OUT OF THE MOUTHS OF BEASTS:
Animal Communication and Perspectives on the Evolution of Language

Erica Cartmill

Senior Linguistics Thesis

Swarthmore College

May 14, 2003

TABLE OF CONTENTS

I: INTRODUCTION

Scientists, philosophers, and theologians have long sought to define properties that distinguish humans from all other living things. It is commonly assumed that the ability to use language is one such property. Although other animals can undoubtedly communicate certain things, animal communication systems fall far short of the complexity and versatility of human language. Almost everything that we associate with human intelligence depends on our unique capacity to manipulate symbols and to use words to represent the world.

Darwinian theories about human evolution raise doubts about our uniqueness and pose particular questions about language as a marker of that uniqueness. If we accept that humans developed from non-human animals in a gradualistic, Darwinian fashion, then language must have ancestral forms in other species. Linguists in the twentieth century have struggled with the idea of a true "evolution" of language. To many students of language, it has seemed impossible that the transition from the situation-specific and often unconscious calls of animals to human language (which requires consciousness and has infinite applicability) could have been gradual or driven by natural selection. The theories of language evolution that have attempted to provide solutions to this problem are generally implausible and mutually contradictory.

Studies of the naturally occurring communication systems of nonhuman animals and attempts to teach human language to other species provide insight into simpler communicative structures. Such studies may help us to determine the most basic components of language and to identify possible precursors. By looking at alternative natural systems as "living fossils" of possible transitional stages in the evolution of speech, and by studying comparative structures of cognition revealed in the behavior of language-trained animals, we may be able to construct a more probable theory of the emergence of language.

II: APPROACHES TO THE EVOLUTION OF LANGUAGE

> Languages may change and ' evolve' in the sense of cultural evolution, but as far as can be determined, this is in the context of a fully developed language capacity. For the prior 5 million years, we can make only very indirect inferences based on the nature of artifacts such as tools and pictures, and on equivocal hints about the structure of the brain and vocal tract.
> — R. Jackendoff, 2002 (232)

The study of the evolution of human language is by nature a theoretical one. There can be no direct evidence of the emergence of language because we have no record of its early stages. While there are widely accepted theories about the structure of current human language, no single theory of its emergence has risen above the rest. It is generally accepted that early language accompanied a substantial increase in cognition and a restructuring of thought, but whether cognition or communication served as the driving force remains heavily debated (St. Clair 1973). In the absence of any hard linguistic evidence, we are forced to build theories based on secondary data from the fields of genetics, anatomy, comparative psychology, and behavioral ecology. The main approaches to language evolution can be divided into the following classes:

1. *Genetic Approach*: Examining the human genome for "language-specific" genes by comparing normal speakers to language-impaired families and the genomes of apes.
2. *Form Approach*: Comparing human language to other naturally occurring communication systems to gain insight into the common traits or functionality that may have provided a fitness benefit.
3. *Function Approach*: Comparing human linguistic abilities to those of language-trained animals tested in laboratories in order to determine the basic elements of

language and the cognitive processes linked to them.

There are also theories that deal strictly with data from human language and are focused more on how we structure full language than on how early language may have emerged (see Chomsky 1957). The most recent theories (which I consider to be the most complete) combine elements from several of the different approaches, combining genetic heritability with studies in comparative psychology or linguistics (Calvin and Bickerton 2001, Jackendoff 2002, Hauser 1997).

Placing language as an evolved trait within the context of other communication systems requires taking a stance on whether it was selected for as a primarily communicative or cognitive process. There are two competing schools of thought, and comparison of human linguistic ability to that of other species is more or less favorable depending on which approach is taken. There are those who believe that language plays a significant role in determining how we structure abstract thought, and that without words to label concepts, much higher-level reasoning would be impossible (St. Clair 1973). Conversely, there is growing support for the theory that cognition evolved before language and was only expressed later after argument structure had evolved in the mind (Calvin and Bickerton 2001, Jackendoff 2002).

Systematic studies or approaches to language evolution have been mainly a product of the last sixty years. Prior to the middle of the twentieth century, symbolic communication was seen as one of the uniquely human traits (along with tool use) and no studies were conducted on naturally occurring non-human communication systems. Language was regarded as something that had arisen among early humans from a necessity to communicate more easily (similar to the theory of Hockett and Ascher 1964). However, Noam Chomsky's (1957) theory of innate linguistic structure (universal grammar) changed the way language evolution was approached because it placed language into the set of heritable traits, meaning that it should have been subject to natural selection.

In a Chomskian framework, linguistic ability is seen as a genetically determined trait or trait complex inherited in a Mendelian fashion, rather than a socially learned system governed by intelligence. This is a problematic assumption from the standpoint of evolutionary biology. In order for linguistic ability to persist and spread in a population, it had to confer some advantage on its initial carrier, despite the fact that the founder individual would have lived and reproduced in an otherwise non-linguistic species. It is hard to see how the possession of language would provide a fitness benefit to a lone speaker in a speechless population. This paradox could be avoided if language had originated through the gradual accretion of small incremental improvements in communicative abilities happening on a group level; however, Chomsky's theories of language did not allow for a slow or stepped evolution because he saw grammar as a purely heritable mechanism unable to be broken down or acquired. As Jackendoff notes, "The common view of Universal Grammar treats it as an undecomposable 'grammar box,' no part of which would be of any use to hominids without all the rest" (Jackendoff 2002). Chomsky proposed that there was a heritable structure of language, and yet denied that natural selection could have led to language. This contradiction between an evolved innate trait and one that seems to have emerged out of nowhere has posed a problem for structural linguists ever since its introduction because it seems unlikely that any system so complicated could have arisen both spontaneously and intact.

The problems that arise from looking at language as a process in the mind of the individual rather than as a socially formed system raise the question of whether language is mostly a system of communication at all. It is possible that the communicative aspects of language arose as side affects of an evolution towards a system for mental representation and abstract thought (Bickerton, 1990, Jackendoff 2002). Bickerton defines the role of language as a cognitive one. "Language. . . is not even primarily a means of communication. Rather it is a system of representation, a means for sorting and manipulating the plethora of information that deluges us throughout our waking life" (Bickerton, 1990).

If language evolution is viewed as a byproduct of cognitive improvements and restructuring, it becomes possible to examine it through the process of natural selection and individual advantage. In this view of things, language is an emergent phenomena that comes into existence as an incidental side effect of the gradual accretion of abilities and traits selected for in another context. It could very well be the case that "the physiological prerequisites of language developed, in proto-human populations, in a manner having nothing whatever to do with their subsequent linguistic expression" (Toulmin, 1972:459). Accordingly, the earlier theories that placed language evolution at the group level have mostly given way to theories grounded in the idea of a heritable universal grammar and the assumption that the primary function of language is the structuring of thought (Hauser 1997). In order to address many of the questions raised by the problem of the seemingly impossible jump from animal call systems to full-blown human language, it is necessary to present several different theories.[1] Current theories can be grouped under three main headings:

1) *Slow Evolution*: natural selection for increased communication (Hockett and Ascher, 1964)
2) *Catastrophic Mutation*: single mutation allows grammar to emerge intact (Chomsky, 1957)
3) *Protolanguage*: intermediate step between no language and full language, essentially comprising semantics without syntax (Bickerton 1990, and Jackendoff 2002)

[1] For a concise summary of other approaches see Hauser, 1996: Chapter 2

These three theories form a foundation for a dialog about the advantages and purpose of language, and can be used as a set of tools to examine work in the companion fields of genetics, anatomy, comparative psychology and behavioral ecology. There are other theories dealing with the development of human vocal production mechanisms (notably Lieberman 1991), which I will address briefly later. The theories that I find the most plausible are those that combine elements from many of the different approaches and include both a period of slow evolution and a rapid radiation. Chomsky himself, along with several other prominent structural linguists (Pinker 1994, Calvin and Bickerton 2001, Jackendoff 2002), is now advocating a slower evolution based on primate group structure that echoes earlier theories (Hockett and Ascher 1964). Because there is an emerging consensus that the physiological changes involved in the origin of language were driven by the burgeoning communication system (Jackendoff 2002), I will focus on the evolutionary theories concerning the emergence of a structured form of communication rather than the capacity for speech itself.

SLOW EVOLUTION

In order to build a theoretical model of how our earliest ancestors acquired language, Hockett and Ascher (1964) first consider the environmental and social pressures that affected the last common ancestor of humans and apes. These proto-hominids were probably smaller than humans, lived in territorial, polygynous bands of 10-30, and possessed some sort of call system similar to those of many monkeys today.

The call system possessed by this common ancestor would have almost certainly exhibited certain characteristics that are still common among many primate species. It would have consisted of a small repertoire of calls that were situation-specific (able only to refer to certain events/stimulus: food, danger etc.) and mutually exclusive (no call contained parts of another). This system would have lacked displacement (the ability to "talk" about things not present) and so calls would occur only in the presence of their referents. Also, the calls would have lacked duality of patterning (the breaking down of the smallest units of meaning into common units of sound so that each meaning unit is recognized by the pattern of smaller parts rather than by an entire unique sound). The system would thus have lacked anything resembling the morphemes and phonemes of language.

According to Hockett and Ascher (1964), call systems and language differ in four principal respects:

CALL SYSTEM	LANGUAGE
1) **closed** – no recombination	1) **open** – recombination handles new circumstances by using elements of other utterances in new ways
2) **no displacement**	2) **displacement** – allows individuals to refer to objects of events outside of the context in which they occur
3) **no duality of patterning**	3) **duality of patterning** – smallest units of meaning are recognized by their unique pattern of smaller sound-units; larger units of meaning are constituted by the arrangement patterns of smaller ones

4) **largely heritable**	4) **requires teaching** – possibly some genetic component for capacity to learn, but must be acquired through social interaction

These two systems are diametrically opposed with respect to these characteristics, and it is hard to see how one could develop into the other naturally. Hockett and Ascher (1964) conjecture that the necessary elements for language were selected for slowly, and that the progress was linked to other physiological and cultural changes (such as bipedalism and hunting) affecting early hominid populations.

Hockett and Ascher (1964) posit that early hominids could have developed new strategies of calling by combining their existing calls to achieve novel meanings. They theorize that a closed call system could have become an open one through an accumulation of event combinations leading to an eventual restructuring of the system. They suggest that the bands of early hominids occasionally faced calls in which more than one call was appropriate. Suppose the band already had calls of the pattern ABCD and EFGH, meaning "food" and "danger" respectively. It does not seem unlikely that there must have been some situations where they encountered both the presence of food and danger. In such circumstances, they might choose to communicate whichever seemed more important at the time, or give first one call and then the other. However, Hockett and Ascher theorize that early hominids might occasionally have combined the calls, giving first part of one and then part of the other, possibly fitting them into an already existing call structure, creating a call of the form ABGH. Such a call would manifest potential or incipient open call structure that allowed for a duality of patterning at the morphemic level.

Emergence of open structure through this sort of coincidence is, in my opinion, highly unlikely. It would require not only the innovation of the signaler at each successive event, but also the innovation of the listener. Such a step would require much trial and error until there became any predictable form or ability to understand the meaning. Any signaler that unsuccessfully combined two calls (i.e., was misunderstood by the rest of the group) would have a significant disadvantage when compared to a signaler who, instead of combining, gave both calls and transmitted both situations. However, if it were possible to make the jump from individual, phonetically isolated calls to combinations of call aspects that carried meanings of both the original calls, then the systematic evolution proposed by Hockett and Ascher might be possible.

If calls could be combined to handle situations in which both signifiers were present, then the smaller units of sound used in the combination would carry the semantic content of the original call. In order for the original call to remain intact, the uncopied part of the call might have taken on a new and specific meaning. If ABGH meant "food + danger," then ABCD would have meant "food + no danger," and EFGH meant "no food + danger." Hocket and Ascher further speculate that this situation occurred for nearly all of the calls in the "lexicon," and that eventually, individuals would start to use combinations of 3 or more calls at once. This stringing together of smaller units of meaning would have effectively "opened" the system from the limitations of the initial call structure.

It would be hard to make the transition to combining calls because of the low frequency of coming across situations in which two different calls had to be expressed. Combinations of more than two calls seems even less likely. However, even supposing the system could be "opened" in this way, the ability to combine smaller units of meaning would still not evade the fact that the calls were elicited only by the presence of the referents. Hockett and Ascher (1964) state that since an open system cannot be entirely heritable, teaching and learning would therefore become much more important. They speculate that there must have been a system for young individuals to practice vocalizing, probably a sort of verbal "play." Play in juveniles involves practicing adult social behaviors with one another, usually out of context, so that they will be proficient at them later. Young horses, dogs, bears and many other mammals play-fight with each other. So it seems possible that the young of early hominids could have "play-talked" with each other in order to gain an auditory feedback period like that seen in songbirds (Brainard and Doupe 2002). However, in hominids, this period would have been used to improve social signaling rather than just to learn to be able to produce the correct sounds. If this occurred, Hockett and Ascher believe that it would have constituted displacement, another one of the basic characteristics of true language.

If young engaged each other in mock situations and play-called in conjunction with their games, then it would show that they were able to refer to objects outside of their immediate surroundings and they would have achieved displacement or essentially abstract semantic content. However, it seems highly unlikely that this sort of play-calling could have persisted, or that it could lead to adult call displacement even if it did persist. If adults heard their offspring "playing" and giving alarm or food calls, they should act as if the calls were real and respond to them. This would quickly lead to a breakdown of the listener response. As in the fable of the boy who cried "wolf," either the young in question would not be believed in future circumstances, or the calls would cease to elicit responses in general.

In some species of monkey, juveniles learn meanings of calls through giving calls to many stimuli and gauging the responsiveness of adult members in the group. In vervets, the young must learn to give alarm calls to martial eagles and not to other aerial stimuli (Cheney and Seyfarth 1990). At first, the juveniles give alarm calls to all types of eagles, other sorts of birds, and occasionally even falling leaves. The adults respond by looking up every time the juvenile calls, but give calls themselves only if they see a martial eagle. If juvenile vervets started to use these calls playfully by giving them when there was no stimulus (even an incorrect one), then the adults might learn to ignore them altogether and the juveniles would lose the benefit of the indirect teaching that the reinforcement of the adults' response provides. Also, if the calls were frequently given out of context by juveniles, they would never learn to use them in the right context unless they began playing with situational calls only after they had already learned them. In such a case, what would be the point of "playing at calling" in order to practice and learn? Play can be very useful in developing some social behaviors such as fighting, but it seems like an ineffective tool for learning the meanings and context of specific calls.

Hockett and Ascher's (1964) speculations on the jump to duality of patterning are equally shaky. They theorize that eventually the available sound space would have filled up and the small units of meaning would start to sound like each other—the morphemic bits would start to run into each other and be confused in the mind of the listener. They assume that, at this point in language evolution, most groups would experience a breakdown in their communication system and probably go extinct as a result. However, they conjecture that at least one group made the miraculous switch to interpreting morphemes not by the overall sounds that composed them, but by the patterns of the sounds. These sounds would then take on individual characteristics and be reusable in different morphemes. These reusable sound units would become individually distinguishable phonemes.

CATASTROPHIC MUTATION

Prior to the Chomskian innovations of the 1960s, it was not difficult to place language evolution within a framework of group selection. A natural progression from primate gestures and calls to human language seemed probable given the advantages of increased communication and abstract thought (Mowrer, 1960, Skinner 1957). The improvements in linguistic ability were thought to have developed at the group level and to have been driven by a need for increased communication.

When Chomsky proposed his theory of universal grammar (1957), the scientific community began to adopt an "all or nothing" approach to structured language. The idea of a heritable "grammar box" made the theory of a slow evolution from calls to language impossible. The grammar box would be useless if incomplete and so would have had to arise in a single step, as there would be no advantage to having one part without the whole. Chomskian linguistics "places language in the individual mind rather than in the social group to which the individual belongs" (Knight et al. 2000). This shift in focus from the social unit to the individual meant that language could now be examined through the process of natural or "individual" selection.

In order for grammar to have emerged as a single unit (as Chomsky posits), there must have been a spontaneous mutation or physiological change that gave one individual the ability to structure its utterances. Because there must have been an advantage to the individual even if no other individuals possessed the same ability, grammar must have been linked to some other process or ability that gave a significant advantage to the individual and its progeny.

Evidence has emerged over the past few decades that has supported the single mutation theory. There are three highly convincing arguments for a rapid jump to language:

1) The rapid expansion of the fossil tool kit: The fossil record shows that there was an extremely rapid development of tools for specific functions around the time of the emergence of *Homo sapiens* (about 200,000 years ago: Klein 1989). This rapid differentiation of tools is thought to have accompanied a boom in language ability (Davidson and Noble 1993). Although there is no evidence for such a linkage, the radiation of tool type occurred without an increase in brain size, so it seems probable that it accompanied the spread of some other behavioral or communicative process that played an important social or cognitive role in the lives of early man.

2) The study of pidgin languages has revealed that while the first generation to form a pidgin between languages uses a linguistic form that contains shared labels for items or concepts but no regularized grammar, their children will use the vocabulary of their parents in a grammaticized way (Bickerton 1981). In essence, structure arises spontaneously in a single generation with no intermediate stages. This serves as a present day example of the "all or nothing" nature of human grammar.
3) Recent studies of mitochondrial DNA have suggested that all humans are descended from one female who lived approximately 220,000 years ago (Hauser 1997). If this is true, perhaps this female was the origin or carrier of the single mutation that restructured thought and altered the vocal tract, making speech possible. And if this is the case, then it is likely that her progeny would have been highly successful and could have out-competed other lineages that lacked language.
4) Even more recent analysis of the human genome has revealed FOXP2, a highly conserved gene that codes for a neural processor (Enard et al. 2002). Humans possess 2 copies of this so-called "language gene" while apes have only a single copy. It is not clear what function the gene actually serves, but there is some evidence that individuals that lack expression of the gene are linguistically impaired.

The single-mutation theory was problematic because it relied on a chance mutation to alter several physically unconnected, complex systems at the same time. This mutation would have had to:

1) Rewire neural circuitry to support syntax as a cognitive tool since it would be involved in both production and reception.

2) Modify the vocal tract to allow for a wider range of sounds to be produced.

That a single catastrophic mutation could have altered all of the secondary structures involved in the production and synthesis of language simultaneously seems impossible. As Hauser notes:

> The probability that such complexity of design is the result of a single mutation is extremely low. Even something *relatively* more simple, like the vertebral column, is coded by hundreds of genes, responsible for segmentation and structural arrangement (Hauser 1997: 41).

PROTOLANGUAGE

There are some who believe that "the communicative advantages of human language are just the kind of cognitive phenomena that natural selection is sensitive to" (Jackendoff 2002, Pinker and Bloom 1990). This view is more in accordance-with Hockett's original analysis of language as a slowly evolving communicative system than with Chomsky's "big bang" hypothesis. However, the theories of language evolution that have arisen post-Chomsky have nonetheless attempted to reconcile Chomsky's view of a heritable linguistic structure with a slower progression that would have provided fitness benefits at multiple stages (Bickerton 1990, Calvin and Bickerton 2001, and Jackendoff 2002).

It is possible to find a compromise of sorts between a Chomskian and Darwinian analysis of language by using a two-stepped evolution. These two steps consists of a slow accumulation and improvement of signals which plateaued and were than supplemented by a mutation that allowed the structures of full language to be applied to the heightened vocabulary. This speculated plateau in between no language and full language is very interesting because there has been some evidence showing that it may well exist. Bickerton (1990) calls this stage "protolanguage," defining it as a single intermediate stage between no language and full language. This protolanguage stage consists of a large lexicon but no structured grammar. It is, essentially, semantics without syntax (Bickerton 1990).

Derek Bickerton proposed the theory of protolanguage after studying the pidgin languages that formed when tradespeople or sailors speaking different languages came into contact for extended periods of time. He noticed that although they constructed a common lexicon of nouns and verbs, it was expressed only in very short word strings and with few grammatical rules. However, the next generation, if they grew up around parents speaking pidgin frequently, used a more regular word order and spoke something that more closely resembled a full language than limited trade banter. Within very few generations, a grammatically limited pidgin dialect will become a creole and will use a regular structure and vocabulary; it will become a full language.

This shows that in modern humans, protolanguages are unstable. We have some sort of internal structure or patterning that (when given exposure to referential labels and opportunity to interact with other speakers) will apply itself to our utterances. Humans will order and structure labels if exposed to them at the right age. There is even evidence that grammar will arise "spontaneously" within groups of people who have no exposure to full language, such as groups of deaf children who have not been taught a sign language (Goldin-Meadow 1983).

Modern humans may be predisposed to grammatical structure, but it is quite possible that our ancestors operated for a long period of time using some sort of protolanguage. Bickerton explains that there are several examples of protolanguage that are visible today (Bickerton 1990). Jackendoff (commenting on Bickerton's work) divides up the examples into the categories of "disrupted language" and "not fully developed language" (Jackendoff 2002). The disrupted category contains:

1) pidgin languages
2) humans that learn language after the sensitive period
3) agrammatic aphasiacs

These categories involve cases where full language was disrupted either externally for want of interaction or instruction, or internally by affecting an area of the brain that controls grammar. Developing language also passes through a stage of protolanguage, as does language that never becomes fully realized. Cases of developing and under-developed language include:

1) early child language
2) second language learning
3) language taught to apes

Language learners pass first through a phase of vocabulary building where they learn labels for objects and ideas and then later progress to a stage of learning how to structure utterances. This is apparent to anyone who has attempted to learn a second language later in life. Before learners become highly proficient at a language, they will use strings of nouns and unconjugated verbs or relational words without much grammar or predictable word order. The same deficiencies can be seen in young children around the age of two (Haynes 1998). Experiments with attempting to teach language to apes have yielded similar results (see Savage-Rumbaugh 1998, Gardner and Gardner 1969). In all cases, the apes have learned many referential signs or symbols, but none of them have shown much ability or inclination to use their signs in a regular order.

It is possible that language developed by first passing through an evolutionary phase of protolanguage. This would allow natural selection to work on individuals and groups with the communicative advantages of an increased vocabulary without claiming that modern grammar was subject to the same slow evolution. Bickerton's (1990) theory still relies on a single leap from protolanguage to language, but because of its intermediate stage, it seems more plausible from an evolutionary standpoint than Chomsky's version.

In response to criticism of his reliance on a catastrophic mutation to develop grammar, Bickerton proposed a theory of slower (yet still rapid) genetic assimilation through cumulative "Baldwin effects."[2] This theory relies on a single mutation that does not necessarily restructure cognition, but makes it possible to express previously-evolved concepts of AGENT, THEME and OBJECT. Bickerton proposes that these concepts evolved from a need to remember debts and credits in a social system that contained reciprocal altruism (Calvin and Bickerton 2001). Several studies have supported the theory that prelinguistic argument structure is needed in social primates with complex social "rules" and stable dominance hierarchies (Cheney and Seyfarth 1998, Dasser 1988).[3]

Building on the theory of pre-existing argument structure, Bickerton and the neurologist William Calvin have more recently developed a theory that attempts to merge the possibility of an explosion of grammar with what we know about the structure of the human brain. Bickerton created a model of the theoretical processes in the brain that contribute to the production of language and attempted to apply it to Calvin's model of neural-linguistic

---

[2] "Baldwin Effect": If organisms are capable of learning some task that is important in their environment, natural selection may favor those individuals who, by virtue of genetic variation, happen to have an innate 'leg up' on learning the task. (Baldwin, 1896 as cited in Jackendoff 2002: 237).

structure. In this way, they developed a theoretical model of the pre-linguistic yet socially-savvy brain [Figure 1].[4]

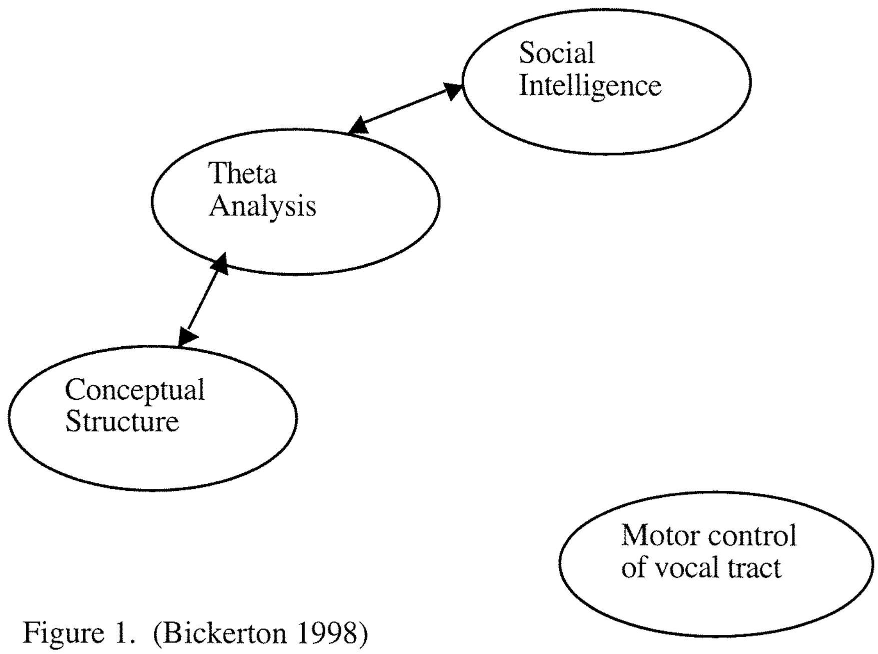

Figure 1. (Bickerton 1998)

Calvin and Bickerton theorize that the connections between the conceptual, theta analysis, and social intelligence centers in the brain were all strengthened through natural selection favoring increased social awareness and planning future events based on past encounters. These centers had to exist in full before a grammar could develop. Thus, although the early hominids were getting smarter and more socially "savvy," their communications did not develop past the point of a "protolanguage." They could produce some calls that were referential (much as some monkeys today do: Zuberühler 2003). Hypothesis testing may even have been possible in a pre-linguistic mind of this sort. Individuals could have formed mental signs for referents and been able to manipulate basic relationships between arguments without being able to trigger the vocal production

---

[3] The role social awareness plays on developing argument structure is discussed in greater detail in section VII.

[4] Figures adapted from Bickerton 1998

mechanism. However, they were unable to vocally express the complicated relational structure between objects that they were able to perceive.

Calvin and Bickerton speculate that eventually there occurred a fortuitous connection between disparate functional units within the brain. Somehow a bridge was formed between the unit controlling conceptual structure and the vocal production mechanism [Figure 2].

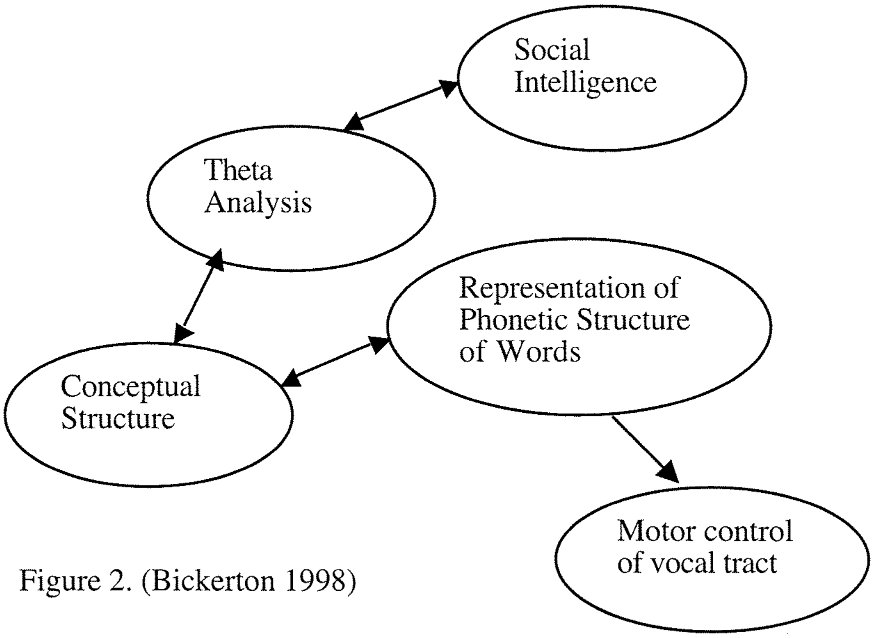

Figure 2. (Bickerton 1998)

Calvin and Bickerton conjecture that this bridge may have formed when the reasoning center grew larger and overlapped or touched the production center. Then neurons in one could have caused the other to fire and these hominids may have found that rather suddenly they were able to vocalize in accordance with their mental mapping. It could also be the case that a somewhat less catastrophic mutation made it possible for the thought center to affect the production mechanism. This theory is an interesting and compelling congeries of many disparate theories. It provides a neurological grounding for a theory of stepped language evolution and addresses the huge difference in non-human linguistic abilities seen in the laboratory versus the field.

However, the problem remains that there is no evidence that such abstract centers exist in the brain. There appear to be different regions that control grammar and semantic content (McMillan 2003), and different types of aphasia can cause humans to revert to protolanguage when one area of the brain is affected and meaningless but syntactically correct babbling when another is; but there is little support for the abstract centers and bridges that make Calvin and Bickerton's theory so clean.

Jackendoff (2002) presents a model that incorporates Bickerton's stage of protolanguage, but allows for a more gradualistic transition to language. The advantage to Jackendoff's structure is that it inserts steps before and after protolanguage, creating a more detailed picture of the possible evolutionary process and affording more scope for the operation of Darwinian natural selection. Jackendoff also gives several possible directions that could be taken at each intermediate step, showing different strategies that could have led to the next major stage. His summary of incremental evolutionary steps is reminiscent of Hockett and Ascher's original theory of a slow evolution [Figure 3][5]. Its most important difference is its postulation of protolanguage as an open-ended phase of language evolution that could have been a semi-stable system persisting for many generations.

[5] Figure 3 has been adapted from Jackendoff 2002

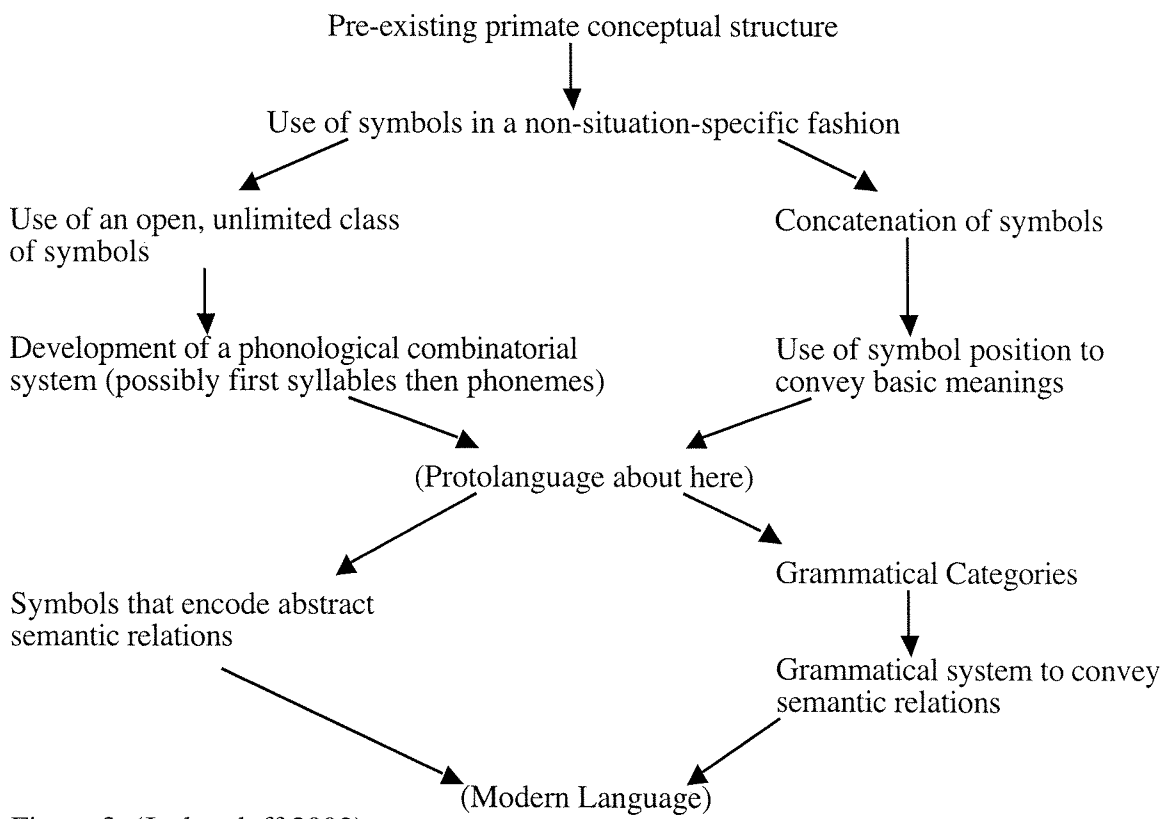

Figure 3. (Jackendoff 2002)

Jackendoff posits that the most important stage in language evolution was not the emergence of grammar, but the ability to use symbols in different ways to refer to different things (Jackendoff 2002). He remarks that "many researchers shortsightedly view [symbol use] as a straightforward and inevitable development of language from such humble beginnings" (Jackendoff 2002: 239). Jackendoff stresses that the initial stage of symbolic representation "consisted of single-symbol utterances, lacking combinatoriality" (Jackendoff 2002: 239). He notes that the use of single-symbol utterances differs greatly between young children and apes. While children can use a single symbol to refer to many things, whether naming an object, asking where it is, pointing it out, asking for it, or noting resemblance to the object, non-human primates seem to be much more limited in their application of symbols (Jackendoff 2002). While an alarm call can be used to alert others to the presence of a predator (Cheney and Seyfarth 1990), it cannot be used to ask about such things as the location of the predator or the time of the last sighting.

There were two major innovations needed to get from single-symbol utterances to language: a large lexicon (open class – unlimitedly large), and the beginning of syntax. The transition from a closed system of holistic signs to an open class system of calls is not a clear one. Jackendoff (2002) proposes a solution that echoes Hockett and Ascher (1964). Jackendoff posits that "proto-syllables" developed and were eventually combined. He defines proto-syllables as holistic vocal gestures that could be combined to make 2- or 3-syllable "words." The emergence and combination of these syllables could have been an intermediate step in the development of an open-class system. Jackendoff believes that there are remnants of these protosyllables in the utterances of very young children. One-year-olds speak using syllables that are not broken down fully into phonemes and prefer words with simple syllables that do not vary in the point of articulation (Levelt 1994). This initial preference of simple syllables without a full duality of patterning (i.e., with no phonemic distinction) may represent an early phase that language passed through before the emergence of phonemes and complex words.

Whether early utterances passed through a phase of proto-syllables or not, concatenation of symbols is a necessary first step in language evolution. Concatenation of symbols can allow one symbol to modify the meaning of another even without syntax (Jackendoff 2002). Consider the utterances of young children such as "mommy cookie" or "cookie mommy." The order of the symbols is not as important as the fact that they are used together in a string (in this case to mean "mommy, give me a cookie"). Apes seem to reach this point, combining several symbols in strings (often with repetition), but do not show a stable word order (Savage-Rumbaugh 1998, Gardner 1969).

Jackendoff speculates that predicable word order might have developed after basic concatenation as an adaptation for increased clarity. Once symbols are frequently combined to create complex meaning, other aspects of grammar may have emerged slowly, selected for to increase clarity of meaning. Language can be used to communicate many things even when it is very underdeveloped in terms of a regular grammar. Evidence from this can be seen in the pidgin languages studied by Bickerton (1990) as well as in the language acquired by migrant workers (Klein and Perdue 1997).

Klein and Perdue (1997) introduce the idea of "Basic Variety" as an intermediate stage of language comparable to Bickerton's protolanguage. They studied migrant workers who picked up aspects of the local language but who were never directly taught. They claim that the distinguishing features of Basic Variety are:

1) Lexical
2) Absence of inflectional morphology (verbs don't conjugate)
3) Omission of arguments (no mandatory subjects or objects)
4) Absence of sentential subordination (no relative clauses, etc.)
5) Simple semantically-based principles of word order (AGENT first and FOCUS last)

These characteristics provide a more complete analysis of the limited grammar use underlying disrupted or imperfectly developed language than the "semantics without syntax" definition of protolanguage. Pidgin languages (such as those studied by Bickerton) also seem to adhere to these rules (Givón 1995).

The limitations of Basic Variety may well have existed for years until a need for increased clarity caused the introduction of ordered grammar. The basic word order of "AGENT first FOCUS last" may, in fact, be a "fossil principle" that served as the first ordering mechanism in protolanguage (Jackendoff 2002).

Another possible protolanguage "fossil" is the use of compound nouns. These are used in a predictable way where the second word serves as the "head" and the first word takes on the role of an adjective modifying the head. The remnants in modern language may serve as an alternative process by which word combinations and concatenation emerged from a lexicon of holistic utterances. Jackendoff (2002) cites examples such as "doghouse" and "housedog" where the order of combination changes the meaning. He also cites "wheelchair, snowman, garbageman, milkman" etc. In every case the combination is used in a slightly different way: to refer to the location of the head (e.g. "housedog"), to denote the matter of the head (e.g. "snowman"), or to imply an action or task (e.g. "milkman").

Compounding in a predictable way may have been the first use of grammar. This is the only sort of morphology found in Basic Variety and in the early utterances of young children (Jackendoff 2002). It is possible that these basic ordering principles could have developed in conjunction with concatenation. There is no reason to believe that unordered strings of symbols were assembled first; it merely seems more plausible to assume that mechanisms of greater complexity were added later when placing developments within an evolutionary progression (Jackendoff 2002). Compounding and concatenation may have been the mechanisms that moved early referential symbols to an open class structure and formed the foundation of grammar. (As will be shown later, the linguistic mechanisms that arose from the first combination of symbols can be used to compare language ability between species.)

The theories of protolanguage and Basic Variety both rely on an early stage of language that was marked by an absence of formalized grammar but which used a large lexicon of referential symbols and occasionally combined them to create complex utterances. Basic Variety and Protolanguage represent essentially the same stage of linguistic ability. However, Bickerton (1990) does not believe that protolanguage is a full language, while Perdue (1997) believes that Basic Variety represents modern language at its most basic. It seems safe to say that protolanguage and Basic Variety possess part of Universal Grammar (phonology), but not all (incomplete syntax).

III: WHY STUDY ANIMAL COMMUNICATION?

> Our fellow creatures can tell us the most beautiful stories, and that means *true* stories, because the truth about nature is always far more beautiful than even what our great poets sing of it...In so far as the 'signal code' of a species of social mammal can be called a language at all, it can be understood by a man who has got to know its 'vocabulary.'
> — Konrad Lorenz, *King Solomon's Ring* (1952)

The main limitation on the field of evolutionary linguistics is that there is no way to tell which theory of language evolution is more "probable." All that can be studied is the fossil evidence of anatomical changes in early hominids and the "living fossils" seen in disrupted or under-developed human language. There is an understandable desire to reach beyond intra-human comparisons and to look for the precursors of human language in the call systems of primates. This often means assuming that because they are our closest living ancestors, they represent a "starting point" for our communication system and many of our behaviors. However, it is important to remember that apes, are not in fact our distant ancestors, but are merely the family that shares the closest common ancestor with humans. It is a mistake to think of their behaviors or communication systems as less-developed versions of our own, since they are equally evolved (meaning "have evolved for the same amount of time"), but represent a different evolutionary path from our last common ancestor.

If animal communication systems do not necessarily represent earlier stages in the evolution of human language, what can be learned about the evolution of language by studying them? There is no single answer to this question, but there are several arguments that parallel those made in the study of cross-species physiological similarities.

In attempting to determine how a highly complex organ such as a human eye could have emerged by gradual improvements from simpler antecedents, Darwin turned to comparison of similar structures in other living organisms (Darwin 1859)[6]. It seems impossible that a mechanism as complex and interconnected as the eye could have evolved gradually rather than emerging fully functional. And since the eye is rarely preserved in fossil vertebrates, it is impossible to recover intermediate stages in its evolution from the fossil record. It was therefore necessary for Darwin to regard simpler eyes in other living species as "living fossils" that modeled potential intermediate stages of the evolution of the eye. This provided him with a greater perspective on how eye-like structures far simpler than our own could still provide a fitness benefit. In a similar manner, looking at the appearance of certain elements of language use or ability in other animals can model potential intermediate stages in the evolution of language.

It is also advantageous to study the parallel evolution of similar structures even if they are not homologous. To return to our eye analogy, it is of great value to compare a human eye to an eye of something like a squid. By doing so, it is possible to see that there are alternative approaches to achieving the same ability (in this case, image-forming eyes). It becomes apparent that the vertebrate approach (with the neural wiring on the inside of the structure) is not the only way the organ can be designed. Studying the parallel evolution of communication in many different species can shed light on different approaches to or means of communication. By studying distantly related species whose last common ancestors were assuredly non-linguistic and most likely non-social, we can determine what defining features of communication are universally adaptive and which are mere historical contingencies. This interspecific comparison allows us to speculate on what the fundamental, adaptively significant features of communication may be and in doing so, learn more about the underlying structure and "purpose" of language.

---

[6] The examples concerning comparison of eye structures are taken from a personal communication with Matt Cartmill (2003)

Studying animal social systems can also lend credence to the theory that cognitive precursors made a sudden grammar explosion possible (Bickerton 1990). If the "catastrophic mutation" did not restructure thought, but merely made possible the expression of a previously internalized "grammar" of social expectations, then it is of great interest to study the cognitive abilities and grammatical competence of other highly social animals with complex brains, especially primates. If there was a selective advantage to ordering the world in terms of AGENT and TARGET, then perhaps a human-like syntactic ability would have evolved before communication improved at all. This could provide a theoretical intermediate between the big bang and the slow evolution theories. If cognition and argument mapping were selected for because of their social benefits, then a single mutation might have allowed the full and sudden expression of an underlying structure that was already there. Perhaps, without the full cognitive syntax (knowledge of third party relationships and cause and effect), such a "language" mutation would have had no selective advantage and would never have spread throughout the population.

Like Darwin's comparisons of animal eyes, a comparative study of animal cognition and linguistic acquisition might help shed light on how such a process could have occurred. If an animal that, in the wild, does not exhibit a complicated communication system can nonetheless acquire much of the structure and meaning generally assumed to be part of language, then it may already possess the ability to think in the same cognitive patterns as those required for basic language ordering.

We must make a distinction between studying naturally occurring communication systems and the investigation of linguistic ability in a laboratory setting. Each of the two approaches has its own merits and drawbacks. *Field studies* of natural communication systems provide a way to investigate alternative communicative strategies, but present many uncontrolled variables and difficulties in interpretation.

Conversely, *laboratory studies* in which animals are taught human-devised symbolic systems allow for precise control over independent variables and permit repetition of problematic situations to resolve problems of interpretation, but they do not provide insight into the conditions under which an animal's communicative abilities arose or the adaptive significance they may have in the natural environment.

In the terms laid down earlier, it can be said that field studies best exemplify the *form approach*, and laboratory studies best address the *function approach*.

The *form approach* is most useful in identifying different solutions to the problem of needing to share information with others in a group. By observing the communication of non-human social animals in the wild, we can often determine what things are most important for them to communicate or be aware of. This helps us understand the needs and pressures that drive and confine language evolution. For instance, there may be an advantage to giving loud calls to advertise your presence to potential mates, but such calls may be disadvantageous if predators are also alerted to your location. Such conflicting pressures typically lead to a system of calls that are "designed" for specific purposes – for example, a mating call that is easy to locate and a predator alarm call that is shorter and harder to trace (for field work on frogs see Ryan,1990).

Field studies on animal communication systems are often hindered by humans' limited perception of the animals' signals. It is impossible to know whether the animals are tuning into the same things that as we do as external observers. There are undoubtedly many visual cues in animal signaling, and these are often very hard for us to pick up on, because we are not innately sensitive to them, and are nearly impossible to quantify because they can not be analyzed as easily as acoustic signals. Also, where vocal signals are discretely bounded, many physical signals are not. How does one decide when something like posture has a communicative value? It is easy to measure periods of vocal communication, but much more difficult to determine where subtle physical cues start and stop.

Vocal signals are analyzed both through passive observation and playback experiments. The difficulty with these approaches is that we are forced to rely on the reaction of the listener to deduce meaning of a call. Many supposedly meaningful signals are assumed to be devoid of content if the receiver does not react in a predictable way each time the signal is transmitted. Playback experiments can control the temporal and acoustic parameters of a signal, but must still rely on the reaction of the listener to determine the call's semantic value.

It is difficult to ascertain how much information is encoded in the signals of a nonhuman species. All that can be observed is the call or signal of one animal and the following behavior exhibited in another. It is difficult to determine whether the signal elicits the behavior or even if there is a direct relationship between the two. If human conversations were analyzed in this way, one might assume that almost all human speech was contact calling, one person vocalizes, the other replies "here I am" and so on (Kiriazis 1997). Communication is not something that can be completely deciphered through a stimulus/response experiment. It is possible that although an animal might not react immediately to another's signal, it has received and processed the encoded information and will act upon it at a later time. It is important to keep in mind the limited ability to crack into a linguistic structure from the outside. Animals may encode information in ways we cannot perceive or about things we are not aware of.

The *function approach* (comparing the use or purpose of vocalizations) is used in laboratory studies in which animals are taught human-devised communication systems. Studies of this type tell us what animals are *capable* of, not what they in fact *do*. In the laboratory, experimenters can interact more directly with experimental animals and construct experiments to test specific questions. Such controlled studies are especially important in determining how different syntactic mechanisms may have been acquired. Determining the relative ease of acquiring linguistic properties that seem to be part of an inseparable unit can provide models of how they could have developed sequentially.

Artificially-created or natural human language can also provide insights into how different animals order and see the world. If animals are capable of using human symbols to communicate meaning, they are performing a task similar to human speech. By studying how exactly they operate within a matrix familiar to or created by us, we can determine *how* different species approach their environment, interactions with others, and (most importantly) communication. Such studies have had some significant successes, but are also limited in what they can measure and what they can ultimately show about other animals.

Laboratory studies have shown that some animals (apes, some cetaceans, and possibly parrots) can acquire and use abstract symbols. However, there is an ongoing debate over the extent of their demonstrated capabilities within the human mediums, and the degree to which their symbol production or understanding of utterances resembles language. One of the problems with using an artificially-taught language to gauge linguistic competence is that animals can always understand more than they can produce. Many animals can follow sometimes very complicated commands, but are unable to produce utterances that combine even a very few symbols. This is much the same as someone learning a second language would behave in response to a fluent speaker. In animals as in human subjects, language comprehension always exceeds production.

Another problem embedded in the nature of laboratory studies is that the animals are asked to do things very unnatural to them. The experimenters often teach animals labels for things that surround them in the lab because they are familiar. However, these experiments often ignore the things that may be of greatest interest to the animals in the wild: other individuals in their group. Most language-trained animals have been raised away from conspecific social groups so that they could be immersed in a linguistically rich environment. The experiments do not focus on what they would want to convey in their normal social groupings but instead force them to interact purely with humans and to reference objects that may be of no interest to them. There is no way to quantify how much this has under-represented the linguistic abilities of the animals tested because of their lack of personal investment in the experiments, but it must be assumed that it had some effect. It is also possible that many animals may have greater linguistic or cognitive ability than they demonstrate because they misinterpret the objective of the test.

One of the main criticisms of linguistic lab work is that it involves very small samples, usually a single animal. It is therefore necessary to ask if the animal tested is representative of its species of if it is an abnormally intelligent individual. It is also quite possible that by raising an animal in a linguistically-saturated human environment, the cognitive and linguistic abilities of the animal improve (Savage-Rumbaugh 1998). Raising an animal around humans may also have adverse affects because it will usually be socially deprived and human-imprinted and, as a consequence, will show very different behavior from normally socialized animals of the species.

Though many experiments of non-human linguistic competence have been done in both the lab and field, it is very hard to reconcile the work done in the two domains. The cognitive science/linguistics approach taken in the lab causes experimenters to dig themselves further into minute points about syntax and what exactly the animals can do without prompting, and rarely addresses the question of how the animals came to have these cognitive abilities. The field scientists, conversely, are concerned with broader concepts of evolution, but rarely manage to run a complete study on social interaction or vocalization.

IV: FUNDAMENTAL STUCTURES OF LANGUAGE

> 'Animals don't always speak with their mouths' said the parrot in a high voice, raising her eyebrows. 'They talk with their ears, with their feet, with their tails - with everything. Sometimes, they don't **want** to make a noise.'
> — Hugh Lofting, *The Story of Doctor Dolittle* (1920)

It is difficult to test for "language ability" in other animals, because it is rarely a question of simple presence vs. absence. Clearly, no other animal exhibits a system of signals bearing any significant structural or functional similarity to human language. However, animals are capable of communicating many things, and in some cases, they can learn to interpret or even generate signals that exhibit some of the design features of language. In order to compare language ability or complexity of communication systems between species, it is therefore necessary to break human language down into its component features and to test for the presence of each independently. This division typically begins with the distinction of semantics from syntax. For a communication system to be comparable and potentially antecedent to a full language, it should manifest semantics, syntax, or both (Greenfield et al 1990 and Kako 1999).

SEMANTICS

Semantically meaningful signals are perhaps the most basic component of language. Semantic meaning is not restricted to acoustic signals. Even human languages can utilize exclusively visual or even tactile channels, as in the cases of writing or signing. Meaning can be determined from animal signals of other sorts, such as body posture or pheromones, although few of them would be considered "semantic" by linguists.

Communicative signals can be broken down into three basic types[7] depending on the relationship between the sign and its referent (O'Grady 1997:591):

1) **Iconic Signs** directly resemble their referent. A school crossing sign is iconic as are onomatopoeic words like *buzz, crack* or *woof* in that they replicate the sound the sounds they represent. An open mouthed threat in dogs or other animals can also be classified as iconic because it resembles the act of biting.
2) **Indexical Signs** spontaneously rise from the presence of their referent. It is often said that there is a causal link between these signs and their referents. They usually serve to alert the listener to the internal or emotional life of the signaler. A cry of pain or growling stomach are both indexical signs, as are many of the sounds made spontaneously during mating or fighting. Indexical signs are not considered to be controlled by the signaler since they arise spontaneously and are not deliberately chosen.
3) **Symbolic Signs** are arbitrarily related to their referents. This category is distinct from the other two in that it requires that the signs be abstract and consciously chosen. Most human words are symbolic, as are some animal calls and signals such as the dances of honeybees.

Though human language is composed of all three types of signs, we are most concerned with symbolic signs or "symbols." The arbitrary association between words and meanings is what we mean when we refer to the "semantics" of a word. In English, there are several cases where words are linked to their referents, but are nonetheless considered semantic. This may well be the case because the words are not *necessarily* linked; they could be replaced by a new term if it was properly explained. Considering the

[7] The list and general descriptions is taken from O'Grady (1997).

abstract nature of human words, perhaps a better substitution for the term "semantic" would be "abstractly referential." For something to be abstractly referential it must 1) carry meaning, 2) be a "consciously" expressed symbol, and 3) not be *necessarily linked* to the referent (i.e. must not resemble the referent or be triggered subconsciously by the referent).

Animal signals are sometimes referred to as *functionally referential,* meaning that they serve as an indicator of a referent, but may not necessitate the conscious choice of sign that human words do. Thus, a functionally referential signal may carry information to a listener, but be a subconscious reaction on the part of the signaler. These signals can be said to have semantic meaning, but should not be classified in the same way as human words.

The way in which individual words are classified in the lexicon is also generally treated as an aspect of semantics. In all human languages, words with specific meanings are grouped together into hierarchically-nested levels of larger categories. For example, The word "oak" has a specific referent, but it also carries the information "tree," then "plants," then "living things," and then finally "nouns." Each word need not directly refer to all of its super-categories, but in some way it represents them. The phenomenon of categorization is extremely important to the structure of language, but is not necessary for a symbol to be abstractly referential. It is possible for a symbol to represent a single occurrence or object in the real world that does not carry categorical information. Proper nouns are referential symbols of this sort. Although they refer to a single entity rather than a class, they are nonetheless arbitrary symbols, and contain semantic information.

SYNTAX

Syntax is loosely defined as the structure that governs how symbols are expressed. Syntax shapes the meaning of strings of symbols through word order or by changing the symbols slightly based on their purpose in the particular utterance. In a natural language, the syntactical system encompasses all of the properties contained in Chomsky's theory of universal grammar. Although syntax (as universal grammar) is often seen as a single unit, in studying animal communication systems or evaluating learned linguistic ability, it is important that we break syntax down into its component functions or properties, so that we may look for the presence of each individually. Three principal properties of human syntax have become widely accepted as the foundations of human linguistic structure (Greenfield 1990, Kako 1999).

1) **Discrete Combinatorics**: symbols can be combined to create new words to refer to novel objects while still retaining their original meanings. For example: a *catfish* is a fish that has some cat-like characteristics, not an animal halfway between a cat and a fish.
2) **Category-based rules**: words can be broken down into categories governed by certain linking rules. For instance, nouns can take determiners, while verbs can have subjects and sometimes direct and/or indirect objects. The existence of such rules is what lets you know whether a sentence sounds as though it could exist in your language or not. Constructions that are grammatically possible but devoid of meaning will still follow category-based rules.
3) **Argument Structure**: utterances take on a certain form by combining symbols according to the categories they fall under. At the sentence level, this structure typically involves three such categories: AGENT (the thing that is acting),

THEME (what the agent is doing), and OBJECT (what the agent is acting upon). Expressions do not have to include all categories; they can imply arguments through context and do not have to take an object (as in the sentence *Fred* (AGENT) *walks* (THEME)). Argument structure is necessary for understanding *who* does *what* to *whom.*

Item 1, discrete combinatorics, is especially important because it allows a communication system to take on the form of an open class system. In order to achieve a structure close to that of human language, a communication system must make the shift from distinct holistic utterances to a system based on combining known units of meaning. This allows a limited lexicon to be "opened up," allowing for near-infinite combinations of meaning. Without unlimited memory for distinct symbols (which would be needed if every new situation were to be given its own label), the lexicon must consist of smaller units of meaning that may be combined in different ways to handle novel or previously unnamed items or events. It is important that these units of meaning retain their original meanings, because if meanings were combined when we combined symbols, then we would only be able to refer to new objects if they represented intermediate stages between previously-labeled items (Kako 1999). In a system based on discrete combinatorics, it would be possible to make the transition from every symbol referring to an entire referent to a syllable-based system where the smaller units of meaning (in our case, morphemes) cannot stand alone.

Although items 2 and 3, category-based rules and argument structure, seem remarkably similar, each can exist in the absence of the other. Category Based Rules are most often used to tell if a sentence "feels right." They determine which words or units of meaning can or are most likely to follow each other. It is possible that we use Category Based Rules to process language more quickly because when we hear a word in category X, our minds expect to hear a word in category Y, and so we often do not have to hear the next word in its entirety before we figure out the role it plays in constructing the meaning of the utterance. For example, if we hear a noun like *Jane* or *the dog*, our minds anticipate a verb, so if the next phoneme we hear is /w/, we then narrow down our expectancy to *walks* , *went* , *will,* and so on. This phenomenon probably plays an important role in language acquisition because it helps to narrow down the possible choices a child has in determining a word's meaning. If there were no way of narrowing down the possible meanings of a word from the entire lexicon, it would take children much longer to learn the meanings of words (Romero 2003).

It is also possible that Category Based Rules could be applied to non-semantic units. This may in fact occur in some non-linguistic communication. Some birdsong follows strict rules that require certain types of sounds to be followed by other types of sounds. For example, in order to be accepted by a female, a male's song may have to follow a pattern of *whistle-trill-trill-whistle-chirp.* If the song varies from the pattern, a female may not recognize it and therefore ignore the male's advances (Searcy and Andersson 1986). This type of birdsong follows the same organizational pattern as that of Category Based Rules, but should probably be considered to be a phonetic ordering rather than an ordering of categories unless the different song units can be determined to have different meanings.

Argument structure looks at the underlying roles different words play within an utterance. It relies on word order (or inflections) to determine the role a word plays in a sentence. Argument structure centers around an action and requires a knowledge of the participants in the action and their respective roles with relation to the action. Edward Kako presents a more linguistically detailed description: "To possess argument structure, a linguistic animal must know (1) how many participants are involved in the event that a verb labels and, therefore, how many *arguments* (in the logical sense) are associated with the verb and (2) how those arguments should be assigned to syntactic positions in the sentence" (Kako 1999: 2). However, I believe that it is possible to possess a non-linguistic argument structure by understanding the thematically-predetermined roles different individuals play in events without vocally labeling them. If an animal can understand the actors in an event and behave differently towards them in the future depending on the roles they played in the interaction, it might be said that the animal has an underlying understanding of argument structure. Similarly, if an individual can remember "debts" in a structured social system (who owes whom a favor), they must have some way of categorizing individuals according to the roles they played in past interactions.

It is possible for semantics to exist in the absence of syntax, as in the utterances of very young children or other cases of modern protolanguage (Bickerton 1990). Syntax can also link together symbols from the lexicon or made-up words so that they carry no meaning as a phrase, but still be a "grammatically correct" sentence. This is demonstrated in Chomsky's famous sentence *Colorless green ideas sleep furiously.* It is also possible for ordered structure like syntax to act on meaningless units of sound if vocalizations are arranged according to regular rules. As noted above, this may be the case with the songs of some birds (Searcy and Andersson 1986). However, syntax without semantic content should not be directly compared to language ability, since there is no communicative advantage to having syntax in the absence of semantics as there is in using meaningful but unordered symbols.

SET-BUILDING

Semantics and syntax are the commonly accepted building blocks of human language. However, they are not the only linguistic or cognitive components that contribute to language. Reusable phonemes, theory of mind, and the construction of sets all play essential roles in our everyday communication.

Phonemic distinction is an important evolutionary step towards achieving a duality of patterning and allowing the lexicon to grow even when the vocal space has been filled. This may be necessary for achieving a linguistic system with fully human capabilities, but it is not a necessary precondition for a system to be compared to language since, according to our previous definition, a language could be structured with holistic morphemes that did not reuse smaller units of sound. In a non-human communication system, it is often impossible to distinguish phonemic units from morphemic ones. It is difficult to determine "meaningful" units of sound from purely acoustic ones in a non-human system. Also, it may be equally challenging to assess the vocal capabilities of other animals, whether the absence of certain sounds is due to physiological limitation, and how sounds are parsed.

Theory of mind is an indispensable part of human language. We are driven to communicate things because we want to share knowledge or experience with others. If we had no idea how the experiences of others differ from our own, we would have a much less clear idea of what needs to be communicated in a particular situation. It is important to realize the distinction between a "full" theory of mind and one that controls self-awareness and association of self with other humans. A "full" theory of mind is not ~~a~~ necessary for language. Traditional theory of mind tests require predicting the responses of others who are ignorant in situations where the test subject is not. Children under the age of 6, as well as many partially autistic adults, fail these sorts of strict tests yet do not exhibit diminished language ability (O'Grady 1997). It nevertheless seems reasonable to think that self-awareness and awareness of the presence of others must be inherently linked to language. One must be capable of synthesizing aspects of one's environment and have a desire to share them with others in order to acquire personal gain or confer advantage onto another who will reciprocate in the future.

The construction of sets of objects is usually considered to be part of semantics, but I propose that it be listed as a separate category. Many human determiners and constructions rely on the ordering or comparison of sets and would be impossible without an ability to construct sets. For example, it would be impossible to say *more X than Y* without the sets *X* and *Y*.

Syntax can be applied to labels that do not contain information about sets. Take the sentence *Mara hit Connor*. This is both a semantically and syntactically viable sentence, and yet does not necessitate the ability to construct sets. There does not need to be a set of hitters or of people whom Mara hit. Thus, set-building should be seen as an ability separate from that of attaching arbitrary (non-iconic) labels and structuring utterances.

On the other hand, set-building does not presuppose linguistic ability. It is possible to have a pre-linguistic ability to classify objects into sets. On the simplest level, set-building is nothing more than the ability to abstract the sign used to label an individual and use it to apply to all similar individuals. This ability is shown by all animals that give alarm calls. If a chicken had to learn that each individual bird of prey fell into the category *predator*, its chances of survival would be greatly reduced. It must have the ability to abstract the characteristics of one known predator to apply to all similar individuals. It is the same for learning what foods to eat. One experience with a poisonous fruit should be enough to avoid all similar fruit in the future — essentially, to avoid the new set [*poisonous fruit X*]. This level of set-building is important, but it is a fairly simple process that can be based on associative learning and the physical similarity of all objects in a set. It is more difficult is to create higher-order sets such as the superset *bird* as compared to more concrete and specific sets such as *robin, nuthatch* or *ostrich. Bird* is harder to understand because its members do not have as many characteristics in common, and the set of all birds therefore contains many individuals that look very dissimilar. Creating super-sets is closely related to understanding relationships between sets or in "nesting" sets.

The ability to nest sets is necessary for labeling relationships between objects. To have a concept of the relationship "mother," a social animal must first be able to discern the sets of "females with young" and "juveniles," and then derive from them a super-set comprised of all females who have given birth and juveniles who are related through the relationship "mother". The relationship "mother" would be understood only if the individual could process MOTHER ([adult females with young], [juveniles]). It may be possible to label the relationship between two individuals from the observation of a single mother and child, but to abstract the relationship to other dyads it is necessary to possess the ability to nest sets.

There is some evidence that abilities of this sort can exist pre-linguistically. Laboratory experiments have shown that longtail macaques are able to recognize mother/offspring dyads out of a set of paired individuals from their social group (Dasser 1988). If it is possible to possess the cognitive ability to create and nest sets of individuals without having linguistic labels for them, then there is support for the theory that the underlying structure of language evolved before the ability to verbalize.

In order to evaluate the linguistic abilities of animals and to directly compare naturally-occurring communication systems to language, we must look for presence of each of the three components of language 1) *semantic content*, 2) *syntax*, 3) *nested-sets* independently. While all of these elements can exist without the other two, it is important to remember that in each case ability must be quantified on a graded scale, not simply scored qualitatively for presence or absence.

V: SEMANTIC STUDIES

> Lack of evidence for truly referential communication in animals is most likely a consequence of *our* incompetence, rather than that of the animals.
>
> — Irene Pepperberg, *the Alex Studies* (1999)

As noted in Chapter 3, the study of meaning in animal communication can be approached through either field studies or laboratory work. Each approach has advantages and limitations. Studying naturally-occurring systems in the wild allows direct comparison to human language as a communication system and provides insight into how animal communication may have evolved into something more complex. However, the meaning of animal signals observed in the field cannot be deciphered by a human observer unless they elicit responses from listening conspecifics. On the other hand, while laboratory studies may be able to determine the meanings of signals that rarely elicit responses, they may also underestimate or overestimate the communicative abilities of a species, since all such experiments rely on very few (usually just one) subjects and bring the animals up in a human-dominated, language-rich environment.

FIELD WORK

The ability to transmit information about the environment allows an individual animal to notify other members of its social group of things they cannot themselves perceive. This description might suggest that communication ought to involve an underlying theory of mind — that communication in itself entails that animals realize other individuals have different perceptions of the world and therefore must be informed of events or objects they cannot directly perceive. However, abstract referential signaling can easily exist without a theory of mind. For example, honeybees (which surely lack a theory of mind) use highly developed abstract signaling systems to communicate the location of food (O'Grady 1997). Semantic vocal symbols such as alarm calls also do not necessitate a desire to "share" information, but could have evolved from indexical signs that were nothing more than screams of fear. They could, over time, have adapted to become predator-specific by conferring more of an advantage to the listeners when the appropriate anti-predator response was elicited.

The problem with the theory of slowly-evolving social communication of this type is that it relies on some sort of group selection. An individual giving an alarm call puts itself at greater risk of predation by calling attention to itself. This fact should mean that individuals who do not vocalize but nonetheless are warned of the presence of a predator would have an advantage over those giving alarm calls. In order for alarm calling to be an evolutionary stable strategy, there must be a way to prevent cheating. Since there are no known cases of animals reprimanding non-calling individuals, there must be an alternative explanation.

Consistent, repetitive signals may have evolved as a way to reduce confusion on the part of the receiver (Cullen 1966). Such reduction of ambiguity is always beneficial. In giving threatening versus subordinate displays, an individual gains a significant fitness benefit by having the recipient correctly "read" its signal and decide not to attack. Krebs and Davies (1993) suggest that signals are specialized pieces of behavior that can be used to manipulate another individual to the signaler's advantage. In their model, the manipulated individual would (over time) develop counter-signaling to avoid exploitation and a complex system of regulated signals would evolve. This model, however, cannot account for signals such as food or predator calls, which do not give the signaler any particular advantage. Such signals provide benefits to the group while possibly incurring costs to the signaler that has to share its food or draw attention to itself in the presence of a predator. While the theories of the evolution of repetitive signaling may be useful in explaining threatening displays or courtship, they are not flexible enough to address the emergence of complex social communication.

It is possible that group selection did play a role in evolving *sender-disadvantageous* signals by favoring those groups whose members produced these calls, thereby increasing the fitness of all the members of the group and out-competing non-calling groups. Kin selection could also have led to seemingly selfless behavior if the larger groups were made up of family units so that every individual in the group would have a greater incentive to alert others nearby.

CHICKENS

Referential communication is of enormous importance for an animal trying to locate a limited food source or to alert others to the presence of a predator approaching from a certain direction or at a certain rate. It has been shown that some primates respond differently to alarm calls given in the presence of different predators (Seyfarth et al 1980). It is, however, difficult to determine (from observing the behavior of the listener) whether the calls distinguish between types of predator or between modes of attack. For example, chickens were once thought to have had alarm calls that specified predator class — an *aerial-predator* call (used for raptors) and a *terrestrial-predator* call (used for foxes and raccoons); but it was later shown that what the animals were communicating was the distance to or the direction towards the attacker (Gyger et al. 1987). When videos of foxes were shown as if these predators were flying overhead, chickens gave the *aerial-predator* call and the whole group crouched down and looked upward (a typical *aerial-predator* response). These results showed that the chickens' alarm calls were specifying the mode of attack, not the class of predator. Encoding information about the direction of attack is likely to be equally successful (if not more successful) in eliciting the appropriate anti-predator response. However, when analyzed linguistically, referring to direction of attack proves to be semantically simpler because it does not necessitate the chickens' having the ability to label a class of predator (i.e., create a set). The chickens' "attack from above!" alarm call is *sender-dependent* — that is, it relies on the perspective of the caller rather than on a common label for a set containing predators of a certain type. In the absence of common labels for sets of objects, there cannot be labels for the relations between the objects (requiring an ability to nest sets). And unless there are semantic labels for both objects and relations, there can be nothing like human language (O'Grady 1997).

MONKEYS

Predator-specific alarm calls are given by several species of primate (Seyfarth et al. 1980, Zuberühler 2000). Klaus Zuberühler demonstrated the presence of true referential labeling in Diana monkeys (Zuberühler 2000). These monkeys are preyed on most frequently by leopards and crowned hawk-eagles. The two types of predators elicit different alarm calls from the monkeys and consequently different evasive behaviors in the troop. In order to determine whether Diana monkeys were communicating the type of predator present or merely a spatial and temporal estimation of the approaching attack, Zuberühler carried out a playback experiment on 23 groups of monkeys in the Taï National Park on the Ivory Coast. To discern if the alarm calling was referencing either distance or elevation, audio speakers were placed at different heights in trees at different distances from troops of Diana monkeys and used to play back either leopard growls or the shrieks of crowned hawk-eagles. Zuberühler found that the appropriate alarm calls were given to the different predators and that neither height nor distance of predator had a significant effect on call structure.

Zuberühler (2003) conducted a second experiment to determine whether the monkeys were capable of ascribing the same semantic meaning to more than one acoustic stimulus—that is, whether the "leopard" alarm call carried the same semantic value as the vocalizations of the leopard itself in the minds of the monkeys.

The alarm calls of Diana monkeys differ by gender. By playing a male's alarm call to groups of females, Zuberühler was able to elicit the analogous alarm call from the females. He used a habituation technique to test semantic value through a series of playback experiments. Three protocols were used: 1) a recording of a leopard growl followed by that of another leopard growl, 2) leopard growl followed by the male monkey's "leopard" alarm call, 3) leopard growl followed by the male's "eagle" alarm call. Zuberühler reasoned that the females should exhibit the same pattern of habituation (weaker response to the second stimulus) as they would for the control situation in which the predator vocalization was repeated, if they perceived the predator's vocalization and the male's "leopard" call has having equivalent meaning (Zuberühler 2003).

Zuberühler found that if a leopard's growl was played followed by a male 's "leopard" alarm call, the females had only a weak response to the male's call. However, when the leopard growl was followed by a male 's "eagle" alarm call, the second stimulus elicited just as strong a response as the first. The results show that Diana monkeys consistently use semantic vocal labels to signify predator category. This implies some ability to construct semantic sets.

While chickens need to have two semantic labels, attack from above and attack from below, their call structure does not require that they have a preexisting categorization of predators. Each predation event is looked at and labeled on the basis of the experience and individual viewpoint of the sender. By contrast, the alarm calls of Diana monkeys are predator-specific; they do not change depending on the direction of the referent; and they can be grouped with predator vocalizations to create sets of signs carrying the same semantic value. Diana monkeys therefore possess a cognitive concept of predator class as well as an ability to categorize acoustic stimuli.

PRAIRIE DOGS

Even if the calls of animals can be shown to carry semantic information, it can be difficult to determine whether or not listeners receive and process all of the information transmitted. It is possible that the ability to encode certain information in signals did not evolve as a communication system, but appeared merely as a side effect of some other trait that was selected for. This possibility can be seen in the seemingly over-complex information encoded in the calls of certain prairie dogs. These ground squirrels have been hunted by humans for over two centuries and consistently give alarm calls at the sight of humans. C. N. Slobodchikoff et al. (1991) performed an experiment on two colonies of Gunnison's prairie dogs, *Cynomys gunnisoni,* using human "predators" of different sexes dressed in different colors to see if the resulting alarm calls would code for the characteristics of the individual testers. During each trial, the human "predator" walked a prescribed route through the colony while a second tester tape-recorded the resulting alarm calls from a short distance away. Three separate protocols were tested. In the first experiment, four testers of equal height, two male and two female, wore matching white laboratory coats and sunglasses and each walked through the colony three times. The goal was to determine if the animals could differentiate between the individual testers even though they were visually similar. The second experiment used the same four testers, but each wore a shirt of a different color (in colors that could be distinguished by prairie dogs, which are red-green colorblind). The third experiment used only two testers (one male and one female) who each walked through the colony 10 times, half of the time wearing a white shirt and the other half a yellow one. This experiment was also analyzed to see if individual testers could be recognized even though their appearances were similar. The alarm calls were compared along 10 different acoustic parameters, and all calls were compared to the mean. Though success differed somewhat between the three tests, the prairie dogs gave the "appropriate" call 94% of the time; that is, the investigators could determine with 94% accuracy, from the acoustic structure of the calls alone, which person had walked through the prairie-dog colony and what color of shirt that person was wearing.

The results of the study suggest that Gunnison's prairie dogs have the ability to distinguish and transmit semantic information about the physical features (including color and aspects of shape) of individual predators. The calls also show some evidence of syntax in the broad sense, because the vocal structure through which they label individual testers is consistent and predictable; they are using a set structure to convey meaning. It remains to be determined whether the prairie dogs are processing the information encoded in their alarm calls. It may be that this information is superfluous; however, an ability to distinguish between individual predators as well as types of predator may very well be advantageous to these rodents. Since prairie dog colonies remain in a fixed location, they are likely to come into contact with the same predators repeatedly. Individual predators undoubtedly have variations in hunting style and skill, and so an alarm call that warns the colony of "the dog that digs into burrows" instead of "the dog that just sniffs around" could in principle be useful in avoiding predation.

Slobodchikoff's work shows that while all the alarm barks of prairie dogs sound similar to the human ear, they carry far more information than originally thought. While it remains to be shown that the prairie dogs themselves register and use all of the information encoded in the calls, this study serves as a warning against the tendency to assume that if humans cannot hear evidence of call variations then they do not exist. The study of wild chimpanzees conducted by Adam Clark Arcadi (2000) exemplifies both the problem of human listeners' aural limitations and the enormous discrepancy between the linguistic abilities chimpanzees can acquire when taught and that which they exhibit naturally.

CHIMPANZEES

Chimpanzees and their close relatives, bonobos, are often the subjects of linguistic and cognitive studies in laboratory settings. It is tempting to assume that their performance on language and intelligence tests will provide insight into how early human ancestors may have acted. However, although controlled tests in which signed and artificial languages are taught to apes have shown that some individual apes can acquire a linguistic system comparable to that of a two-and-a-half-year-old child (Savage-Rumbaugh 1998), there is no evidence that wild chimpanzees possess a communication system that is any more complex than those of other primates.

To determine whether or not wild chimpanzees conveyed information in their vocalizations, Adam Clark Arcadi conducted a field observation of a community of animals in Kibale National Park in Uganda (Arcadi 2000). Limiting his analysis to the vocalizations of males (who call more frequently than females), Arcadi recorded 83 hours of male calling over a period of two months. By assigning each call made to one of sixteen categories and labeling it as either submissive or non-submissive, Arcadi was able to compare types of calls and responses between individuals. In comparing calling rates, percentage of call responses and percentage of calls that "echoed" the initial call, Arcadi found that most vocalizations were caller-initiated or "spontaneous," and that individuals responded to only about 10% of calls. Most importantly, of the response calls that were made, the majority were of the same type as the call heard. This last finding implies that chimpanzees are merely imitating the call heard previously and not *replying* in the human sense of the word.

Since other primate species are known to "countersing" with their neighbors or "duet" with their mates (Deputte 1982), the results of Arcadi's study imply that chimpanzee calls are not more complex than those of other primates though their ability to acquire language would suggest otherwise (Savage-Rumbaugh 1998, Gardner 1969). Arcadi admits that some of the subtle complexities of the calling may have been lost due to lumping several calls into one category, or may simply not have been discernable by the human ear as acoustically different calls. However, if we assume that the study was at least approximately representative of the nature of wild chimpanzee vocalizations, it shows that while chimpanzees may possess the underlying capability to use language, they do not develop or use it naturally.

These studies show that there is a huge diversity in the use of meaningful signals employed in natural populations. Though some species are capable of labeling predator class or encoding specific details about individual predators, animals in the wild do not exhibit anywhere near the degree of linguistic proficiency they can acquire under artificial enriched conditions.

LABORATORY STUDIES

Work with "talking" animals has been done for centuries. However, many of the early successes demonstrated not true linguistic competence, but the animals' ability to detect subtle cueing from their trainers. This left the scientific community skeptical of all language-training in animals.

Skeptics of linguistic studies done with animals claim that successful animal/human communication is nothing more than subconscious cueing on the part of the trainer. This false but outwardly impressive ability of animals to use language has acquired the moniker the "Clever Hans Phenomenon;" named for one of the most famous fakes. Clever Hans was a trotting horse owned by a German mathematics teacher, von Osten, at the turn of the century. Von Osten spent years trying to train Hans to understand written and spoken German, solve simple arithmetic problems, and distinguish dissonant from consonant musical chords (Cartmill, 1990). Hans would answer by nodding his head, tapping a hoof, pointing with his nose, or picking up an object in his teeth. He performed amazingly well, responding correctly almost every time provided the experimenter was a familiar trainer. "Clever" Hans was observed by a panel of several experts in psychology and animal training. These experts concluded that Hans was not receiving cueing, but could not figure out how a horse could possess such amazing abilities. Finally, the young psychologist Oskar Pfungst solved the mystery. "By varying one condition of the testing situation after another, Pfungst was able to establish that Hans's trainer and questioners were indeed giving Hans visual cues that guided his responses, but that these cues were so subtle that only Hans had noticed them" (Cartmill 1990: 179).

The trainer, von Osten, would unconsciously move his head or tense his muscles almost imperceptibly when Hans neared the correct answer. This movement was too slight to be detected by a human observer, but was all the horse needed to determine when to stop pawing the floor or nodding his head (O'Grady, 1997).

This discovery cast a shadow over all legitimate animal intelligence experiments, forever affecting the interpretation of results of animal language studies. Although there has been amazing progress in the last 10 years with the design and use of artificial languages and with structuring experiments to better test semantic and categorical information, no study has been fully accepted by the scientific community. Most are quick to criticize or doubt even convincing work in the field because there have been many fakes in the past, and because many animal-owners use anecdotal evidence to over-estimate language comprehension and usage.

VIKI

The first structured study to determine if an animal could acquire human language was undertaken in the 1950s by the Hayeses, who adopted an infant chimp named Viki and raised her alongside their own child. They hoped that if brought up as a human, Viki would acquire language. Although Viki swiftly exceeded the abilities of the human child in both agility and dexterity, she managed to acquire only four words (Hayes and Hayes 1951). Her inability to make human speech sounds was thought, at the time, to indicate chimpanzees' inability to learn language. It was not until the late 1960s that a second attempt was made to teach an ape language. This time American Sign Language [ASL] was used as the linguistic medium, and the results proved that it was the species' vocal limitations, not their cognitive ones, that had caused the failure of the Hayses' experiment.

WASHOE

Allen and Beatrice Gardner acquired an infant female chimpanzee, Washoe, and attempted to teach her sign language (Linden 1970). They believed strongly that chimpanzees could use their manual dexterity to communicate with humans if given a chance. The Gardners were aware that "even caged, laboratory chimpanzees develop begging and similar gestures spontaneously, while individuals that have had extensive contact with human beings have displayed an even wider variety of communicative gestures" (Gardner and Gardner 1969: 665). Since sign language is a fully structured human language, they hoped that the chimp's progress could be compared directly to that of human children. They hoped that Washoe's ability to acquire the combination of iconic and abstract signs in ASL would provide insight into the referential capabilities of chimpanzees.

The Gardners raised Washoe in an environment similar to that of human children. While her trainers were able to attend to Washoe only in shifts, they attempted to interact with her as parents would with a human child and to communicate with Washoe and each other exclusively in ASL. The goal was to expose Washoe to a linguistically rich environment, much "as one chatters at a human infant in the course of a day" (Gardner and Gardner 1969: 666). The only sounds that her trainers made were non-words such as laughter and grunts that could be easily imitated by a chimpanzee.

Washoe's trainers used instrumental conditioning and molding as their primary teaching tools: shaping Washoe's hands into the shape desired for the sign and then demonstrating the sign themselves. Over the course of the study, Washoe learned 55 separate signs, most of them labels for objects in her environment. She also picked up at least two signs (*flower* and *toothbrush*) through pure and delayed imitation (Gardner and Gardner 1969). These signs were not initially elicited from Washoe but were used by her trainers during her daily routine. She began using both signs spontaneously to ask for the objects. This, the Gardner's claimed, was evidence that she could use signs purely out of a desire to communicate (Gardner and Gardner 1969). Eventually, Washoe combined signs to communicate directed desires, as in *come here* and *tickle Washoe* (Linden 1970). The Gardners also claim that Washoe could use signs out of familiar contexts to be humorous or insulting. Anecdotes testified to Washoe's ability to "swear" by adding *dirty* as an adjective onto the name of a trainer she was mad at (Linden 1970).

Washoe was also capable of transferring her learned labels to novel objects in previously established noun classes. The Garders state that "when introducing new signs, [they] used a very specific referent for the initial training…Washoe has always been able to transfer her signs spontaneously to new members of each class of referents" (Gardner and Gardner 1969: 670). This fact is important in evaluating Washoe's abstract signaling capabilities. Because she can apply a semantic label to novel members of a referential class, she must have an ability both to create sets and to semantically label them. Washoe's ability to acquire signs is seemingly limited to objects, certain verbs, and a single proposition (*on*), but her capacity to acquire human semantic labels and to apply them to all members of a referent class place her use of semantics well within the basic framework of language.

The Gardner's work with Washoe renewed interest in the language capacity of apes. Their successes laid the groundwork for future linguistic studies that would not have to begin with proving that a non-human can acquire words, but could instead begin to assess the language capacity of apes in a structured way. The most successful of these studies has been Sue Savage-Rumbaugh's work with the bonobo, Kanzi.

KANZI

Kanzi's study is unique in the field of language-trained animals because he began to acquire language naturally without being actively taught, much in the same way that a human infant would. Savage-Rumbaugh used "Yerkish," a set of abstract symbols made up of referents, verbs, locations and so forth arranged on a computer keyboard, as the primary mode of instruction. She set out to teach Kanzi's adoptive mother, Matata, to communicate simple desires using these symbols, but met with little success. However, when Matata was sent away for breeding, Kanzi (then a young juvenile) approached the computer board and used the Yerkish symbol to ask for food (Savage-Rumbaugh, 1998). Kanzi had never been taught to use Yerkish, nor had he ever exhibited interest in it before. It was only with his foster mother gone that he was forced to interact directly with the keepers and accordingly changed his means of expression in order to do so. When Savage-Rumbaugh began actively training Kanzi, she found that he acquired new signs not through laborious reinforcement (as other chimps have), but through unforced "dialogue" with his trainers during which they used Yerkish symbols and spoken English simultaneously whenever possible (Kako 1999).

Kanzi acquired a vocabulary of over 300 signs and gestures. He uses 256 Yerkish symbols as well as manual signs, which he mostly uses for verbs. He often combines the two media, using the Yerkish symbol as the AGENT and the sign as the ACTION (Savage-Rumbaugh, 1998). Savage-Rumbaugh has tested Kanzi extensively on his ability to label objects and actions and on his understanding of complicated commands in English. More than 13,000 of Kanzi's "utterances" have been recorded and analyzed and there is no question that Kanzi understands and uses abstract referential signs. What is most remarkable about his linguistic ability is that it seemed to come almost naturally to him. No other animal has shown such a predilection for acquiring language. Although he may be an anomaly amongst apes, Kanzi's ability to acquire language out of a desire to communicate with his trainers (Savage-Rumbaugh, 1998) may provide great insight into the early evolution of language.

ALEX

The study of the communication of birds usually deals with either songs or alarm calls. Irene Pepperberg broke this pattern by attempting to teach human language to an African Grey parrot, Alex. Alex was taught by the model/rival method in which two trainers demonstrated the desired dialogue, one asking questions and rewarding the second with bits of food if she got it right. This showed Alex the testing process from the onset of his training and encouraged him to answer the questions and get the food by imitating the model.

Alex has been trained to answer questions about a set of small objects placed before him on a tray. By teaching him labels for shape, material, and color of the objects, Pepperberg hoped to learn more about the cognition underlying avian signaling. How would he categorize objects and for which aspects of the objects could he acquire labels? Alex has learned thirty "object labels," (key, nut, wood, shoulder, water, banana, etc.) and has the ability to use them in phases involving "no," and in expressions of desire: "want X." In addition, he can distinguish nouns from adjectives. He can utilize "attribute labels" such as colors, numbers, and shapes and uses numbers to distinguish objects by their number of corners (Pepperberg 1999).

There is no question that Alex has learned to label the objects with which he is presented. He can apply labels to the entire referent class even when an object is particularly worn or otherwise dissimilar from the typical referent— e.g., highly-chewed clothespins or many different types of paper (Pepperberg 1999). Alex often uses words without prompting, to communicate desires (*want X*) or babble to his trainers (*hello, goodnight, etc.*), but it is his performance in tests of label recognition and categorization that shows his true abilities.

Alex is tested through exercises in which he must label common traits of items or match the correct object label to the description he is given. This procedure is very much like similarity/dissimilarity exercises designed for young children (e.g., "one of these things is not like the other…"). Alex has shown an ability to both match an object with the description provided, and to determine what qualities several objects have in common or how they differ (Pepperberg 1999).

The main limitation in working with Alex lies not in what he is able to do, but what he is willing to do. He can be stubborn and reluctant to participate in the experiments, and at one point refused to work for almost two years.[8] Linguistic laboratory work with language-trained animals often encounters this problem. When an experiment hinges on the performance of a single individual, progress is inevitably determined by the animal's eagerness to perform.

AKE & PHOENIX

Although dolphins are closest to humans in brain/body ratio, are easily trainable, and are agreed to be highly intelligent, they were not initially used in attempts to teach animals language because their means of signal production are so different from our own. The early approach to language acquisition, in which chimps were raised as human infants, has no applicability to dolphins. Dolphins cannot form human words (as Alex can) and are certainly unable to use a human sign language (as Washoe did). To test the linguistic ability of dolphins, Louis Herman developed two training "languages," one employing an acoustic medium, the other a visual one. One dolphin, Phoenix, learned to respond to a communication system based on computer-generated acoustic clicks played from an underwater speaker. The other dolphin, Ake, learned a gestural language based on the arm and hand movements of a trainer standing at the edge of the tank (Herman 1984).

The dolphins were trained to associate their acoustic or visual symbols with referents consisting of objects that could be placed into their pool (*PIPE, BASKET, HOOP* etc.). They were also taught modifiers that could be used to specify object location (*SURFACE, BOTTOM, RIGHT, LEFT*) and actions that could be carried out with the objects (*OVER, UNDER, TOUCH, FETCH* etc.). They also learned the words *YES, NO, ERASE* (ignore previous command) and *QUESTION* (used to request information from the dolphin) (Kako 1999). Ake and Phoenix were presented with strings of signs that asked them to carry out different actions involving the objects in the pool. Some of their commands involved a single referent (*PIPE FETCH)*, while some incorporated both a direct and indirect object (*BASKET FETCH PIPE*—meaning take a/the pipe to a/the basket).

[8] Dorothy Cheney, personal communication.

The dolphins were able to successfully carry out most of the commands and were consistently able to correctly identify all objects that had labels in their "languages." What is even more impressive is that when the test sentence used an unfamiliar label, they chose an available novel object to pair it with, and one of the two (Ake) was able to report when the labeled object in the command was not present amongst the objects in the pool. Herman states: "These various abilities evidenced that the words of the languages had come to represent symbolically the objects and events referred to in the sentences," (Herman 1984: 129).

Several researchers have employed similar protocols in attempting to teach sea lions gestural language but have been less successful (Herman 1989, Gisiner 1992). These animals evince less ability than dolphins to acquire abstract labels and respond correctly to the use of novel labels. Gisiner noted that: "The sea lion's responses also showed that she made little, if any, use of logical or semantic properties of the signs" (Gisiner 1992:78). Herman concludes: "For the dolphin, but apparently not the sea lion, gestural signs were attached to concepts rather than to entities" (Herman 1989: 19).

The ability of primates, parrots and dolphins to acquire human semantic labeling is impressive. The question then remains: if the underlying cognitive ability to acquire more advanced communication exists in some nonhuman animals (especially in apes), why does it not appear naturally? It is possible that language-trained animals' cognitive abilities increase substantially as a result of living in a highly enriched, linguistic environment, and that "normal" animals never acquire these cognitive advantages. It could also be the case that normal, socialized animals develop non-verbal means of social communication the value of which we have yet to determine.

The results of the field studies on birds, primates and rodents paired with the results of language-training studies (Pepperberg 1999, Savage-Rumbaugh 1998), lend support to the theory that the underlying cognitive structures of language evolved first and only later acquired the link to production mechanisms (Jackendoff 2002). While it may be useful to regard these naturally occurring systems as "living fossils" of different stages in the evolution of language, it is important that they be approached not as *stages* of language, but as non-linguistic or pre-linguistic systems that primarily utilize different aspects of semantic signaling.

VI: SYNTACTIC STUDIES

> If the imitative tendency of the parrot could be coupled with the quality of intelligence of the chimpanzee, the latter undoubtedly could speak.
>
> — R. M. Yerkes (1925)

Syntax has often been claimed as the aspect of language that is truly unique to humans. The concept of universal human grammar first introduced by Chomsky (1957) is still prevalent. This theory discourages comparison of syntactic qualities in animal communication to the full syntax of human language because it looks at syntax as an all-or-nothing quality.

Laboratory studies with language-trained animals have attempted to assess the ability of certain animals to acquire the main aspects of human syntax (discrete combinatorics, category-based rules, argument structure). These results of these studies have not been as promising as those of experiments looking for abstract referential labeling. Most of the studies revolve around word order, and while some animals are able to comprehend connections between labels when they are posed in command form, the use of syntax in production has been almost negligible (Savage-Rumbaugh 1998, Kako, 1999). Even the most promising animal subjects fail to combine words or symbols with any regularity.

Several field studies have looked at syntax-like structures of repetition and unit ordering in natural communication systems. However, many of these focus on the structure of utterances that have not been shown to carry any particular semantic value.

It is interesting to note occurrences of structure in the absence of meaning, but difficult to abstract any relation to the evolution of communication, since non-referential signals are probably not being used to enhance communication.

A similar problem arises in studies where the calls encode information, but it is not clear if the listeners are receiving the information — for example, in Slobodchikoff's (1991) study of Gunnison's prairie dogs. In this particular case, the alarm calls were found to follow a certain pattern where different qualities of the physical description of the predator were encoded in different ways, according to temporal position or pitch. This regular structure could be compared to human sentence structure and word order because the different categories of descriptive information were always expressed in the same place (showing something like *category-based rules*), but referencing different things. If the animals are not using the information encoded in the calls, the ordered structure may simply be the result of another type of neural ordering that affects vocal production but has no real effect on comprehension or enhanced communication.

It is also nearly impossible to determine what structural aspects of a species' vocalizations are semiotic. Field studies focus more on determining what can be encoded in different calls than on how the information is structured. Due to the dearth of field studies dealing with meaningful syntax in natural call production, it is necessary to restrict the discussion of syntactic ability to the animals that have been introduced to human syntax in the laboratory.

Some aspects of syntax seem to be more easily acquired than others. Some limited competence in utilizing discrete combinatorics has been seen in several of the language-trained species, notably bonobos, parrots and dolphins (Kako 1999). These subjects have also demonstrated limited competence with category-based rules, though more in comprehending human commands than in production. The same can be said for tests involving argument structure; when following commands based on word order, the animals (Ake and Phoenix, Kanzi) perform well above chance, but are not capable of employing this type of structure in production (Kako, 1999).

There is at least one other element of language structure that has been systematically tested in animals, that of syllable construction. The evolutionary step to breaking holistic utterances (*words*) down into reusable syllables is a difficult one. It requires generativity (seeing utterances as combinatorial—much the same as Hockett's (1964) duality of patterning) (Jackendoff 2002). Jackendoff stresses the importance of the emergence of syllables in the evolution of language, identifying syllable use as a distinctly human trait. He states that ape language experiments have not tested this type of generativity; there has been no systematic approach to combining defining characteristics of semantic symbols. The Yerkish symbols used by Kanzi are atomistic and represent holistic utterances. Also, Jackendoff notes, new signs in ASL are taught to apes as separate gestures rather than by using defining characteristics of place handshape and movement (Jackendoff 2002).

However, teaching ASL by breaking it down into its syllable-like or phoneme-like components , as Jackendoff proposes, seems unnatural. Deaf children are taught individual signs in the same way as hearing children are taught words: either by modeling or by direct instruction. Children are taught to read by breaking words down into syllables, but to demand a linguistic dissection of signed language seems an unfair requirement to impose on the ape language experiments.

There has been no combined approach by animal language researchers to address each of the categories of syntax similarly so that usage may be compared across species. Therefore, each experiment must be evaluated independently by a criterion that encompasses the different levels of syntax. Edward Kako (1999) examined the syntax use of Kanzi, Alex, and the dolphins Ake and Phoenix according to his three components of syntax: discrete combinatorics, category-based rules, argument structure. The following discussion, while based on his comprehensive analysis, also addresses experiments not covered in his paper and includes perspectives from each of the researchers involved in the subsequent projects.

WASHOE

Washoe could produce only single signs or simple two-sign combinations. However, when her progress was compared to the early language development of children, her two-word combinations matched those of two-year-olds. The Gardners believed that if the word order of Washoe's signed utterances showed that she was addressing new referents by combining known signs or using the relation between the signs to carry semantic meaning, then Washoe was indeed demonstrating discrete combinatorics and perhaps category-based rules as well (Gardner and Gardner 1969). For example, if "Washoe tickle" could be shown to mean something different from "tickle Washoe," then a certain level of syntactical ability could be shown to exist. Washoe often signed "me tickle" before tickling one of her trainers or friends, but this is not enough evidence to prove that she understood the meaning of the word order. Even when she began to make three-word sentences and could distinguish between "you tickle me" and "I tickle you," there was nothing to prove that she did not merely learn to associate those two structures with different scenarios. She did not carry this argument structure over into situations that required different verbs or referents.

The sentence structure of Washoe's utterances was compared to that of human children through the data used by Roger Brown in his study of child vs. chimpanzee word acquisition. Brown created extensive tables of word order and association, hypothesizing that the relationships between words in two-word phrases were independent of the specific words in the phrase. This implies that even very young children's utterances evince an understanding of linguistic structure. Brown found that 75% of human children's combinations followed his model; the Gardeners found that 78% of Washoe's constructions followed it as well. "Brown contended that the *child's* preference for certain word orders is indeed evidence of an emerging syntax" (Linden 1981: 46). This conclusion should transfer over onto Washoe's equivalent productions; however, the Gardners were hesitant to make any conclusions about Washoe's syntactical abilities because of potential flaws in their experimental design. The goal of their experiment had not been to analyze Washoe's syntax, but to see whether a chimpanzee could acquire human words (Gardner and Gardner 1969).

KANZI

Kanzi has shown a remarkable ability to distinguish different meanings conveyed by different word orders in spoken commands. His level of comprehension is equal to, if not above, that of a two-and-a-half-year-old child (Savege-Rumbaugh 1998). Kanzi has been tested quite vigorously on his comprehension of ordinary, spoken English, and has done remarkably well. To make sure that he understood the *connections* between words and was not merely responding to object labels alone, Savage-Rumbaugh tested him according to two main protocols.

First, Kanzi was given a variety of novel sentences and question forms involving complicated constructions or intransitive verbs that required Kanzi to be able to acquire information from the structure of the sentence, rather than just from the object labels themselves, in order to respond correctly when novel structures or labels were substituted (Kako, 1999). Previously-unseen combinations such as "Put the money on the mushrooms," and different question words like "could you" or "can you" were used. Kanzi's performance was tested against that of a two-year-old girl named Alia. Kanzi responded correctly on 74% of trials, Alia on 65%.

In the second experiment, Kanzi was tested on a series of sentences aimed at determining whether or not he possessed an argument structure. He was given paired sentences where the direct object and indirect object labels were switched. This protocol required that that Kanzi determine the goal of the sentence from the position of the word in the sentence, not the object label itself. Commands such as "Put your *hat* on the *ball*" vs. "put the *ball* on the *hat*" were used as well as sentences like "take the *rock outdoors*" vs. "get the *rock* that's *outdoors*" (Kako 1999). Although Kanzi was not as successful with this experimental design as he had been on some of the previous experiments, he preformed well above chance and certainly possesses an ability to determine thematic role from position in the sentence.

While Kanzi evidently possesses *discrete combinatorics* and arguably *argument structure* (Kako 1999), there is less convincing evidence surrounding his possession of *category-based rules*. The problem again lies in the fact that Kanzi seems capable of understanding far more complicated structure than he is able to produce. He never combines more than three words at a time, and never more than two lexigrams; therefore, it is hard to test him on the structure of his "utterances." Savage-Rumbaugh (1998) claims that Kanzi has created his own regular word order, thereby showing that he is using *category-based rules*. However, the order that Kanzi employs is not *category-based*, but *modality-based.* In his two-word combinations, Kanzi prefers to use both a gesture and a lexigram and always places the lexigram first. The creation of this rule shows that Kanzi does not employ the category of the word (object, verb, etc.) into the structure of his utterances. Even if he understands that role can change based on position in an English sentence (as the second experiment above leads us to believe), in his own utterances the rule "lexigram first" overrules any knowledge of *category-based rules* (Kako, 1999).

ALEX

Alex employs categorization of nouns as well as discrete combinatorics in his use of object descriptions. He knows that an object can be described by its color or its number of corners or both together, and will correctly label objects when asked about their color or shape. Alex's use of descriptions such as "four-cornered blue paper" or "three-cornered yellow hide" show that he is able to combine different words without losing any of their original meaning.

It is especially interesting that Alex is able to correct himself when he "misnames" an object (Pepperberg 1999). When presented with a piece of pink paper (a novel object), Alex called it "rose hide." However, after he was able to chew on it, Alex changed his prior categorization and labeled it "rose paper." Alex thus understands which word is linked to which physical property. This knowledge of the relation between word position and type of physical characteristic shows that that Alex has a sense of category-based rules (Kako 1999). Alex also uses two verb expressions that likewise demonstrate this knowledge. His utterances "wanna go X" and "wanna Y" are always followed with the appropriate class of noun (Kako 1999). Alex fills in X and Y by choosing from the appropriate subset of his lexicon; X can only be a *location*, and Y can only be an *object*.

Alex definitely possesses *discrete combinatorics* (he labels objects by their multiple properties in novel combinations) and *category-based rules* (he labels objects *adjective* + *noun*, and uses *want* + *object*, and *go* + *location*). However, though his correct use of "wanna go X" vs. "wanna Y" shows that he understands the limitations the structure places on X and Y, Alex does not know enough verbs to prove that he understands argument structure (Kako 1999).

AKE & PHOENIX

Ake and Phoenix arguably have shown a more advanced understanding of syntactic rules than any of the other language-trained animals. They certainly possess comprehension of discrete combinatorics. Their ability to use modifiers for object labels such as "RIGHT PIPE" as well as their processing of commands where an action is combined with a noun (and each retains its original meaning) show that each label retains its original meaning (Kako 1999).

More significantly, Ake and Phoenix are able to respond accurately to commands based on word order and argument class (Kako 1999). In order to test the dolphins' ability to understand word order and classification of labels, Herman presented them with many differently-structured commands. He concluded that "...comprehension, at levels far above chance, was shown for all of the sentence forms and sentence meanings that could be generated by the lexicon and the set of syntactic rules, and included the understanding of: (a) lexically novel sentences; (b) structurally novel sentences; (c) semantically reversible sentences that expressed relationships between objects; and (d) sentences in which changes in modifier position changed sentence meaning," (Herman, 1984: 129).

The dolphins' performance on the test situations involving familiar structures with novel labels and familiar labels arranged into novel configurations was above chance. Kako write that "this suggests that their knowledge of structure is to some extent based on *categories* of words —objects, actions, modifiers —rather than on relations among specific items" (Kako 1999: 8).

Ake's knowledge of argument structure was tested by the insertion of extraneous verbs into her commands. Her responses to these tests were unpredictable; she would usually chose to ignore one of the arguments and fall back on the familiar command structure but she also occasionally ignored the command. Herman believed that if Ake recognizes the anomalies in the sentences, she should refuse to carry out any portion of the command (Herman 1993). However, it is unreasonable to ask a trained animal (who is rewarded for successful trials) to refuse to act on a command. I believe that Ake's tendency to "repair" the ungrammatical commands by eliminating one of the objects or verbs—rather than, say, retrieving both objects or executing both verbs—shows that she understands that verbs can be assigned only two objects.

Ake and Phoenix's ability to recognize a connection between thematic role and word order shows that they can recognize argument structure. In tests involving a transport object and a destination object that could ostensibly be reversed, they performed the correct transportation 99% of the time. Kako notes "their remarkable accuracy suggests that they understood the links between syntactic position and thematic role," (Kako 1999: 9).

All of the language-trained animals described above have shown an ability to acquire some aspects of human syntax. They are able to combine (or react to the combination of) discrete symbols to encode information about the relationship between the referents. They also possess an ability to extract thematic role or classification of an object from its place in a sentence (though, admittedly, to differing degrees). It is reasonable to suspect that this capacity for structured utterances is employed in some way in their natural communication systems.

It is possible that there is much more order in natural communication than we are able to detect. Perhaps visual signals are incorporated into the structure of vocalizations, or perhaps the animals are cueing into some acoustic signal we deem insignificant. It is also possible that this ability to acquire linguistic structure comes not from the ordering of communicative signals, but from a pre-linguistic cognitive framework.

VII: SET-BUILDING AND ARGUMENT MAPPING

> The building of alliances takes time, and exploiting them to the full takes more time. It places a heavy load on the memory, too, if you have to remember over weeks and months who you owe and who owes you.
> — Derek Bickerton, (2001)

It is possible that the abilities of language-trained animals to categorize referents by creating mental sets and to utilize argument structure is the product of a pre-linguistic cognitive ability to order and group objects. As was seen in Chapter Two, both Calvin and Bickerton (2001) and Jackendoff (2002) believe that grammatical utterances could have emerged from a pre-linguistic structure through a fortuitous mutation that allowed the structure to be applied to vocalizations.

Bickerton (2001) believes that syntax derives from a need to assign and remember roles in interactions grounded in reciprocal altruism. As Calvin and Bickerton put it, "The practice of reciprocal altruism created the set of abstract categories and structures that, once they were joined to a structureless protolanguage, yielded the kind of syntax that all modern human languages exhibit" (Calvin and Bickerton 2001: 126).

This mental ordering need not be based purely on reciprocal altruism in order to be beneficial. It would convey a similar advantage in a hierarchical system in which individuals find it useful to be able to remember whom they have interacted with and in what capacity. This would make it easier to explain the cognitive ordering abilities of animals such as chimpanzees, which have been shown to possess some syntactic capabilities but do not exhibit reciprocal altruism. In a social system where resources are occasionally shared voluntarily or affiliative behaviors such as grooming are employed, an individual animal can increase its personal success if it does not allow other individuals to exploit its time or resources. In such a society, the majority of the individuals would benefit from a mechanism to detect cheating: " Animals with partners who don't cheat will do better (suffer less stress, win more fights, gain more sexual access, and thus probably leave more progeny behind) than animals whose partners cheat and get away with it," (Calvin and Bickerton 2001: 128).

There would also be an advantage to recognizing relationships between other individuals in one's social group. For example, if an individual engaged in an aggressive interaction with another individual who had alliances with other members of the group, that individual would be well served if it possessed an ability to avoid the "friends" of its opponent directly following the altercation. One could also increase one's standing in the group by choosing alliance partners based on whom they were "close" to. Humans make these sorts of "networking" decisions on a daily basis, but it may be possible for pre-linguistic minds to employ similar judgment processes in their social interactions.

This process of strategically choosing partners based on their standing or their history of interaction with other individuals would require some form of basic argument mapping (knowing *who* did *what* to *whom*) as well as recognizing the nature of the relationships between members of the group. Correctly "reading" social relationships and standing within a group allows an individual to avoid costly aggressive encounters, much as displays and tension-releasing affiliative behaviors afford a means of avoiding (or repairing) aggressive encounters that may threaten social relationships (Cheney, Seyfarth and Smuts, 1986).

Social relationships in nonhuman primate societies are usually based on one of three factors: kinship, relative dominance, or friendship (Tomasello and Call, 1997). Temporary coalitions may also be formed to circumvent a high-ranking individual or to gain access to a prized resource (Tomasello and Call, 1997). This phenomenon allows coalitions of low-ranking individuals to gain access to resources they would have been denied individually. The formation of coalitions requires an advanced knowledge of the social structure of the group. In order for an individual to form an alliance with one animal in order to compete with another, it must know not only whether the other individuals are lower or higher-ranking than it, but also the rank of the other animals in relation to each other. It must possess an awareness of so-called "third-party relationships" (Tomasello and Call, 1997).

Possessing an understanding of the nature of relationships between other individuals in a group would allow an animal to exploit the social structure of the group for personal advancement and gain. Such an understanding would facilitate decisions about which individuals to groom, solicit or challenge based on whom they were "related" to and therefore who else would be affected by the behavior (Tomasello and Call, 1997). With this sort of third-party knowledge, an individual would also have the ability to transfer aggression or solicitation to the close relatives of an animal with whom they had had a recent encounter (Tomasello and Call, 1997). These behaviors, along with the underlying social awareness that they presuppose, imply thought processes of a degree of complexity often attributed only to humans, in which the animals must learn relationships by merely watching, and not experiencing, encounters between individuals.

BABOONS

Several field studies have tested monkeys' abilities to recognize third party relationships. Cheney and Seyfarth (1998) used a series of playback experiments to test female baboons' ability to perceive the social relationships of other individuals. They tested several pairs of unrelated adult females by playing calls of disputes (sound sequences containing an aggressive grunt from one animal and a scream from another) and measuring the reactions of the two listeners. Dyads were tested on playbacks where one of three situations obtained: 1) both of the recorded individuals were related to the females in the listening dyad, 2) the one giving the aggressive grunt was related to one of the listening females, or 3) neither one was related to the females. The listening females were videotaped during and after the playback. Their response to the taped disputes was determined by whether and how long they looked towards the speaker and whether or not one of them looked at the other (the speaker was positioned so that they could not do both simultaneously). Twenty-six female dyads were tested, and their behavior was tracked for fifteen minutes following the playback to see whether any sort of reconciliation or redirected aggression took place. Cheney and Seyfarth found that the females looked at the speaker more often and for longer periods of time when the calls involved kin, and that the subordinate female related to the caller that screamed looked proportionally longer. The females also seemed to have an awareness of the relationship between the caller and their dyad partner. Dominant subjects looked at the other female more often during the trials involving two kin callers, and subordinate females looked at the other female for a shorter period of time when the callers were both non-kin. Dominant females were also more likely to supplant the other dyad female when the playback involved two kin.

These results indicate that female baboons have the ability to vocally identify individuals in their groups and recognize the relationships between them. They have the ability to take information from the social content (aggressor and victim) of the call sequence and then to use that information to determine the recipient and nature of their following behaviors. When a dominant female hears a dominant grunt from a relative followed by a submissive scream from another individual, she is likely to supplant a relative of the submissive caller. In contrast, the relative of the submissive caller is unlikely to approach a close relative of their kin's aggressor following playback. This behavior suggests complex behavior prediction based on third-party social awareness.

Anecdotal data suggest that primates have a much greater social awareness than had previously been suspected. Frans de Waal (1998) describes numerous events within a captive chimpanzee colony that suggest that the individuals have mastered a complex and ever-changing social environment. He describes events in which third parties intervene to solve disputes or reconcile conflicts between two individuals — for example, when one female "kidnapped" an infant and its mother displayed angrily until a third female retrieved the infant and returned it to its mother. His examples also nicely illustrate the social perception that strategic alliance-forming requires. De Waal describes a change in the male dominance hierarchy resulting from a takeover. The lowest-ranking individual had an alliance with the young mid-ranking male; but as soon as the young male took over, the low-ranking male began to make friendship displays to the previously-dominant male. By forming a friendship with the less dominant of the two, the low-ranking male was allowed to mate occasionally. However, the low-ranking male had to be aware of the shift in power in order to consciously sever one friendship and attempt to forge another.

In order to recognize relationships between third parties, individuals must possess transitive reasoning capabilities based on the ability to group (i.e., create mental sets of) other individuals. It is theoretically possible that, through a long period of trial and error based on direct interaction, an animal could determine "relatedness" between two individuals by engaging in a dispute with one and having the other intervene each time. However, this would be impractical and costly, and the reinforcement of the second individual's approach might occur only rarely. It is more plausible to assume that such individuals possess a transitive cognitive ability.

MACAQUES

Dasser (1988) tested captive longtail macaques for this sort of abstract understanding of social relationships and transitive reasoning. He showed them photographs of pairs of individuals from their social group and rewarded them for choosing the picture of the one mother-offspring pair out of the set. After the animals were habituated to choosing the related pair, Dasser tested them by asking them to choose between another mother-offspring pair and an unrelated pair. He also performed a similar procedure in which the animals were asked instead to choose a photograph of the correct offspring when presented with a photograph of a mother. He found that many macaques were able to perform both tasks reliably, although some individuals could not master the test and so were excluded from the study.

The ability of the macaques to correctly choose novel mother-offspring pairings shows an abstract understanding of the social relationship "mother-child" independent of the individuals involved. It could be the case that this understanding is a template for "relatedness" that allows the mental insertion of individuals and thus the ability to create sets of individuals categorized by their relation to one another.

PIGEONS

This capacity for seeing abstract relationships between members of a set seems to be a highly advanced cognitive ability close to human thought processes. However, this transitive reasoning ability seems to not be limited to primates. Von Fersen, Wynne and Delius (1991) performed a transitive inference test using pigeons. They trained pigeons on several pairs of adjacent symbols from a set ranked in the order A>B>C>D>E, and conditioned them to choose the "highest-ranking" symbol each time. They then tested them on a novel combination B+D to see if they had understood the abstract hierarchy of the set. Reliably, the pigeons would choose B. This would make it seem as if the pigeons understood the relative rank of the letters within the set. However, the birds could have been using "value transfer" instead of true transitive inference. An understanding of the set based on the principle of value transfer would say that since B is associated with A in the conditioning, it must be high-ranking, and since D is often associated with E (which is the lowest-ranking letter), it must be low-ranking. This approach to the set does not require the same sort of abstract relationship between B and D that an understanding of the overall hierarchy of the set does. The pigeons do have some degree of hierarchy perception, because when the set is made circular (through the insertion of the conditioned pair E>A), they do not choose B more often than D.

Many social vertebrates seem to have some awareness of social relationships—they are able to recognize individuals visually or acoustically and will often adjust their behavior based on their past history with that individual (for examples involving fur seals see Insley, 2000; for Siamese fighting fish see Whitfield, 2002). However, primates seem unique in their ability to see abstract patterns of relatedness applied to an entire set of objects. There are animals who have shown remarkable abilities at handling shared characteristics within sets (e.g. Alex: Pepperberg1999) and many species that can determine social relationships and dominance when they are engaged in first-person interactions (Shettleworth1998).

There is some degree of social perception required for all animals that live in groups. However, a true awareness of third-party relationships requires the ability to create sets of individuals (based on their association) and some form of argument mapping (to understand how those associated with an individual will respond related to whether the individual is the "winner" or "loser" of an encounter). A cognitive ordering of this type would allow as well as be driven by complex social phenomena not based on kinship (open-choice coalitions, friendship).

VIII: CONCLUSIONS

While we can be fairly certain that no other species possesses a language ability equal to that of humans, we are unable to identify a single characteristic of language that sets us apart. Natural communication systems of widely differing types (avian, rodent, cetacean and primate) employ functional referential signaling. Some species, like the Diana monkey, show evidence of abstract referential signaling and ability to group acoustic signals into classes (Zuberühler 2003). Studies with language-trained animals have proved that some species possess the ability to acquire human semantic labels and use them to reference both single objects and object classes (Savage-Rumbaugh 1998, Herman 1984).

Several aspects of syntactic structure (discrete combinatorics, category-based rules and argument structure), thought to be limited to human language, have likewise been employed by language-trained animals (Kako 1999). Although evidence of these structures has yet to be found in natural communication systems, studies with Kanzi, Alex, and Ake and Phoenix demonstrate that the underlying cognitive patterning necessary for such structures exists in a wide range of species.

Studies involving recognition of third-party relationships as well as evidence of redirected aggression and aided reconciliation suggest that incipient forms of the grouping, set-building and argument-mapping abilities underlying language exist in the pre-linguistic mind (Cheney and Seyfarth 1998, Zuberbühler 2003). These pre-linguistic cognitive-ordering abilities may be the underlying structures that later emerged as syntactic ordering in early language.

Evidence pointing to the recent emergence of language (within the last 500,000 years) (Enard et al. 2002), suggests that many of the mental-ordering capabilities employed by language evolved before significant improvements in communication appeared. The advantages to possessing a capacity for categorization and argument mapping are manifest in complex social systems. The fitness benefit to being able to avoid cheaters and construct socially-savvy alliances would allow these mental-ordering abilities to radiate throughout a population.

While a slow evolution of language dependent on intra-group communication may not be an evolutionary stable strategy, the evolution of cheater detection and recognition third-party relationships is. By employing pre-linguistic mental-ordering, it would be possible for grammar to "appear" with few communicative antecedents. The catastrophic mutation that allowed the expression of grammar would not have had to restructure many systems simultaneously. It would have merely provided pre-existing structure to be expressed in the vocalizations of early hominids. The theories of a stepped evolution of language punctuated by a period of proto-language (Bickerton 2001, Jackendoff 2002) support this theory because they rely on a pre-linguistic syntactic structure. Humans may be the only species to possess language, but the underlying structures of language are phylogenetically much older.

ACKNOWLEDGEMENTS

I am indebted to Drs. Robert Seyfarth and Dorothy Cheney for sharing with me their time, advice and resources as well as for inspiring me to pursue questions they themselves have struggled with. I would also like to deeply thank my thesis advisor, Dr. Eric Raimy, for his guidance and continued help with the process.

I would like to extend my deepest gratitude to my father, Dr. Matt Cartmill, for his advice, counsel, undying support and editorial contribution to this work. I would also like to thank my mother, Dr. Kaye Brown, for her guidance and support given continuously throughout the process.

Last, I must thank Sydney Beveridge and Mara Gustafson for their commiseration and motivation.

REFERENCES

Arcadi, A. C. (2000) Vocal responsiveness in male wild chimpanzees: implications for the evolution of language. *J. of Human Evolution*. **39**: 205-223

Bickerton, D. (1998) Catastrophic evolution: the case for a single step from protolanguage to full human language. Chapter in *Approaches to the Evolution of Language*. Hurford, J., Studdert-Kennedy, M., and Knight, C. [Eds.]. Cambridge, England: Cambridge University Press.

Bickerton, D., (1990) *Language and Species*. Chicago: University of Chicago Press.

Bickerton, D., (1981) *Roots of Language*. Ann Arbor, Michigan: Karoma

Brainard, M. S. and Doupe, A. J. (2002) What songbirds teach us about learning. *Nature*. **417**: 351-358

Calvin, W. and Bickerton, D. (2001) *Lingua ex Machina: Reconciling Darwin and Chomsky with the Human Brain*. Cambridge, Massachusetts: The MIT Press.

Cartmill, M. (1990) Human uniqueness and theoretical content in paleoanthropology. *International Journal of Primatology*. **11**: 173-192

Cheney , D.L. and Seyfarth, R.M. (1998) Recognition of other individuals' social relationships by female baboons. *Animal Behaviour,* **58**: 67-75

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

Cheney , D. L., Seyfarth, R. M. and Smuts, B. (1986) Social relationships and social cognition in nonhuman primates. *Science*. **234**: 1361-1366

Chomsky, N. (1957) *Syntactic Structures*. The Hague: Mouton

Cullen, J. M. (1966) Reduction of ambiguity through ritualization. *Phil. Trans. R. Soc. B* **251**: 363-374 (as cited in Krebs and Davies 1993)

Darwin, C. (1859) *The Origin of Species*. Reprinted. London: Penguin.

Dasser, V. (1988) A social concept in java monkeys. *Animal Behaviour*. **36**: 225-230

Davidson, I. and Noble, W. (1993) Tools and language in human evolution. Chapter in *Tools, Language and Cognition in Human Evolution*. Gibson, Kathleen R., and Ingold, Tim [Eds.]. Cambridge, England: Cambridge University Press. pp 363-388

de Waal, F. (1998) *Chimpanzee Politics*. Maryland: Johns Hopkins University Press.

Deputte, B. L. (1982) Duetting in male and female songs of the white-checked gibbon (*Hylobates concolor leucogenys)* In (C. T. Snowdon, C. H. Brown and M. R. Petersen, Eds) *Primate Communication,* pp. 67-93. Cambridge: Cambridge University Press.

Enard, W., Przeworski, M., Fisher, S. Lai, C. S., Wiebe, V., Kitano, T., Monaco, A. P., Pääbo, S., (2002) Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* advance online publication, 14 August 2002.

Gardner, R. A., and Gardner, B. T. (1969) Teaching sign language to a chimpanzee. *Science*. **165**: 664-672

Givón T. (1995) *Functionalism and Grammar*. Philadelphia: John Benjamins. As cited in Jackendoff (2002)

Gisiner, R. and Schusterman, R. J. (1992) Sequence, syntax, and semantics: responses of a language-trained sea lion (Zalophus californianus) to novel sign combinations. *Journal of Comparative Psychology.* **106**: 78-92

Goldin-Meadow, S. and Mylander, C. (1983) Gestural communication in deaf children: noneffect of parental input on language development. *Science.* **221**: 372-374

Greenfield, P. M., and Savage-Rumbaugh, E. S. (1990) Grammatical combinations in *Pan paniscus*: Process of learning and invention in the evolution and development of language. In S. T. Parker and K. R. Gibson (Eds.), *"Language" and intelligence in monkeys and apes: Comparative developmental perspectives* (pp. 540-578) New York: Cambridge University Press.

Gyger, M., Marler, P. and Picket, R. (1987) Semantics of an avian alarm call system: the male domestic fowl, *(G. domesticus) Behaviour*. **102**: 15-40

Hauser, M. (1997) *The Evolution of Communication*. Cambridge, Massachusetts: The MIT Press.

Hayes, K. J. and Hayes, C. (1951) *Proc. Amer. Phil. Soc.* **95**: 105

Haynes, W. O. (1998) *Communication Development: Foundations, Process, and Clinical Applications*. Baltimore: Wilkins and Wilkins

Herman, L., Kuczaj, S., and Holder, M. (1993) Responses to analogous gestural sequences by a language-trained dolphin: evidence for processing of semantic relations and syntactic information. *Journal of Experimental Psychology: General*. **22**: 184-194.

Herman, L., Richards, D. and Wolz, J. (1984) Comprehension of sentences by bottlenosed dolphins. *Cognition*. **16**: 129-219

Herman, L. (1989) In which Procrustean bed does the sea lion sleep tonight? *The Psychological Record*. **39**: 19

Hockett, C. F. and Ascher, R. (1964) The Human Revolution. *American Scientist.* **52**: 70-92.

Jackendoff, R. (2002) *Foundations of Language: Brain, Meaning, Grammar, Evolution.* New York: Oxford University Press.

Kako, E. (1999) Elements of syntax in the systems of three language-trained animals. *Animal-Learning-and-Behavior.* **27**: 1-14

Klein, R. G. (1989) *The Human Career: human biological and cultural origins.* Chicago: University of Chicago Press.

Klein, W. and Perdue C. (1997) The Basic Variety, or: Couldn't Language be Much Simpler? *Second Language Research* **13**: 301-47

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

Krebs, J. R. and Davies N. B. (1993) *An Introduction to Behavioral Ecology.* Cambridge, MA: Blackwell Science Ltd.

Levelt, C. (1994) *The Acquisition of Place.* Leiden: Holland Institute of Generative Linguistics. As cited in Jackendoff (2002)

Linden, E. (1981) *Apes, Men, and Language.* New York: Penguin Books.

Lieberman, P. (1991) *Uniquely Human.* Cambridge, MA: Harvard University Press.

Lofting, H. (1920) *The Story of Doctor Dolittle.* 1988 edition. New York, Dell Publishing.

Lorenz, K. (1952) *King Solomon's Ring.* New York: Thomas Y. Crowell Company.

McMillan, C. (2003) The neural basis for semantic memory: introduction and generalized quantifiers. Lecture. University of Pennsylvania. April 16, 2003.

Mowrer, O. H. (1960) *Learning Theory and the Learning Process.* New York: Wiley.

O'Grady, W., Dobrovolsky, M., Arnoff, M. (1997) *Contemporary Linguistics: An Introduction.* Boston: Bedford/St. Martin's.

Pepperberg, I. M. (1999) *The Alex Studies: Cognitive and Communicative Abilities of Grey Parrots.* Cambridge: Harvard University Press.

Pepperberg, I. M. (1991) Referential Communication with an African Grey Parrot. *Harvard Graduate Society Newsletter*: Spring 1991

Pinker, S. and Bloom P. (1990) Natural Language and Natural Selection. *Behavioral and Brain Sciences* **13**: 707-26

Romero, M. I., (2003) Lecture. University of Pennsylvania. February, 2003.

Ryan, M. J. (1990) Signals, species, and sexual selection. *American Scientist.* **78**:46-51

Savage-Rumbaugh, Sue, and Shanker, Stuart G. Taylor, T. J. (1998) *Apes, Language and the Human Mind*. New York: Oxford University Press Inc.

Searcy, W. A., Andersson, M. (1986) Sexual selection and the evolution of song. *Annual Revue of Ecological Systems*. **17**: 507-533

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

Skinner, B. F. (1957) *Verbal Behavior*. New York: Appleton-Century Crafts.

Slobodchikoff, C. N., Kiraizis, J., Fischer, C. and Creef, E. (1991) Semantic information distinguishing individual predators in the alarm calls of Gunnison's prairie dogs. *Animal Behaviour*. **42**: 713-719

St. Clair, R., (1973) Language in thinking. *Foundations of Language*. **12**. 459-461

Tomasello, M. and Call, J. (1997) *Primate Cognition*. New York: Oxford University Press.

Toulmin S. (1972) *Human Understanding*, vol. I. Oxford: Clarendon Press. As cited in Jackendoff (2002)

von Fersen, L., Wynne, C.D., Delius, J.D. and Staddon J.E.R. (1991) Transitive inference formation in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*. **17**: 334-341. As cited in. Shettleworth, S.J. (1998) *Cognition, Evolution and Behaviour*. New York: Oxford University Press.

Whitfield, J. (2002) Nosy Neighbors. *Nature*. **419**: 242-243

Yerkes, R. M. and Learned B. W. (1925) *Chimpanzee Intelligence and its Vocal Expression*. Baltimore: William and Wilkins.

Zuberühler, K. (2000) Referential labeling in Diana monkeys. *Animal Behavior*. **59**: 917-927

Zuberühler, K. (2003) Referential signaling in non-human primates: cognitive precursors and limitations for the evolution of language. *Advances in the study of behavior*. In Press