants) show left hemispheric dominance for processing the vocalizations of their conspecifics and Rauschecker et al. (1995) have shown that rhesus macaques possess cortical neurons that are selectively responsive to the vocalizations of conspecifics. These and other parallels can be used to identify general evolutionary principles that lead to the development of referential communication systems.

10. The relationships between human language and other forms of animal communication are undoubtedly complex. This suggests that it is appropriate to adopt an ethological perspective on language, exemplified by Evans and Marler (1995, p. 341) when they write:

    - Language is dependent upon such a complex package of cognitive and anatomical features that at first it seems to be the exception to the otherwise clear pattern of continuity apparent between humans and other vertebrates, particularly the higher primates. It is, however, possible to isolate the different functional attributes of language ... and then to search for these traits in the cognitive and communicative abilities of other animals. This approach takes advantage of one of the most important legacies of the early ethologists, who demonstrated that comparative studies allow us to discriminate between attributes that are phylogenetically ancient and those that have evolved much more recently ...

    That our sympathies lie with this approach will be clear from the remainder of this chapter.

What is a trait?

11. Any attempt to investigate relationships between the abilities of humans and those of other animals is faced with a challenging question: When are the traits of different species instances of the same (type) trait? In some cases this question can be answered in a reasonably straightforward way on functional grounds. The upper appendages of birds, bats, and dragonflies are all considered wings because they play roughly the same role in enabling these organisms to fly. In this case a trait (the possession of wings) is identified on functional grounds

while questions of homoplasy versus homology are decided on grounds that indicate separate evolutionary development, such as the considerable anatomical differences between the wings of birds and the wings of insects.

12. When traits have easily recognized functions their reidentification across species is a relatively straightforward matter. Questions of homoplasy vs. homology can also be settled in a relatively straightforward manner if there are obvious anatomical structures associated with the functions. But language provides a special difficulty for the comparative method because the capacity for language has very many functions whose dependencies on each other are not clear. Naively one might ask: Is the capacity of humans to communicate via language the same (type) trait as that of vervets to communicate via alarm calls? But this question is not as straightforward as it seems; before answering it one must determine which of the multiple functions of human and animal communication systems fix the identity of the trait. The determination of functions can perhaps be placed on an objective footing (see Allen and Bekoff 1995) but the choice of functions used to determine the identity of a trait seems to be a largely pragmatic decision -- it is interest-relative. Comparisons between human language and vervet communication are further clouded by the fact that while there is much overlap in function, there is also divergence at many points. To be sure, humans can do things with their languages that other animals cannot do with their communication systems. But so too can hummingbirds do things with their wings that eagles cannot do with theirs, and they share a common trait despite these differences.

13. Those concerned with defending the uniqueness of human language often deny the equivalency of the traits studied in other animals. With respect to ape-language research this has led to lengthy battles over the definition of "language". But it is our view that such disputes are off the point. It's not so much a matter of what language is, but what language does that is important for its evolutionary analysis. And because there are many things that language does, and many features of language that support these functions, it is a mistake to dismiss the relevance of animal studies wholesale. Different features and functions of language can be investigated from an evolutionary perspective, and there may be different answers about the usefulness of comparing each feature to the capacities of nonhuman animals. Some features of human language may turn out to be homologous with animal communication while others may turn out to be homoplasies, and still other features may turn out to be entirely unique to human language. But this latter category does not entail that there is no sense in comparing them; if the wings of hummingbirds have unique features it does not follow that it is fruitless to compare hummingbird wings to the wings of other birds or to the wings of insects.

14. Participants on both sides of this discussion recognize that disputes about the definition of language are usually pointless. Pinker, for example, declares that the debate "over what qualifies as True Language" is "fruitless and boring" (p 347). It is worth noting, however, that this follows an earlier statement that "Genuine language ... is seated in the cerebral cortex, primarily the left perisylvian region" and is thus distinguished from primate vocalizations that are (he alleges) involuntarily controlled by subcortical structures (p 334). If the debate about "True Language" really is fruitless and boring then it's really just as pointless to hold up control by the cerebral cortex as a criterion of "genuine" language. These facts are, of course, relevant to questions about homoplasy vs. homology. But the more interesting topic is the extent to which various functions of language can be teased apart, and then shown to exist in different phylogenetic groups, not whether humans are categorically different from other animals (see also Hauser 1996).

15. Those who take a hard line about the relationship between human language and nonhuman animal communication also suggest that the study of animal communication is of no interest for understanding the evolution of human language unless the traits are homologous; mere homoplasy -- parallel or convergent evolution -- is deemed irrelevant. But this is too hasty. It is

true that homologies provide more direct evidence about the actual historical trajectory along which current traits developed, but it is false that homoplasies provide no relevant information. Ethologists are interested in general principles of evolution and selection. Homoplasy between the traits of different organisms results from similar selective pressures. Thus the study of homoplasies can reveal general principles about the effects of selection on organisms. A better understanding of those general principles can be used to inform the construction of hypotheses about specific historical trajectories. Comparative studies of vocal development in birds and humans provide a wealth of examples of how fruitful this approach can be (Hauser 1996).

Reference and Language

16. In trying to defend the view that reference is a basic functional property of language we note that many words in human language refer to actions or objects that are external to the speaker. The 20th century philosopher Ludwig Wittgenstein famously argued that philosophers (including his own younger self) mistakenly tend to think of reference as the sole function of words in a language, a view of language that he associates at the beginning of his *Philosophical Investigations* (1953) with St. Augustine's theory of language acquisition. While we agree that it is a mistake to forget that words do help to fulfill other functions, we want to suggest that the ability to refer is basic, and that in a very interesting sense all the more sophisticated functions of language are dependent on this ability. Because of the basic role of reference one might also suspect that the capacity to refer to external objects or events is phylogenetically older than other features of language, and that it might therefore be a property of other forms of animal communication. This, in turn, suggests that it may be fruitful to study the role of reference in other forms of animal communication in order to understand the evolution of human language.

17. Pinker, indeed, grants that the apparently referential vervet alarm calls are as good a place as any to begin thinking about the evolution of human language. He writes (p. 352): "Perhaps a set of quasi-referential calls such as these came under the voluntary control of the cerebral cortex, and came to be produced in combination for complicated events; the ability to analyze combinations of calls was then applied to the parts of each call." Underlying Pinker's description of the vervet calls as "quasi-referential" seems to be the now familiar concern with the alleged lack of voluntary control coupled with the fact that vervet communication lacks a structured, combinatorial syntax. Although the combination of voluntary control and complex syntactic abilities plausibly allows organisms to exploit a wider range of communicative abilities for various biological functions, the issues of signal reference, volitional control of signal production, and combinatorial syntax can be treated independently. Lumping different features together and considering them all essential to language reinforces the view that human language is unique. After all, nothing else has *precisely* this combination of features. In a similar vein, absolutely any trait could be deemed unique to a particular species. But appeals to such a weak standard of uniqueness may obscure evolutionary continuity for the various features considered independently.

18. The hardliner may continue to deny that the communication systems of other species share interesting features with human language. But if features exemplified by the vervets are conceded to be (at least) homoplasies for similar features of human languages, then this weakens the case for the uniqueness of human language. To move towards treating those features as homologies would be to deny the independent evolution of human language. Using Pinker's analogy of the elephant's trunk, the corresponding issue here is whether the common ancestor of elephant and hyrax passed on any traits that predisposed its descendants towards the evolutionary development of a trunk given selective pressures favoring a prehensile snout. If not then the trunk can be viewed as a trait that evolved independently in the elephant branch of the evolutionary tree. Similarly, at issue for language is whether the prehominid ancestors of humans possessed and transmitted traits that predisposed their descendants towards the

evolution of human language, given selective pressures favoring such a development (either directly or indirectly). To settle this question it is necessary to do comparative work involving humans, apes, monkeys, and non-primates.

19. Hardliners about the claim that language is a uniquely human trait may seek to bolster their case by pointing out that the apparent reference of animal signals is assessed by criteria that do not apply to the referential uses of words in human languages. In particular, biologists seeking to establish that the signals of nonhuman animals are referential often apply a criterion that Quine's fictional *Homo javanensis* would satisfy but that modern human languages do not satisfy. For example, Macedonia and Evans write (1993, p.179): "The `production' criterion is that referential signals should exhibit a degree of stimulus specificity. This requires that all eliciting stimuli must belong to a common category....One clear correlate of the `production specificity criterion' is that referential signals should not occur at appreciable rates in inappropriate contexts. We would not expect them to be produced in the absence of the putative referent." But as a matter of fact, in cases of modern human language one might very well expect that signals are used more often in the absence of their referents than in their presence. One of the advanced functions of a language is to allow us to talk about people and things in their absence. Without this capacity, gossip would be next to impossible! So, are biologists wrong to place such a restriction on reference? Or are they talking about a different phenomenon entirely? We think that the answers to these questions are both negative. It is the same phenomenon, but the cases of reference that satisfy the *javanensian* "production specificity criterion" are (as Quine versified) more primitive in evolutionary development, and it is perfectly acceptable for biologists to seek evidence for these more primitive capacities.

20. That signals are often present in the absence of their putative referents is evidence either that those signals do not refer to the putative referent, or that the organism using those signals has advanced beyond *javanensian* reference. It would be a mistake to conclude that because the signal "airplane" is most commonly used in the absence of direct stimulation by airplanes (and therefore fails the *javanensian* production specificity criterion) that it does not refer to airplanes. It is important when applying this criterion to recognize the fact that the absence of *javanensian* reference alone does not indicate that there is no reference. One must consider the evidence for both of these rival hypotheses: that reference has advanced beyond the here and now, or that reference is absent. One place to look for evidence of the former is in the development of reference in the young of the species. A suggestion that might (if treated with caution) guide this research is that ontogeny recapitulates phylogeny.

21. This old saw, that ontogeny recapitulates phylogeny, is at best a rough heuristic and at worst thoroughly misleading (Deacon 1991). Nonetheless, in this case there are some interesting comparisons to be drawn between human language acquisition, and the role of *javanensian* reference in the evolution of reference. A standard philosophical view of language acquisition is traced in the very first paragraphs of Wittgenstein's *Philosophical Investigations*. The "Augustinian" picture is suggested by St. Augustine's claim that he learned his language from adults who pointed to things and named them. The process of ostensive definition is caricatured by Wittgenstein in his example of "the block world". In this world a teacher points to stones cut into different shapes and utters a word, allowing a pupil to learn the correlation between the words and the stones. The blockworld language has three key features: (1) it contains referring terms (loosely called "names"); (2) it has no syntax; (3) it is acquired by a process of ostensive definition. Wittgenstein complains that the Augustinian picture of language stresses the importance of naming and plays down other aspects of language. He points out that even in such a restricted language there is much more to language than naming; there are also, for example, complex relationships between linguistic expression and action, such as fetching or carving a particular shaped stone. Different "language games" may even involve varying responses to the same expressions. By stressing naming, the Augustinian picture gives an incomplete and therefore misleading view of language.

22. Wittgenstein's premise --- that there is more to language than just the names of objects --- is correct. However, his ensuing de-emphasis of the naming function of language may be as misleading as the view it is intended to replace. While Augustinian ostensive definition may not apply to all of language, it may still provide the best account of a foundational stage of language acquisition. So, while there may be more to learning a language than learning the names for various referents, learning those names may be a necessary stage in the process of learning a language.

23. Learning to associate names and referents is necessary for organisms who are not born with complete knowledge of their vocabulary and its use. Such organisms must learn to correlate the referring terms with those events or objects beyond the speaker to which they refer. Without knowing that "dog" refers to dogs, the speaker of English is at a loss to understand sentences and utterances which use the word "dog". But how central is ostensive definition for the process by which a competent language user comes to learn the relationship between terms and referents? What role in language learning does ostensive definition play? Is the naming of objects a necessary ontogenetic precursor to other language functions? And, if it is, we're led to wonder if ontogeny does recapitulate phylogeny. Is what we find in many nonhuman species a primitive blockworld language? If so, was something like this blockworld language a step in the evolution of human language? Did language evolve by first being something akin to a primitive blockworld language that we might find in nonhuman species?

24. Does the previously-mentioned communication system of vervet monkeys lend credence to the suggestion that animals might use something resembling a blockworld language? Cheney and Seyfarth (1986) report that infant vervets learn to refine the application of their calls to correspond to the call use of adults in the group. Initially an infant will make the bird of prey alarm whenever something is moving in the sky. However, adults in the group ignore the call in most cases where the object in the sky is not a predator. If it is a predator, the call is typically repeated by an adult. It seems likely that nothing like explicit pedagogy is occurring between adults and infants, and Cheney and Seyfarth do not claim, on the basis of their data, that there is a connection between adults' repetitions of calls and infants' learning when to use them. Caro and Hauser (1992) do, however, present evidence that adult repetition is necessary for the infants to acquire the same patterns of use as the adults. Despite the lack of explicit instruction, we believe that there are some important similarities to the blockworld language game. The reinforcement provided (perhaps unintentionally) by adult vervets depends on infants and adults sharing attention to the same environmental features. The limited vocabulary of the vervets is similar to the limited vocabulary of the block world. And although the extent to which the vervets could be said to have a language is clearly very limited, nonetheless the referential function of the vervets' communicative utterances appears to be shaped by an ostensive process, albeit one that does not involve any overt pointing. But how is this fact about the acquisition of the vervet communication system related to the acquisition of human language? Is their apparently Augustinian acquisition process merely vaguely similar to human language acquisition? Is it an homologous process? Or is it homoplastic? How might advocates of the three views discussed above answer these questions? (The questions assume, of course, that there is a shared function of vervet "language" and human language. The evidence supports this assumption: both have the function to refer to objects and events external to the "speaker".)

25. Those who propose the hardline view that human language is a unique and recent adaptation are likely to argue that there are still considerable differences between the (possible) conditioning of young vervets by adult reinforcement and the kind of ostensive definition that Augustine described involving active pointing and naming of objects. But pointing is a gesture that is highly species-specific to humans. Its function is to draw attention to some feature of the environment. In the case of the vervets there is no need for adults to draw the attention of an infant to the cause of the infant's vocalization because they already share attention to the relevant environmental condition. We return to the issue of shared attention further below. But for now

our point is that the absence of pointing does not preclude other mechanisms for establishing shared attention. Once there is shared attention, then adult vocalizations bear the same relationship to the attentional object or event in both humans and vervets.

26. Comparative psychologists and cognitive ethologists are likely to be more comfortable with the idea of drawing inferences about phylogeny from ontogeny. Such inferences must be treated with extreme caution, but if the language learning of humans passes through an Augustinian phase and a similar mode of learning is found in animal communication systems, then the idea that this mode of learning is phylogenetically quite old demands further investigation. Comparable to Greenfield's view that language development is evolutionarily related to the development of the ability to reason hierarchically about objects one might also view both *javanensian* reference and learning from ostension as necessary precursors to more sophisticated linguistic abilities. The cognitive ethologist may also argue that the development of these abilities in other species provides important evidence about the possible precursors of the more varied language abilities found in hominids. Not only is there reason to think that the vervets (and other organisms) possess a communication system that shares features with the primitive blockworld language---a language that contains referring terms and is learned ostensively, but lacks syntactic structure---but there is also reason to think that these capacities are advantageous. For example, all other things being equal, the organism who can warn her offspring of an approaching predator stands a much better chance of passing her genes on to future generations than does the organism who lacks even the capacity for *javanensian* reference to predators.

Three Kinds of Reference

27. The questions we have asked this far have had to do with the relation between the referential properties of animal communicative systems and reference by human language. But we could approach this issue from another direction. We might ask about the attitude reflected in the ditty which begins this paper: Can organisms whose communication skills match those of the speaker of a blockworld language (vervets, for example, or perhaps, very young humans) manage to refer to objects that are not present? Or is the user of a blockworld language doomed to refer only to objects that are present to the senses? Under what conditions can reference extend to objects that are absent? To investigate these questions we consider three varieties of reference.

28. The first kind of reference to consider we shall call "mimetic reference"; this occurs when a signal closely resembles the referent so that it is capable of directly causing the same kind of response. As Dawkins and Krebs (1978) point out, a substantial part of signalling in nonhuman animals is the evolutionarily designed attempt by one animal to use another animal's muscle power to achieve the ends of the first. So an angler fish that attracts prey by means of a lure has certain behavior patterns and looks a certain way so that it may conserve its own energy while using the muscles of its prey to do work for the angler fish. This is an example of mimetic reference. And, in a nontrivial sense, the angler fish's lure makes reference to something (food for the prey) that is not present. Likewise, the broken-wing distraction display of plovers described by Carolyn Ristau (1991) constitutes a reference to an injured animal, and in doing so it provides an attractive alternative prey for predators near the plover's nest. Both the broken-wing display and the angler fish lure work by providing a stimulus that is not discriminated by the intended audience from a stimulus that would accompany the presence of a suitable prey item. This works because there is a non-arbitrary relationship between signal and referent. There may, however, be a considerable difference in the degree of sophistication with respect to voluntary control. Ristau's results suggest that the plover's deployment of its broken wing strategy depends on a complex assessment of intruder intention that is partly based on previous experience with individual predators. The point that interests us here, however, is that mimetic reference is possible in the absence of complex intentionality and in the absence of a blockworld language

and it may in some species lack the sophistication of even involuntary, emotional communication systems. Although this kind of reference is, of course, available to possessors of a blockworld language, it seems appropriate to treat it as an independent trait.

29. The second kind of reference we shall call "proxy reference"; this occurs when signals function as proxies for their referents in the sense that the signal elicits the same kind of response that the referent would but does so by a different cognitive mechanism. This is in contrast to mimetic reference which works by stimulating the same sensory/cognitive pathways as the referent would. Proxy reference represents a step towards the more arbitrary relationship between words and their referents that is common to all human languages. Many species of birds and mammals have such signals but much more work is necessary to find out the extent to which the arbitrary relationships between signal and referent are innate or learned. Proxy reference may also be exploited to refer to absent or non-existent referents to provide a benefit for the signaller. For example, in mixed-flock species of birds, Munn (1986) observed that members of a sentinel species would sometimes emit an alarm call when in direct competition with birds of other species for a particular prey item. In many cases there was no predator present, and when the other birds took anti-predatory action, the individuals who gave the spurious alarm calls were able to catch the insects for which they were competing. (This behavior seems to persist because the costs of ignoring a threat far outweigh the benefits of eating a bug.) There is also anecdotal evidence of similar behavior by monkeys during intergroup conflicts (Dennett 1983). All that is necessary for successful proxy reference is the ability to make the link between a signal and the object or behavior that signal refers to. Proxy reference is thus within the capability of organisms limited to a blockworld language.

30. The third sort of reference is what we call "conceptual reference"; this occurs when signals may refer to external conditions without it being normal for such uses to elicit the responses that the referent itself would elicit if it were present. This kind of reference is the norm for everyday human conversation. When someone says the word "tiger", the listener is not expected to get scared, begin evasive maneuvers, or to look for cat food. Because of the independence of conceptual reference from behavioral response, conceptual reference enables a whole new range of references to things that are not physically present. Bob is able to tell Jane about the spectacular sunset he saw last week and thereby make her think that he is a romantic. Whereas, if his description simply evoked her typical responses to sunsets, the point about his character might be lost (although bonding could still occur). Is conceptual reference possible for an organism that is limited to a blockworld language? There seems to be no principled reason the Augustinian organism would be unable to exploit conceptual reference. Indeed one can imagine a Wittgenstinian language game where the utterance of "slab" would invoke different responses according to whether or not it is uttered in the presence of a slab. In this case, the word is not simply a proxy for the object. Whether actual Augustinian organisms do show evidence of conceptual reference is another question---one that we believe could be profitably investigated by ethologists and comparative psychologists. The demonstration of conceptual reference would not amount to a demonstration of more than a blockworld language, and the absence of conceptual reference would raise interesting questions and suggest interesting research into why that capacity is absent.

31. Because both proxy and conceptual reference require the organism to make an arbitrary connection between signal and referent, and because it is often ecologically inadvisable to hardwire such connections, both these kinds of reference are likely to involve the kind of learning from ostension that concerned Wittgenstein. In order to learn the name of something ostensively, the referent of the name must be clear to the student. This ostensive definition requires at minimum some shared attention---both participants must be attending to the same condition in the environment. If the infant vervet and the adult are attending to different conditions when the adult repeats the infant's warning cry, the infant will fail to learn what are the correct referents of the warning cry. Indeed Caro & Hauser (1992) observed a case of where

an infant vervet saw a herd of stampeding elephants, gave a leopard alarm call, and then immediately after, the alpha male vervet saw a leopard and gave a leopard alarm call. The infant persisted giving leopard alarm calls to elephants for several months. One interesting issue for future study is the evolution and development of shared attention. How much does the infant know innately about what to pay attention to, and how much is learned? And how is the ontogeny of shared attention related to the ontogeny of language? We turn to these questions below. Before doing so we want to ask what the shared attention is attention to.

Attention and Reference

32. Our tendency as language-using adults is to think of reference stereotypically as reference to objects; but we shall argue that both phylogenetically and ontogenetically, reference to behaviors may be more basic than reference to objects. Consider, first, human infants. Studies of the attention of human infants show that they are more likely to attend to lights moving in an animated pattern (as if they are positioned at the joints of a moving animal, human or nonhuman) than they are to attend to lights that are either static or moving linearly. (Johannson 1973). This suggests that the attention of human infants is drawn naturally to objects engaged in certain kinds of behavior rather than to objects themselves. Furthermore, this suggests that the motion that human infants find most captivating is animate motion, motion, that is, that might be of something animate in the infant's environment rather than the motion of something inanimate (such as a tree branch blowing in the wind). A natural predisposition to attend to animate motion more readily than inanimate motion or to static objects would provide evolutionary advantages under certain ecological conditions. Some of those things that are moving are predators. The organism that is naturally predisposed to attend to animate motion is the organism that is going to monitor a predator more closely. Does this predisposition extend to non-human primates? Do vervet infants prefer to attend to lights moving in an animate pattern over lights that are static or lights that are moving in a linear pattern? We don't know of any research specific to vervets but Johansson's methods applied to cats showed that they can discriminate animate motion from non-animate motion (Blake 1993) and Hauser and Carey (this volume) discuss their own experiments with monkeys (tamarins) that seem to differentiate between animate motion and other forms of motion.

33. There is some evidence from field studies that non-human primates are similarly concerned with motion. Part of the vocal behavioral repertory of the vervet is the moving-into-the-open (MIO) grunt. This is used by a vervet to indicate that it is about to move into an open area, or that it is following another into an open area. The grunts do not indicate the open areas themselves, for they're not performed in most contexts when the vervets are simply near or in an open area, but are reserved for contexts involving motion into the open areas. The referential properties of the grunts would be a mystery if the vervets were limited to referring to objects. There is no one object (or type of object) that is the referent of an MIO grunt, instead there is a change in the spatial relationship between the open area and various vervets. What moves into the open might be the animal producing the grunt, or it might be another vervet. The MIO grunt captures the changing relationship between the two. Because this is one signal that refers to different objects in different situations in which the common element is not just how the objects are related, but also how that relationship is changing, the vervet could not successfully make an MIO grunt with a blockworld language that refers only to objects, and not behaviors. If we hypothesize instead that their language takes behaviors or movements as its primitives, then we could have an atmosphere which facilitates understanding the MIO grunt, and one which doesn't demand any language of greater complexity than blockworld language. (While this result may please those hardliners who prefer to distance human linguistic abilities from those of nonhuman animals, we suggest below that the support this provides for the hardliner about human uniqueness is less than may seem at first blush; remember, evidence also suggests that human infants attend to events rather than objects.)

34. Other ethological evidence supports the conclusion that the vervet communication system takes events rather than objects as a primitive. Careful analysis of the ontogeny of vervet alarm calling shows that the infants don't make mistakes about everything in the sky (when making bird of prey alarm calls) or everything in the trees (when making beast of prey alarm calls); instead analysis of the ontogeny of eagle alarm calls shows that infants' "mistakes" (when making bird of prey alarm calls) are most common for nonpredatory species diving rapidly from the sky or closely approaching the vervets, and that such errors are not associated merely with morphological similarity (Seyfarth, Cheney, and Marler 1980; Seyfarth and Cheney 1986; Cheney and Seyfarth 1986; Hauser 1989). Because these are behaviors that may reasonably be associated with predation, and because moving objects are more easily discriminated from background than static objects, it makes sense that vervets would be innately disposed to react to such events. Again, it is the nature of the events that is noteworthy from the point of view of the vervet, not the nature of the object performing the behavior. In order that the infant vervet learn which objects are the correct referents of the alarm calls, it is apparently predisposed to pay attention to objects behaving in a certain manner. This would also explain how it could happen that an infant vervet would learn to make the leopard alarm call to elephants (as observed by Caro & Hauser 1992): what it was predisposed to notice was a kind of behavior, not a kind of morphology. This is also consistent with discovery by Evans & Marler (1995) that a chicken shown a moving image of a raccoon on a video monitor mounted overhead will make aerial predator calls at a higher rate than terrestrial predator calls. This is not to say that morphology is irrelevant: alarm calls were more reliably elicited by video footage of a raptor on the overhead monitor than by footage of the raccoon. Similarly, realistic, artificially-generated raptor-shaped images were significantly more effective in eliciting alarm calls than disk-like images with the same surface area moving at the same velocity. (Evans & Marler don't tell us but it would also be interesting to know how the animals would respond to video of raptors in the eye-level monitor.)

35. The evidence from chickens indicates that event type of the stimulus is a more important determinant of call type than is object type. (Remember that the raccoon and the disk moving overhead both produce aerial predator alarms, albeit at a lower rate than hawks or artificial hawk shapes.) The evolution of reference to events of these categories as distinct may have been driven by the different anti-predatory strategies that are appropriate for chickens faced by these different predators (Macedonia & Evans 1993). Similarly, infant vervets seem to begin with an action-oriented classification scheme. In vervets, reinforcement of infant vocalizations may lead to a classification scheme that is based more on perceptual characteristics abstracted away from behavior. The ability to categorize and refer to objects independently of behavior would be an adaptive trait when the costs of responding to false positives (such as non-predators behaving in a predatory fashion) or of failing to respond to false negatives (such as predators behaving in non-typical ways) are relatively high.

36. If both human infants and nonhuman primate infants naturally (perhaps even innately) attend to behaviors rather than to objects, this may be a result of a common cognitive attentional ability. We see two possibilities here. Either this trait (the trait of attending to behaviors) is found in both human and nonhuman primate infants because it is an homoplastic development, or because it is an homologous development. The first is certainly possible: as suggested above, there is clear advantage to attending to behaviors rather than to objects. But the second is equally possible: this development may well have first come about in some ancestor common to both human and nonhuman primates. What is needed to answer this question is some comparative study of these cognitive abilities in both humans and other primates. Such a study would attempt to map the co-occurrence and dissociations between the different types of referential abilities described above. This makes possible inferences about the phylogenetic relationships between these traits.

37. To see how this applies to the particular case of human language, consider the importance of

infant attention for the development of language. Without shared attention between infant and adult, learning from ostension is impossible (in fact, this was Wittgenstein's point in objecting to a blockworld language). In order to learn some basic vocabulary, both parties to the interaction must each be attending to the same thing. Because attention to motion is a very basic necessity for a wide range of organisms (but not all), it is not surprising to find that mechanisms that depend on *shared* attention should be built on top of the capacity for attending to motion. If the infant's attention is naturally drawn to behaviors, an adult will be able to draw upon that in her ostensive definition. It seems likely, then, that studying the cognitive mechanisms of attention is important task for understanding the development of language. Supposing that language depends on shared attention, and that mechanisms for establishing shared attention are found in many nonhuman organisms, if it is further found that these mechanisms are especially well developed in primates, then language can be seen as built on top of an ability that we share with other primates. Even if the shared attention mechanisms of primates are merely homoplastic and not homologous to the human trait, the comparative studies may nonetheless establish that the capacity for shared attention is a necessary precursor to learning relationships between signals and referents. Thus, it is too hasty of hardliners about uniqueness to dismiss the relevance of animal studies on the grounds that the communicative capacities of animals are mere homoplasies of human language. And while it is possible that human attentional mechanisms are only analogous to those found in other primates, this conclusion should be supported by comparative studies of humans and other primates, not by an a priori conviction about the nature and origins of human language.

Concluding Remarks

38. We've been pursuing two apparently distinct lines of inquiry in our discussion of the evolution of reference, one about reference to absent objects, and another about the primitive objects of reference. Now it is time to draw these lines together. When an adult vervet makes an alarm call the other vervets respond to that call as if they were responding directly to the predator (their behavior differs, of course, with the kind of alarm call). This is what we have called "proxy reference". Some vervets have apparently learned to exploit proxy reference to their advantage by relying on alarm calls to cause the same reaction in other vervets that a predator would cause. The same is true for members of sentinel species of birds: they have learned to use their alarm calls to their advantage when there is no predator. This kind of behavior can only be advantageous if the consumer of the alarm call treats it as a sign caused by the presence of the predator, not as a false signal of some absent object. Compare this use of an alarm call to shouting 'fire' in a theater in order to get a seat. The false 'fire' cry only works if the consumers of this cry treat it as a reference to something present, rather than a reference to something absent. The successful use of false alarms depends on the audience treating them as signals of a present, not absent object. The sentinel bird and the (anecdotal) monkey are able to rely on their audience responding in that way.

39. What we learn from the discussion of the primitive objects of reference is that these alarm calls, at least when they are being learned, refer not to the predators per se, but to predator-like behavior. This realization helps us complete an ontogenetic picture. For vervets, at least, reference starts with reference to certain behavioral patterns. Through a process that involves ostensive definition this gets modified so that it applies to either specific behavioral patterns or to a combination of behavior and morphology (we think the latter is more likely, but this is an open question and one that should be easily testable). It seems likely that only after the ability to refer correctly has been learned is the vervet able to refer to absent objects in the false signalling sense discussed above. And only after the vervet can refer to absent objects in false signals, could it engage in the kind of reference that we have labelled as "conceptual". Notice that even if the vervet could engage in conceptual reference, the signals it produced would evoke behavior as if the referent were actually present. Its audience, in other words, is unable to comprehend

reference as anything beyond proxy reference. There are two lessons here: one is that successful reference is a product of an interaction between the producer of the reference and the producer's audience; the other is that there are developmental stages in the ontogenesis of reference. Does ontogeny recapitulate phylogeny? Are the ontogenetic stages reflections of the phylogenetic development of reference? If so we might expect to find organisms that have developed referential abilities of the first stage, but not the second, the way vervets have apparently developed abilities at the second stage and not the third. This also raises questions about human development. We have pointed out that there are these stages in the ontogeny of reference for vervets; are there the same stages in the ontogeny of human reference? Finally, this raises questions for the evolution of language: if we find that these stages do occur in humans as well as in other species, this would suggest that the development of language is neither highly unique nor highly modular. We think comparative psychology and ethology are on the right track here; these are interesting questions and the comparative evidence is well worth looking at.

40. Further support for the idea that reference to behaviors is more basic in the early stages of the evolution of language comes from the work of Tanner and Byrne (1996) who have studied the spontaneous gestures of lowland gorillas. Their analysis reveals that these gestures typically refer to desired behaviors such as play or copulation. They remark that:

    - It is likely that the earliest iconic depiction by human ancestors was of actions rather than objects, since it is representation of action that we find in the gestures of extant apes.

    Tanner and Byrne are engaging in exactly the kind of research that would shed light on the questions we're raising here: a comparative study which looks at the development of reference. We find it encouraging that their results confirm the hypotheses we draw from other research into the development of reference.

41. Thus far we have raised more questions than we have answered. This suits our purposes well, for we are more interested in presenting interesting directions for future research than in defending specific theories. In this vein, it is also worth pointing out some implications of these ideas about the evolution of reference for a puzzle about ostensive definition that exercised Wittgenstein and Quine, and continues to exercise those who are interested in the acquisition of language by children. Human children show a remarkable ability to identify the correct reference of learned count nouns such as "dog" or "bird" (Hall 1994). To explain this ability it is common to appeal to the concept of a "medium-sized" or "basic-level" category specified in terms of perceived similarity to a morphologically-identified prototype (Rosch 1978). But the designations "medium-sized" and "basic level" can seem suspiciously ad hoc for in some cases the category is at the taxonomic level of species or family (e.g. "dog") and sometimes at the level of a class (e.g. "bird") and the only reason for picking these levels as basic is that these are what the language-learners in fact latch onto. The degree of abstraction in these categories does, however, correspond closely to those identified by Cheney and Seyfarth as the referents of vervet monkey alarm calls (e.g., "leopard", "snake", or "eagle") which also correspond to different taxonomic levels, and they are considerably less abstract than the categories identified by Marler and Evans as referents of the alarm calls of chickens (e.g., "avian predator" and "terrestrial predator"). It is possible that these translational differences reflect nothing more than a lack of knowledge about the vocalizations of chickens and vervets. But if the differences are real, then data about the ontogeny of vervet alarm calls suggest an interesting hypothesis about the evolution of communication systems and the ability to conceptualize objects independently of their typical behaviors.

42. The hypothesis is that if the ontogeny of sophisticated referential skills recapitulates the phylogeny in a transition from action-based categories to feature-based categories, then it may not be necessary to postulate morphologically-specified, innate "middle-sized" or "basic-level" categories to explain how children settle on a reasonable level of interpretation for count nouns.

The behavioral differences between dogs and cows (although belonging to the same taxonomic order) may have been evolutionarily salient to humans in ways that differences between behavior in different species of birds was not and thus humans might innately be disposed to categorize initially according to such behaviorally specified categories. An evolutionary and comparative approach to the notion of reference and similarities to various nonhuman communication systems may thus help provide a specification of the basic categories that facilitate the earliest stages of language acquisition. Furthermore, worries about just how it is that language-learners manage to latch onto the correct referents may be mitigated by appeal to behaviors. Because motion against a background is more easily discriminated than the boundaries of a static object, it may be possible for language learners and language teachers to rely on the innate tendency to attend to behaviors in order to bootstrap reference to its first stage. The next stage, reference to objects, may then be learned as a consequence of exposure to uses of a word or other signal in cases where the object is not behaving in its typical fashion.

43. In the end we've told a partial story about the complex mosaic of traits that make up human language. We've tried to pick on one of these traits, reference, and describe the varieties of forms that it might take in organisms belonging to various species. Our understanding of these forms has been helped immensely by empirical research that was conducted with an explicitly comparative agenda and that would have been unavailable from a purely hardline approach. Finally, we have a second verse that we wrote for Quine's ditty that summarizes our findings, and although we're still a bit embarrassed about revealing it, here it is:

The well refined and tutored mind
Of *homo sapiens* is
Evolved to treat things inconcrete
From moving references.

Acknowledgments

We thank Marc Bekoff, Denise Cummins, Heather Gert, Patricia Greenfield, and Marc Hauser for comments.

References

Allen, C. and Bekoff, M. 1995. Function, natural design, and animal behavior: Philosophical and ethological considerations. *Persp. Ethol.* **11**:1-46

Blake, R. 1993. Cats perceive biological motion. *Psychological Science* **4**: 54-57.

Caro, T. M., and Hauser, M. D. 1992. Is there teaching in nonhuman animals? *Quarterly Review of Biology* **67**: 151-174.

Cheney, D. L. and Seyfarth, R. M. 1986 Vocal development in vervet monkeys. *Anim. Behav.* **34:** 1640-1658

Cheney, D. L. and Seyfarth, R. M. 1990. *How Monkeys See the World: Inside the Mind of Another Species.* University of Chicago Press, Chicago

Dawkins, R. and Krebs, J. 1978 Animal Signals: Information or manipulation? In J.R. Krebs and N.B. Davies (eds.) *Behavioural Ecology: An evolutionary approach*, First edition. Sinauer Associates, Sunderland, Mass.

Deacon, T. W. 1991. Anatomy of hierarchical information processing. *Behav. Brain Sci.* **14**: 555-557.

Dennett, D. C. 1983 Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. *Behav. Brain Sci.* **6**: 343-390

Dennett, D. C. 1995. *Darwin's dangerous idea: Evolution and the meanings of life.* New York: Simon & Schuster.

Descartes, R. 1637/1985 *Discourse and Essays*, published in *The Philosophical Writings of Descartes*, Vol 1., Cottingham, John, Stoothoff, Robert, and Murdoch, Dugald, trans. Cambridge: Cambridge University Press.

Evans, C. and Marler, P. 1995. Language and animal communication: Parallels and contrasts. In *Comparative Approaches to Cognitive Science.* J.-A. Meyer and H.L. Roitblat (eds.) MIT Press, pp. 341-382.

Greenfield, P. 1991. Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behav. Brain Sci.* **14**: 531-595.

Hall, D. G. 1994. How children learn common nouns and proper names: a review of the experimental evidence. In *The Logical Foundations of Cognition.* J. Macnamara and G. E. Reyes (eds.). New York: Oxford University Press, pp. 212-240.

Hauser, M. D. 1989. Ontogenetic changes in the comprehension and production of vervet monkey (*Cercopithecus aethiops*) vocalizations. *J. Comp. Psych.* **103**: 149-158.

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

Hauser, M.D. and Andersson, K. 1994. Left hemisphere dominance for processing vocalizations in adult, but not infant, rhesus monkeys: field experiments. *Proceedings of the National Academy of Sciences* **91**: 3946-3948.

Hauser, M.D. and Marler, P. 1993a. Food associated calls in rhesus macaques (*Macaca mulatta*). I. Socioecological factors. *Behav. Ecol.* **4**: 194-205.

Hauser, M.D. and Marler, P. 1993b. Food associated calls in rhesus macaques (*Macaca mulatta*). II. Costs and benefits of call production and suppression. *Behav. Ecol.* **4**206-212.

Johansson, G. 1973. Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics* **14**: 201-211.

Lieberman, P. 1995. What primate calls can tell us about human evolution. In *Current topics in primate vocal communication.* E. Zimmerman, J. D. Newman, and U. Jurgens (eds.). New York: Plenum Press, pp. 273-282.

Locke, J. 1995. Linguistic capacity: An ontogenetic theory with evolutionary implications. In *Current topics in primate vocal communication.* E. Zimmerman, J. D. Newman, and U. Jurgens (eds.). New York: Plenum Press, pp. 253-272.

MacNeilage, P. F., Studdert-Kennedy, M., and Lindblom, B. 1991. Primate handedness reconsidered. *Behav. Brain Sci.* **10**:247-263.

Marler, C. and Evans, P. 1995. Bird calls: just emotional displays or something more? *Ibis* **138**: 26-33

Marler, P., Karakashian, S., and Gyger, M. 1991. Do animals have the option of withholding signals when communication is inappropriate? The audience effect. In C. A. Ristau (Ed.), *Cognitive ethology: The minds of other animals. Essays in honor of Donald R. Griffin.* Hillsdale, New Jersey: Lawrence Erlbaum, pp. 187-208.

Munn, C.A. 1986. The Deceptive Use of Alarm Calls by Sentinel Species in Mixed-species Flocks of Neotropical Birds. In R.W. Mitchell and N.S. Thompson (eds.), *Deception.* New York: SUNY Press, pp.

169-175.

Pinker, S. 1994. *The Language Instinct.* New York: William Morrow and Company.

Quine, W. V. O. 1953. Identity, Ostension and Hypostasis. In *From a logical point of view.* Cambridge, MA: Harvard University Press.

Rauschecker, J. P., Tian, B., & Hauser, M. 1995. Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* **268**: 111-114.

Ristau, C. 1991. Aspects of the cognitive ethology of an injury-feigning bird, the piping plovers. In C. A. Ristau (Ed.), *Cognitive ethology: The minds of other animals. Essays in honor of Donald R. Griffin.* (pp. 91-126). Hillsdale, New Jersey: Lawrence Erlbaum.

Rosch, E. 1978. Principles of categorization. In E. Rosch and B. B. Lloyd (eds.), *Cognition and Categorization* (pp. 27-48). Hillsdale, New Jersey: Lawrence Erlbaum.

Seyfarth. R. M., Cheney, D. L., & Marler, P. 1980. Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour* **28**: 1070-1094.

Seyfarth. R. M., Cheney, D. L. 1986. Vocal development in vervet monkeys. *Anim. Behav.* **34**: 1640-1658.

Tanner, J. E. and Byrne, R. W. 1996. Representation of action through iconic gesture in a captive lowland gorilla. *Current Anthropology* **37**: 162-173.

Wittgenstein, L. 1953 *Philosophical Investigations*. New York: Macmillan