SEMANTIC STRUCTURE AND THE CONSEQUENCES OF COMPLEXITY: THE EVOLUTIONARY EMERGENCE OF GRAMMAR

by

Joe D. McFall

(Under the Direction of Brent Berlin)

ABSTRACT

Many linguists advocate the existence of an innate neurobiological, genetically encoded module for processing grammatical structure. This conclusion is reached by assuming such a structure to be the only possible explanation for the apparent high rate of language acquisition in an environment of impoverished linguistic input during development, the existence of universal linguistic features, and the infinitely productive, recursive nature of syntax (i.e., digital infinity). This paper will explore the evidence for which such a biological module is considered an explanatory logical necessity, and attempt to offer a possible alternate explanation for the same evidence that does not rely on an assumed genetically-encoded module. It is suggested that an evolutionary model of language supports a more parsimonious alternate explanation for universal features and the recursive nature of syntax, and as such is preferable to the nativist position.

INDEX WORDS: Language evolution, Generative syntax, Memetics, Cognitive evolution, Computational models

SEMANTIC STRUCTURE AND THE CONSEQUENCES OF COMPLEXITY: THE EVOLUTIONARY EMERGENCE OF GRAMMAR

by

Joe D. McFall

B.A., Emory University, 1998

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF ARTS

ATHENS, GA

2004

SEMANTIC STRUCTURE AND THE CONSEQUENCES OF COMPLEXITY: THE EVOLUTIONARY EMERGENCE OF GRAMMAR

by

Joe D. McFall

Major Professor: Brent Berlin

Committee: Zachary Estes
Paula Schwanenflugel

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2004

TABLE OF CONTENTS

CHAPTER 1

INNATENESS

Every route has a starting point and an ending point. To chart the course of the evolution of language, at least two questions must be asked. The first question is: when did the capacity for language begin to emerge in our species? This question tells us our starting point. The second question is perhaps more important, if not more difficult to answer, than the first: what is the ending point? In other words, where are we now? In order to answer the first question, we must be able to propose an appropriate model of language as it is used in our species. To do so, we need to answer the second question. It seems that one way to explore the question of language as-it-is is to return to the first question. So, then, it seems that these two questions must be answered simultaneously with complementary answers. Of course, the internal coherence that we find in this complementarity must be taken as a whole pattern and matched to data from descriptions of communication in other species, language as a complex, external, sociocultural phenomenon, language as an internal, cognitive phenomenon, and the basic tenets of biological and cultural evolution. A theory of language evolution must be an internally coherent description, and must fit with what we know.

I will attempt to show that an extreme innatist point of view, such as Chomsky's, commits one to a "crypto-creationist" account of language evolution. A crypto-creationist account of language evolution would propose that human language seemed to appear fully-formed at some point in our ancestral history, whether by macromutation or miracle, or both. While Chomsky has only recently softened this point of view, Pinker has attempted to reconcile an innatist position with an account of language evolution that assumes evolutionary continuity without miraculous macromutations (Pinker and Bloom 1990; Pinker 1994). Pinker asks how a hard-wired, syntactic module might have evolved in the brain. I will ask whether it is necessary to posit such a module as Pinker describes, and whether such an account fits the known data.

As an innatist viewpoint of language evolution starts with the presupposition that language is a hard-wired module, let us first analyze the argument in support of this position. The innatist claim stems from a set of premises about the nature of language and language acquisition, as follows:

1) Rate of language acquisition – children acquire language completely, at a very high rate
2) Age-dependence – Adults have a more difficult time acquiring a second language than do children. While children acquire the language of their surrounding community fluently, even if the language is not their native one, adults who learn a second language rarely become fluent. There seems to be a critical age at which one can fully acquire a language.
3) Poverty of stimulus – A child seems to acquire general syntactic rules of a language from inadequate linguistic stimuli.
4) Convergence of grammars within a linguistic community – language users within a linguistic community show little grammatical variance.
5) Language universals – Unrelated languages seem to exhibit common structural features

Notice, first, that none of these premises offered in support of a hard-wired Universal Grammar offer evidence regarding the human brain itself[i]; rather, the argument for innateness is supposed to follow logically and necessarily from these premises. The structure of the argument is such that a Universal Grammar is considered by those who accept the argument as the only way to explain all of these alleged features of language and its acquisition. If one were to find an alternate model of language, as it relates to the human brain, that would explain each of these features, then the argument for innateness would be weakened in that the primary assumption of the argument is that no other explanation is possible; if such an alternate explanatory model offers a better fit with the known data, then certainly such a model would be preferable.

While each of the above premises will be dealt with further in this paper, Sampson (1997) attempts to refute each of them individually. Regarding premise (1), Sampson questions Chomsky's notion of what constitutes speed of acquisition. Children acquire language at a high rate, but relative to what? Sampson distinguishes between two of Chomsky's claims regarding the rate of language acquisition. Chomsky's usual argument is that the rate of language acquisition is *absolutely* fast, whereas he occasionally argues that language is *relatively* fast (Sampson 1997: 35). A claim of absolute speed, of course, is irrelevant unless accompanied by a prediction about the expected rate of language acquisition were language not hard-wired. In other words, even a claim of an *absolutely* high rate must be relativized. Chomsky offers no such comparison. Rather, where Chomsky claims a *relatively* fast acquisition of language, he

[i] Controversy surrounds the recent discovery of FOXP2, a gene that has been called 'the language gene' by popular press reports. While this gene has been associated with a kind of specific language impairment occurring in the lineage of an English family, some evidence shows that their impairment may not be specific to language, but associated, at least in part, with more general, non-linguistic cognitive functioning. For example, the phenotype associated with a mutation at FOXP2 seems to cause a lower *non-verbal* IQ in those afflicted compared to unafflicted members of the family. See Vargha-Khadem et al. (1995) for details.

tends to use spurious analogies such as that of language acquisition as compared to learning physics. For example, Chomsky writes:

> Grammar…[is] acquired by virtually everyone, effortlessly, rapidly, in a uniform manner…Knowledge of physics, on the other hand, is acquired selectively and often painfully, through generations of labor and careful experiment, with the intervention of individual genius and generally through careful instruction…(Chomsky 1976, quoted in Sampson 1997: 27)

This argument by analogy is most clearly flawed in the sense that the two systems being compared are ultimately not comparable. Our understanding of language upon acquisition is likely a kind of unconscious knowledge; we are able to use language in ways that can be judged grammatical by other speakers within our linguistic communities without explicit knowledge of alleged rules underlying the structure of our speech. In this case, Chomsky is comparing unconscious knowledge of a system (i.e. language) to explicit knowledge of a system (i.e. physics). A more appropriate analogy to this unconscious knowledge in the realm of physics might be the ability to, for example, throw a ball at a target; it is not necessary to take a course in Newtonian physics to be able to hit the bulls eye. The ability to throw a ball at a target is a skill that is learned, like language, through observation, practice and trial-and-error; it is not an innate ability. So in terms of a relatively fast rate of acquisition, this analogy, as Chomsky presents it, does not hold.

The argument for the age-dependence of language acquisition can be considered equally vacuous, if only because it is possible that the ability to integrate and reproduce the structure of systems through learning declines with age, as well. Any evidence in support of an innate grammar that describes language acquisition, specifically, as dependent on age must be able to differentiate language learning from mechanisms of general learning. If this distinction is made, it must be shown that while the ability to acquire language does decline with increased age, general learning abilities do not. If evidence exists to support this notion, then that evidence supports the proposal that language can be understood as a domain-specific module; however, neither Chomsky nor others on whose work Chomsky bases his theory (e.g. Lenneburg 1967) offers evidence that would be incompatible with a non-nativist stance in this respect (Sampson 1997).

Perhaps the most crucial argument for an innate Universal Grammar is the poverty-of-stimulus argument. Chomsky suggests that "the nature of primary linguistic data…consists of a finite amount of information about sentences, which, furthermore, must be rather restricted in scope, considering the time limitations that are in effect, and fairly degenerate in quality." (1965:31). To paraphrase Chomsky, the linguistic data to which a child is exposed during language acquisition is not itself sufficient, because of the short duration of exposure and the low quality of the data, for the acquisition of full linguistic competence, without positing an innate device. While this argument may be the most crucial to the innatist position, it is also the most contested. One of the problems of the poverty of stimulus argument is that it is derived from mathematical formalisms. These formalisms are used, for example, to predict what kinds of errors children should logically make, assuming they use simple induction for acquisition, and assuming a relative lack of feedback for ungrammatical utterances. As Tomasello (1995) suggests, however, these hypothetical predictions about what kinds of errors a child should make are products of the theoretical paradigm that assumes an innate grammar module; if one assumes a less formalistic acquisition process, one based on "concrete words and semantically based classes of words" rather than abstract formalisms, then we would have no reason to hypothesize the errors that would 'logically' occur under the innatist paradigm (1995: 145). Furthermore, Pinker's (1994) claim that children do not receive 'negative evidence' (i.e. correction of ungrammatical utterances) does not seem to hold. Tomasello, for example, cites evidence (see Bohannon and Stanowicz 1989; Farrar 1992) that suggests that there are any number of ways that adults can use subtle cues to call a child's attention to ungrammatical utterances.

The possible existence of language universals is another crucial element of the innatist argument. Universals are generally considered *absolute* or *statistical*; that is, a universal characteristic is absolute if it is found in *all* languages, barring none, and is statistical if it is found in *most* languages. The existence of absolute language universals is a necessary property of an innate, genetic grammar module. Without universals, the innatist claim falls apart because this claim purports that aspects of syntax are genetically preprogrammed, biological traits, like arms, legs, lungs, etc. The problem is that while many statistical and implicational "universals" have been discovered, absolute grammatical universals seem to be so rare as to seem trivial. Even the best candidates for absolute universality, according to Pinker, the categories of noun and verb, are problematic in that some languages do not seem to use these strict grammatical categories (Tomasello 1995).

The point of the preceding paragraphs has not necessarily been to prove, with utmost finality, that the innatist position is incorrect. Rather, it has been to show that the innatist position, as proposed by Chomsky and Pinker, is based on assumptions that are at the very least weak, and at most indefensible. What matters most to the argument proposed in this paper, however, is that the innatist position commits one to an account of the evolution of language that does not seem to hold, especially in light of the flaws of the position itself.

Until fairly recently, Chomsky has held the position that language could not have evolved biologically by natural selection, and therefore must have emerged discontinuously as a sudden mutation. Hauser, Chomsky and Fitch (2002) modify that position significantly. Their essay, written in part to consummate the "interdisciplinary marriage…inaugurated over 50 years ago" between biology and linguistics (ibid.: 2), strips the essence of human language to one formal characteristic: the capacity for discrete infinity yielded by syntactic recursion. Discrete infinity is commonly defined as the capacity of a system to generate an infinite set from a finite set of elements. This, of course, is not a practical description of any human language, but rather an abstract, formal, theoretical description of all human language.

Hauser, et al. primarily consider their object of study to be internal language (I-language, i.e. language in the mind/brain), as opposed to language as it is used among individual speaker/listeners (i.e. external- or E-language). Furthermore, I-language, and more specifically the language faculty (FL), is understood in a both a broad and a narrow sense. The language faculty in the broad sense (FLB) consists of at least a sensory-motor system, a conceptual-intentional system and an internal computational system. The language faculty in the narrow sense (FLN) "is the abstract linguistic computational system alone, independent of the other systems with which it interacts and interfaces" (ibid.: 2). On this model, the FLN is responsible for syntactic recursion. Thus, for these authors, the core characteristic of language, in its most narrow sense, is its capacity for recursion. While this limitation on the definition of language might seem extraordinarily restrictive to both linguists and non-linguists alike, Hauser et al. isolate this formal property of language for several reasons. First, Chomsky has long held the position that semantics and syntax are distinct, hence his celebrated adage "Colorless green ideas sleep furiously." This sentence suggests, to Chomsky, that a sentence can be both grammatical and meaningless. Pinker agrees, stating that “grammar [is] a code that is *autonomous* from cognition” (1994: 87; italics in original).

Another reason for isolating the formal property of recursion from the broader aspects of the language faculty, from an evolutionary standpoint, is that most, if not all the mechanisms of FLB can be found as homologous capacities, differing only in quantity and not in quality, in other species. Only FLN is uniquely human, in a qualitative sense. What this isolation of FLN accomplishes is that, assuming FLN to be a biologically-based module, it becomes unavailable to cross-species comparative study and thus renders moot the question of how natural selection can account for its evolution. This fundamental point seems to have been the basis for Chomsky’s alarming silence regarding the question of the evolution of a proposed Universal Grammar module; it seems to be a conspicuous hole in generative theory. Chomsky has, at times, brushed aside the question of UG’s evolution:

> The answers may well lie not so much in the theory of natural selection as in molecular biology, in the study of what kinds of physical systems can develop under the conditions of life on earth (1988: 167).

Elsewhere, he has suggested that:

> …a mutation took place in the genetic instructions for the brain, which was then reorganized in accord with the laws of physics and chemistry to install a faculty of language (1998: 17).

For this reason, Chomsky has been seen by many as a “crytpo-creationist” (Pinker 1997b). What creates this problem is that such a genetic mutation would still be subject to selection, implying either that language takes it form because it presumably was the most adaptive of several competing genetic mutations relating to language structure, or else because this sudden mutation arbitrarily emerged without competitors and happened to endure. Chomsky seems to reject the first possibility:

> In studying the evolution of mind, we cannot guess to what extent there are physically possible alternatives to, say, transformational generative grammar, for an organism meeting certain other physical conditions characteristic of humans. Conceivably, there are none – or very few – in which case talk about evolution of the language capacity is beside the point. (Chomsky 1972: 97-98)

Chomsky does well here to reject the notion of competing grammar mutations. Arguably, if there have ever been alternate mutations, they might still exist today in some tiny, isolated, grammatically perverse, linguistic/genetic community. Chomsky seems to think that language as we know it, given the constraints of the human body and general laws of physical form, has the only possible structure.

Presumably, then, language as we know it emerged suddenly, fully formed, and augmented some prior system that lacked, at the very least, recursive syntactic structure, and probably also lacked compositionality and hierarchical structures, as well. Furthermore, what Chomsky asks us to believe is that this emergence was due to a genetic mutation that encoded these specific structural characteristics into the communication system, a mutation in an ancestral Eve that must have made her a lonely individual indeed, given that every listening ear she encountered sadly lacked the brain hardware to decode the structure, and thus the meaning, of her utterances. Hauser, Chomsky and Fitch differ only slightly from Chomsky's previous suppositions regarding the evolution of language. The authors avoid talking about what role natural selection might have played in the evolution of a biological, recursive processor for use in syntactic language, and focus on the possibility of the exaptation of other cognitive functions that may or may not be shared by other extant species. In any case, syntax is still seen by Chomsky as a biological, species-specific trait that appeared saltationally and could very well have been immune to the forces of natural selection, given its apparent uniqueness.

Pinker and Bloom (1990) attempt to reconcile this Chomskyan-style innatism with a selectionist account. Pinker and Bloom claim that language has the appearance of complex adaptive design which, if language is a biological trait like the vertebrate eye (Pinker and Bloom take this condition for granted), could only be explained by natural selection. Pinker and Bloom's aim in this particular paper is to "argue that language is no different from other complex abilities such as echolocation [in bats] or stereopsis [in monkeys], and that the only way to explain the origin of such abilities is through the theory of natural selection" (1990: 708). While this position, for many scholars, was, at the time, and continues to be a welcome divergence from Chomksyan non-selectionist hocus-pocus, there remains multiple problems. First of all, language is nothing like echolocation or stereopsis: as Tomasello states in his commentary on Pinker and Bloom, "human languages differ among cultures, take several years to acquire, depend in a basic way on the ambient social environment, and show large individual differences in skill level at maturity" (1990: 759). While Pinker and Bloom argue quite convincingly that language has the appearance of complex adaptive design, their fatal flaw, along with Chomsky and many others, is the primary assumption that the core of syntactic language is a biological, genetically encoded, neural device. These authors all neglect to consider, as this paper argues, that syntax might not be encoded in the brain at all, but externally, in language itself. Pinker and Bloom come very close to this consideration in their discussion of lexical variation:

> Once a mechanism for learning sound-meaning pairs is in place, the information for acquiring any particular pair, such as 'dog' for dogs, is readily available from the speech

> of the community. Thus *the genome can store the vocabulary in the environment…*
>
> (1990: 716; italics added, attributed to Tooby and Cosmides (1989)).

This paper argues that a grammar is not necessarily encoded biologically/genetically, but rather, like the specific semantic content of lexical items, is learned (i.e. socially based) and stored as complex semantic/symbolic relationships associated with lexical items. The challenge of this paper, then, will be to account for the premises, listed above, on which the innatist claim rests, while positing a claim that does not necessarily involve a genetic/biological, hardwired grammar module, all from an evolutionary viewpoint. Accordingly, then, the next section will use a comparative methodology to show what kinds of behavior other primate species exhibit that might be homologous to our own linguistic behavior, and how a protolanguage might have emerged from such abilities.

CHAPTER 2

PRIMATES TO PROTOLANGUAGE

Pinker (1994) specifically rejects the importance of data concerning the communication systems and skills of other primates, for various reasons. Primarily, Pinker claims that full human language could have evolved by natural selection since the time that our lineage diverged from the common ancestor that we share with modern-day chimpanzee species, indicating that it is not necessary to compare chimpanzee communication to human communication. He appropriately claims, furthermore, that if there are any features in primate communication that can be compared to those found in human language, these features could very well have evolved separately; that is, these features could very well be considered analogous characteristics, as opposed to homologous characteristics. Pinker's reasoning here seems as justifiable as the opposite conclusion, that of supposing any shared communicative features between primates and humans to be homologous. Given the amount of work done on primate communication systems, occurring both naturally, in the wild, and artificially, in the laboratory, however, it seems unreasonable to wholly ignore this data.

Pinker writes off non-human primate data because he deems it unnecessary for his language-as-instinct hypothesis, but goes on to suggest that since chimps cannot be taught to speak or gesture in human-like grammatical 'sentences,' any evidence of expressive communication has no relation to human language. In a sense, he is correct; chimp communication is not full human language. However, the prevailing assumption motivating the comparison between primate and human communication systems is that the chimpanzee species are the closest living relatives to our own species. If we are to speculate about the communication systems of our distant ancestors, then there seems to be no reason not to compare ours to that of the great apes. Pinker's justification for ignoring primate studies is analogous to supposing that there is a possibility that singing evolved independently in different, closely related species of birds; following Pinker's line of reasoning, this possibility is therefore a likelihood, and it makes no sense to compare song systems of different species of birds.

For the purposes of this paper, I will assume that there is a degree of probability that at least some shared features of human and primate communication systems were in place before the divergence of our species from theirs. Given that we are, in fact, primates ourselves, and closely related to the apes, this position does not seem unjustified. This discussion will focus on evidence of semantic/symbolic and referential/intentional systems of primates in the wild. The purpose of this discussion is to illustrate the groundwork of a proposed intermediate protolanguage: a communication system used by our hominid ancestors after our lineage split from the common ancestor shared by humans and our other primate cousins. Although this discussion is largely speculative (as is much of the theoretical discussion in this field), it is ultimately intended to show how the way was prepared for full human language to emerge in our species. The format for this section, then, will be first sketch the structure of communication and cognition in extant primate species, then to propose a structure for an intermediate protolanguage.

The most prominent, and certainly one of the most often cited, examples of primate communication is that of the call system of the vervet monkey. The vervet monkey call system consists of three specific vocalizations which alert others to the presence of predators (eagle, snake or leopard). When the call is sounded, vervet monkeys respond appropriately, hiding under bushes from an eagle, climbing a tree to escape a leopard, or simply fleeing an area in the case of the presence of a snake (Cheney & Seyfarth 1990). What remains interesting, and indeed controversial, about the vervet monkey case, is whether or not these animals are using a referential communicative system, and if so, what kind?

While primate alarm calls can be said to be meaningful, they are clearly a far cry from human language. Most obviously, they lack compositional syntax. An alarm or food call is expressed as a holistic utterance, an "indivisible package" (Marler 2000). An alarm call cannot be broken down into parts of speech, ordered in a particular way to express something meaningful. It seems likely, then, that the meaning of a vervet call cannot necessarily be said to refer to a particular object, but more likely refers to a whole event. Humans also sometimes use alarm utterances in a similar fashion (e.g. yelling "fire!" or "shark!"). The utterance, in the case of both humans and other primates, calls attention to the presence of an object of concern and recommends that a particular response be taken.

Even though the word "shark" contains semantic, symbolic meaning, if it is yelled on the beach, nobody stays in the water cognizing the concept itself as they would if, say, someone yelled "t-shirts!". If a swimmer sees a fin circling nearby, he doesn't wait to confirm its identity, he doesn't forget the name of the thing he sees, and he certainly isn't silent. Our swimmer's reaction is an emotional one: he yells "shark" and gets to shore as quickly as possible. He's not talking *about* a shark, he's pointing to it verbally and suggesting that everyone do what he's doing: fleeing to shore. Similarly, vervet monkeys cannot be said to be talking *about* an eagle when the alarm goes up, but rather pointing to its presence. Alarm calls make use of *deictic reference*, as opposed to *symbolic reference*. That is, the call itself is functions as a pointer, and its meaning is derived from the immediate context.

Both the production of and the response to an alarm call are emotional ones. Even in humans, one can imagine both of these processes as almost involuntary reactions to the situation. Vervet monkey alarm calls seem to be both emotional and involuntary. Deacon likens a vervet's production of an alarm call to human laughter (1997: 57). Like laughter, alarm calls are involuntary emotional expressions produced as reactions to environmental or social stimuli (1997).

Like laughter in humans, however, primate alarm calls do not seem to be totally uncontrollable. Like laughter, alarm calls are more likely to be expressed in the presence of other conspecifics. Like laughter, the alarm call has an arbitrary relationship with the state it communicates. McCune (1999) proposes that the intentional state (I-state) of a calling vervet

> probably includes focus on the specific predator with continuing peripheral awareness of the presence of other vervets, an affective tone relevant to the imminent danger, and an intention to take protective action (290).

An alarm call produces an I-state in other individuals that is similar to the state of the individual making the call. McCune states that

> [c]ommunication by matching I-states may be achieved adventitiously, without the speaker's intending to communicate by means of the vocal expression (290).

As such, communication is achieved by a sort of empathy, with an emotional state transferred through possibly unintentional vocalization. While alarm calls differ semantically from human language at least in that alarm calls do not make use of symbolic reference and in that alarm calls do not seem to be wholly intentional, that these issues are disputed implies a grey area in these dichotomies.

The alarm call systems of vervet and diana monkeys do not seem to be homologous with characteristics of human language, however, given that apes, with whom we share a more recent common ancestor than we do with monkeys, do not use such specific calls. Chimpanzees, however, “communicate in more flexible and interesting ways with gestures than with vocalizations” (Tomasello 2003: 99). A chimpanzee uses gestures to attract attention to him or herself, or as ‘incipient actions,’ ritualized gestures intended to initiate play, grooming, sex, etc. (Tomasello 2003). As such, chimpanzees use gestures to affect the behavior of others, usually toward the self, and not to call others’ attention to an outside object. Chimpanzee gestures are used imperatively, not declaratively.

As far as what type of communication system our early ancestors were using, one can only guess. It seems likely that human language could very well have emerged from an intentional gestural system similar to that of our closest relatives (see, e.g. Corballis 2002). Corballis (2003) suggests that, given the reliance of gesture on iconicity, speech may be more efficient and available for modification or clarification than is gesture. It is possible that vocalizations with a rudimentary referential system similar to that of vervet monkeys, emerging analogously as an expansion to an intentional gestural system, was evolutionarily adaptive to our savannah-dwelling, possibly pack hunting, socially inclined and increasingly bipedal hominid ancestors. The features of a protolanguage, then, might have been convergences, by both analogy and homology, of the kinds of communicative systems in use by extant primate species.

This is the point in the timeline at which we venture into dark waters. It will remain impossible to verify any speculation regarding how primate communication became specifically human communication. At this point, the story one tells about the evolution of language depends on how one conceives of human communication systems. Any story must rely, nonetheless, on what we know about the fossil record, which, unfortunately, contains only circumstantial evidence about communication. Thus, any formulation of what happens next must at least be in accord with this evidence. Furthermore, one must account for the emergence of every linguistic feature that our primate ancestors do not share with us. While I aim to account primarily for the features considered by innatist viewpoints to be their primary evidence of the genetic/biological basis of syntax, especially the characteristic of discrete infinity, my proposal is that these features emerged from increasing semantic and lexical complexity associated with the lowering of the larynx and increasing brain size. As such, my focus will be on communication in terms of semantic structure and the consequences of complexity.

It was noted previously that other forms of primate communication specifically lack semantic features found in human language. Strictly speaking, all that an individual needs in order to communicate is, first, a message, and then a medium through which to express the message. The efficiency and complexity of the medium, however, constrain the complexity of the message. We have seen that communication in non-human primates is very limited. It seems apparent that some apes communicate intentionally, (Tomasello 2003), and that some monkeys communicate specific information about their external environment (Cheney & Seyfarth 1990; Zuberbühler, Cheney & Seyfarth 1999). Apes don't seem to communicate information about their external environment, but they do seem to communicate about their internal environment. That is, apes often communicate to express a desire, and thus to affect the behavior of others to fulfill that desire. A major step in the evolution of language is when the imperative utterances of apes, our closest living relatives, became declarative. Given that the utterances of some species of monkey seem to be declarative, if in a limited way, it does not seem unlikely that such kinds of utterances could have evolved, by analogy, in our lineage as we left the trees and descended to the savannah. Certainly an alarm call system is adaptive for social groups, and an alarm system based on gesture just would not serve such a purpose. The first system of utterances could have been the kinds of expressive emotional reactions to threats that we find in vervet monkeys.

Coupled with the ability to communicate intentionally (i.e. proactive rather than reactive), this vocal system could have quickly expanded from an alarm call system to something more general in nature.

The key point here is that our ancestors began to intentionally express both internal and external states. If the two media of expression (i.e. gesture and vocalization) were indeed initially distinct, they began to blend, to merge. While the ape gesture system is based on visual iconic reference, the alarm call system of vervet monkeys is based on arbitrary, conventional sound/meaning pairs. The first intentional (i.e. not reactive) utterances were probably expressions of desire (for food, sex, grooming, etc.), and were probably iconic gestures paired with arbitrary, conventional vocalizations. Already possessing the ability to vocalize about external threats, our ancestors began to refer to other objects, individuals or events as well, thus the emergence of the word.

I propose that the first utterances, the first words, were holistic, that they were used to refer to whole events or internal states. The range of available utterances, and probably the range of events or states accessible for uttering, were certainly limited at this early stage. Four million years ago australopithecines walked erect through the grasslands, sharing food within nuclear family structures, having more children than their ape forbears and weaning their young for longer periods of time (Donald 1991). It seems likely that communication became increasingly complex, if only in small steps, during the two million years between the emergence of australopithecines and that of the oldest known habilines. Australopithecines had the relative brain capacity of slightly more intelligent apes. Many things seem to have changed, however, with the emergence of *Homo habilis* two million years ago. There are two key behavioral features that changed in the two million years between australopithecines and habilines: they began to hunt and they began to make tools with which to hunt. Along with the emergence of these new behaviors, the rate of the increase of the pre-hominid cranial capacity began to accelerate. While this might not suggest that pre-hominids starting waxing philosophical, as it were, a larger brain means fewer limits on the complexity of semantic representations available for communication. Furthermore, there is some evidence that there was an enlargement of Broca's area, an area in the left frontal lobe associated with speech production, in *habilis*, possibly pointing to greater control over oro-facial movement (Corballis 2002). If pre-hominids were already communicating internal and external states through means of gesture and vocalization, and if a larger brain with more robust neural architecture related to speech production and vocal control was emerging, then the only constraints on vocabulary expansion were the position of the larynx and the range of internal representations accessible for expression; the kinds of things available to talk about were limited by the vocal medium and the available messages. However, the range of available sound/meaning pairs was certainly more broad than that of an iconic gestural system, and this range grew as the brain expanded and the larynx lowered. This seemed to be happening during the 500,000 years between the emergence of *Homo habilis* and the emergence of *Homo erectus*, at which point the brain had grown to twice the size of that of our closest living relatives, the great apes (Donald 1991).

Homo erectus was, at its time, the most successful and intelligent primate that had ever lived. It engaged in group cooperative hunting, migrated extensively, made complex stone tools and used fire to cook food. It's relative cranial capacity was 80% of our own. It certainly had some form of communication. Donald suggests, furthermore, that erectus possessed a new skill: mimesis.

> Mimetic skill or mimesis rests on the ability to produce conscious, self-initiated, representational acts that are intentional but not linguistic...[Mimesis] involves the *invention* of intentional representations. (1991: 168-169).

Donald cites mimetic skill as "logically prior to language," and has the properties of intentionality, generativity, communicativity, reference, autocueing and the unlimited modeling of perceptual events (1991: 171). Donald bases this conclusion, in part, on evidence of complex tool-making abilities. For tool-making to be part of a cultural inventory, it must be intentional, generative (i.e. parsable into simple procedural components), and able to be communicated or taught; "the act must be distinguished from its referent," meaning that there must be an understood dissociation between a representation of an event and the event represented; a mimetic act should be able to represent an "unlimited number of perceptual events," constrained, presumably, only by the representational medium, and must be "reproducible on the basis of internal, self-generated cues" (Donald 1991: 172).

I have proposed that by the time *Homo erectus* emerged, 1.5 million years ago, there was a communicative system in place that was based on holistic utterances, vocalizations that signified objects or whole, unparsed event structures. As such, this protolanguage was largely syntax-free, in that propositional meanings were not expressed and derived based on compositional meaning. The lowering of the larynx allowed for a broader range of phonological distinctions, and thus the limitations of the vocal medium itself were reduced dramatically relative to prior species. The increase in cranial capacity allowed for a more robust representational system, evident in the emergence of mimetic skill. This alludes to an increasingly efficient and complex ability to express meaning vocally.

One way to express why vocabulary expands in such a situation is to consider lexical meaning (received message) as the analog output of a digital phonological filter (medium), through which analog semantic representations (input or intended message) are expressed. A representational message (sent or received) is analog in that it is not built from discrete units; phonology acts as a digital filter in speech production in that strings of discrete units, unrelated to the message, mediate between input and output. Assuming that the expansion of the brain reduced limitations on the complexity of semantic representations (input), the lowering of the vocal tract allows a wider range of discrete phonological units to be uttered, thus fractionating and complexifying the digital filter and allowing a more complex output. Put simply, if one were only able to utter two syllables, efficient expression of meaning through vocalization would be severely limited to only a few concepts. Increasing the syllabic inventory by a few more available discrete units would dramatically increase the number of possible strings. For example, if our speech were limited to two possible syllables, and if word length were limited to one, two or three syllables, we would only be able to express a maximum of fourteen words, barring multiple meanings and including repeated syllables within a word. With an inventory of four syllables, our maximum number of lexical expressions would increase to eighty. Thus, phonology sets limits on the number of possible expressions in a holistic protolanguage. As these limits are lifted, this posited protolanguage expands exponentially in its complexity, increasing vocabulary and thus paving the way for a transition to a syntactically complex language. The expansion of the phonological inventory thus has a relationship with the expansion of vocabulary, and was likely to have co-occurred with the expansion of the brain in Homo erectus. It is indeed possible, if not probable, that the biological characteristics of even a slightly lowered and more controllable vocal tract and an increased brain mass exhibited selective pressure on each other, through the expansion of an adaptive communication mechanism, in a Baldwinian bootstrapping effect.

The emergence of a protolanguage is conceived as such, then, and its structure would have been not just a more robust version of the communication systems of extant primate species, but would have also included representational elements, allowing the intentional use of holistic vocal symbols in addition to an iconic gestural system and a responsive, empathetic, state-sharing alarm call system. These holistic vocal symbols would probably have referred to objects, individuals and/or whole events in conjunction with deictic specifiers like intentional manual or facial gestures and possibly use of the body for mimetic representation. Such representational or symbolic utterances might have been used in the facilitation of coordinated hunting, or possibly in social (i.e. grooming, pedagogy), sexual or ritual displays (see, e.g., Okanoya 2002; Deacon 1997). Indeed, with the increasing reduction of constraints on vocabulary size, this protolanguage eventually would have expanded to all facets of *erectus* life. Within little more than a million years, *sapiens* began to emerge.

CHAPTER 3

MEMES, LEXEMES AND SYNCHRONIZATION

What was happening in that 1.2 million years between the emergence of *erectus* and the archaic emergence of *sapiens*? During this period, the brain expanded again to nearly its present size and the vocal tract began to resemble its modern form (Donald 1991). Protolanguage was on its way to becoming fully complex human language. At this point, there are two issues to discuss. Now that protolanguage had established itself as an external communicative form, and presumably internally, associated with a semantic representational system, both the external and the internal are important to consider in explaining how syntactically complex language emerged. First, we shall deal with the external: language is conceived here as a complex, adaptive, self-organizing system, itself subject to non-biological evolution. It has been established that language is complex; it's very complexity leads some to draw nativists conclusions about syntax. Language is adaptive in that it adapts to its users, i.e. it is shaped by the constraints placed on it by its interaction with the human brain, the human vocal tract and human social structures. It adapts through self-organization, which seems to be analogous to biological natural selection.

It is this self-organizing property of language that I wish to address at this point. This characteristic of language is key to the emergence of complexity. As Chomsky distinguishes between I-language and E-language, but takes only I-language as his object of study (e.g. in Hauser, Chomsky and Fitch 2002), the object of this discussion will be formal properties of E-language.

A language is a standardized system of information exchange between users. Language is functional in that it is used to serve the purpose of communication. Language users rely on conventional symbols combined in structured, predictable ways. Linguistic communication systems vary between cultures; however, systematically, there are universals in that all humans use their vocal apparatus to produce a medium through which to communicate internal semantic or propositional representations. Furthermore, all humans have a degree of commonality in terms of what kinds of objects, events or states are subjects of communication, and in terms of communicative goals. Language, as a medium, can be considered, in a sense, an entity that is autonomous from its users, in much the same way that a medium such as the internet can be considered as such. It is not autonomous in that is totally separable from the human user, however; language does have a relationship with its users, as does the internet. Language takes its shape in its interaction with the biological (including neurological) machinery of human speakers. What I mean by saying that language is autonomous is just this: that language can be spoken of as interacting, as having a relationship, as being distinct in a sense from its users. What this autonomy allows is language's variation, its rich complexity.

Language, then, is by no means strictly bounded by its interaction with speakers; if it were, there would be a lesser degree of variation between languages. It is, however, constrained by its relationship with humans. Its structure as a medium is constrained by the physical apparatuses utilized by its speakers. It must be user-friendly; it almost has the appearance of having been designed to be user-friendly, i.e. designed for communication through human bodies. The reason for this appearance of design is because language has evolved into the human biological, cognitive and cultural niches over generations, and it continues to do so as these niches, each one its own complex adaptive system, change over time.

Language, as an autonomous communicative medium, can be considered a memetic system. A meme is simply a unit of cultural transmission, just as a gene is a unit of biological transmission. Just as genes encode information guiding biological design, memes encode information guiding behavior. Memes, however, are non-biological, and therefore beyond the guidance of the genes to varying degrees. Memes are only loosely constrained by biology because a meme is external, a meme is cultural, a meme is a unit of *imitation* (Dawkins 1976). Imitation, akin to Donald's (1991) mimesis, can be considered a species-specific characteristic of human behavior (i.e., of extant species). Blackmore (1999), following Dawkins (1976), uses the term 'imitation' in a broad sense: imitation takes place whenever a particular belief, idea or behavior (i.e., a meme) is copied from one individual or another. As such, memes, like genes, are replicators that use human cognitive systems as their vehicles (Dawkins 1976). Memes are copied, with varying degrees of fidelity, from brain to brain, through imitation.

As replicators, memes follow a universal Darwinist principle. There are three main features of an evolutionary system: variation, selection and retention (i.e. heredity or fidelity) (Blackmore 1999). Given these three features of a system of replicators, an evolutionary algorithm takes place and the system necessarily evolves. Language clearly follows this algorithm. Language is memetic because this algorithm occurs externally, constrained only to a certain extent by the body and the brain. If human behavior is based partly on biology and partly on culture, then genes program innate behaviors and memes program behavior that is not hardwired by the genes or otherwise influenced by the biological and functional design of the human body. Memes, then, are a set of second replicators, genes being the primary replicators for biological species. These units of cultural transmission may have emerged as an evolutionarily advantageous adaptation (e.g. mimetic skill), but have the capacity to evolve independently of the biological system, serving the purpose not of being advantageous for the genes of individuals, but evolving, i.e. undergoing replication with retention, variation and selection, as do the genes, to serve their own propagation. This is not to imply that memes have a mind of their own, as it were; neither do genes. However, memes are considered a second system of replicators precisely because their propagation is not necessarily based on what might be advantageous for the genes.

Blackmore suggests that the cognitive environment that makes memes possible, the capacities for learning and imitation, may have emerged by biological natural selection. Blackmore, quoting Dennett, suggests that the emergence of these second replicators could have subsequently pressured the human brain into evolving to better accommodate them:

> The haven all memes depend on reaching is the human mind, but a human mind is itself an artifact created when memes restructure a human brain in order to make it a better habitat for memes. (Dennett 1991: 207, quoted in Blackmore 1999:22)

This implies that the capacities for learning and imitation are found in our species to the extent that they are because of our relationship with culture, or our relationship with this external, evolving memetic system.

Regarding the evolution of language in our species, then, the short version of what I am proposing is as such: that if, as Pinker suggests, the only way to explain adaptive complexity of a biological organ is through natural selection, and if language is not a purely biological system, as the memetic model suggests, then we must account for the adaptive complexity, or the appearance of design, in language, in part, by the external evolution of language. As Blackmore states regarding any evolutionary algorithm: "When this algorithm gets going, the inevitable result is that design is created out of nowhere" (1999:12). I suggest that the grammatical structures we find in human language emerged from the ever-increasing complexity of a prior hominid protolanguage. What this entails is that, at a certain point, protolanguage began to take on the three features of an evolving system, i.e. variation, selection and retention.

Linguistic variation relies on the fact that language is not completely "on the leash" of the genes (to use Blackmore's term). Variation in a holistic protolanguage, then, would only be allowed by a system of arbitrary signs used to refer to objects, events, propositions or states. This system would necessarily be unhinged from innate behavior; in other words, this system cannot be totally instinctual (as the vervet alarm call system might be). The structure of the protolinguistic system that I have proposed fits these criteria. I have suggested that protolanguage emerged as a merging of a prior iconic, intentional gestural system, as can be found in chimpanzees in the wild, and an assumed system of alarm calls similar to that of the vervets. In other words, the arbitrary vocalizations of an alarm call system became intentional in the way that gestures are for apes. Communication with these intentional, non-iconic, holistic utterances might have been facilitated by iconic gesture initially (as linguistic communication often still is), but as vocabulary expanded in concert with a falling larynx and an expanding brain, these vocalizations began to take some precedence over gesture.

In order for these utterances to fulfill their purposes, that the receiving individual understands the message, whether the intended signals are imperative, declarative or otherwise, they must be transmittable to another individual with some degree of fidelity to local semantic and phonological conventions. Several formal mathematical models of the emergence of lexical conventions illustrate various minimal criteria sufficient for the emergence of a stable, standardized lexicon. Zuidema and Westerman (2003) deduce from their models four characteristics of an optimized lexicon: *specificity*, *coherence*, *distinctiveness* and *regularity*. Specificity and coherence refer to a one-to-one mapping between a particular form and a particular meaning, standardized across a community of speakers; in other words, an optimal lexicon has one word for each referent, one referent for each word, and these word/referent pairs are conventionalized. Furthermore, a lexicon with maximally distinctive forms for dissimilar referents (i.e. distinctiveness), and with similar forms for similar referents, "such that misinterpretations are still better than random interpretations" (i.e. regularity), is considered highly optimized (Zuidema and Westerman 2003: 393). Zuidema and Westerman demonstrate that in a population of speakers striving for optimal communication on the individual level (i.e., successful communication of a signal between speakers), these four criteria for an optimal lexicon emerge. Noise created by the environment or by the embodiment of the speakers themselves serves to "impose a topology on both the meaning and the form space of their communication system" (Zuidema and Westerman 2003: 400). This means that the shape of the emergent system is informed by its relationship to its users in their environment. While the noise that results from this relationship has the effect of stabilizing the communication system, a change in the relationship of the communication system to its users in their environments serves to introduce variation into the system. Smith (2004) finds very similar results, demonstrating that an optimal lexicon can emerge given only a learning bias in agents to avoid homonymy. Furthermore, a homonymy-avoiding agent introduced into a population that favors homonymy would eventually destabilize the bias in favor of homonymy, assuming these biases to be genetically determined, and the previously homonymous communicative system, in which at its extreme a single utterance would map to all possible meanings (an ineffective system indeed!) would quickly begin to tend toward optimality. Assuming an evolutionary advantage to effective communication, only populations either with no bias for or against homonymy or with a bias against homonymy would be evolutionarily stable and resistant to an influx of individuals with other biases. Put simply, the more individuals there are avoiding the invention and use homonyms, the more effective their communication system will be, and if effective communication is selected for, the more likely it is that individuals avoiding homonyms will dominate.

What both of these models suggest is twofold. First, the lexicon organizes itself in optimal ways based on the constraints of its users, and, second, certain kinds of constraints (in this case a homonymy avoidance bias) can be selected for biologically in populations where more effective communication is a general advantage. This process illustrates how the three characteristics of a universal Darwinist evolutionary algorithm can emerge in a holistic protolanguage, and how memetic evolution can drive genetic evolution. Variation and selection emerge from the noise created by the changing relationship of the communicative system to its users. Retention, or heredity, emerges as a lexicon standardizes itself for a population of speakers: felicitous transmission of messages through conventional form/meaning pairs is the key to an optimal lexicon. These three characteristics guarantee that the system will undergo evolution, which would tend toward a more efficient and effective communication system. Furthermore, as Smith's model demonstrates, individuals who are better equipped to acquire a more optimal lexical system would be selected for in a population if efficient and effective communication is advantageous, demonstrating how 'the memes can drive the genes,' as it were. I will return to this point later in this paper, with respect to the emergence of grammar.

CHAPTER 4

TRANSITION FROM PROTOLANGUAGE TO LANGUAGE

What I hope to have shown is that a simple holistic protolanguage, based initially very loosely on a combination of communication styles of other extant primates, could have emerged as a stable communication system in *Homo habilis*, evolving externally while gently nudging genetic evolution to favor better communicators. This vocal system initially might have consisted of a very limited number of alarm calls inherited from australopithecines, then expanded with the early pre-hominids to accommodate intentional vocalizations. Over hundreds of thousands of years, as the vocal apparatuses took shape, and as brain mass increased, this holistic vocal system continued to expand based on the increase in available phonological and cognitive resources, while at the same time tailoring itself to those resources through self-organization.

The processes that I have described can be understood as processes of *mutual synchronization*. Two complex, adaptive, evolving systems, namely the human body and external communicative systems, undergo processes of synchronization with each other. Language synchronizes with a population of embodied communicative agents through self-organization according to the constraints and demands of its users, while at the same time each individual language learner is synchronizing with the communicative conventions of the community. We have seen how a standardized lexicon emerges in a community of speakers. Figure 1 illustrates how an individual speaker might synchronize with the conventions of a community. In this diagram, the sender of a message has an internal representation of a proposition based on a given perception. To prepare this intended message for delivery, the sender hypothesizes a surface form. This hypothesized surface form is informed by the linguistic community. In other words, the surface form is hypothesized in terms of imitation: the correct form will ideally be a faithful reproduction of forms previously heard by the speaker. This form is expressed through the physical apparatus of speech (i.e., the phonological filter), and the message is uttered. The receiver of the utterance (i.e. the receiver of the message) gives feedback, intentional or otherwise, on the fidelity of the received message to what they think the sender intended. This can be as simple as whether or not the receiver's behavior has been affected by a request or demand (i.e., did my communication fulfill its purpose? Did the receiver do what I wanted them to do?). In other words, if I as a speaker intended to affect another's behavior, did my message achieve its effect? This feedback doesn't have to vocal. Based on that feedback, and based on the input from the linguistic community, the sender can alter or tweak the different parts of the message sending system. That person can tweak the relationship of the hypothesized surface form to both the propositional representation and the phonological filter, i.e. does this utterance mean what I think it means? and is it supposed to sound like this? Does the hypothesized surface form (a) match the conventional, standardized, community wide symbol for a propositional representation of that nature and (b) does that surface form match the community wide standards of what it should it sound like? This is the process of synchronization. This is how a message sending system like a human synchronizes itself with a community of message sending systems, because of this constant feedback in terms of message fidelity (to a standard).

This applies to protolanguage. A perception feeds into a propositional representation of that perception, which feeds into a hypothesized surface form, which feeds through a phonological filter to be expressed as an utterance. Synchronization, then, applies to both the communicative system of an individual and the conventional communicative system of a community. Furthermore, an individual's synchronization with community standards relies on imitation in a memetic sense, and thus the external memetic system takes shape based on the cognitive constraints and communicative demands of its users.

What I have tried to capture is the basic process by which a holistic protolanguage emerges and becomes standardized within a community of evolving users. As to the stability of such a protolanguage, it is likely that it quickly stabilized with the first set of shared, arbitrary, intentional utterances. However, there were limits on the expansion and evolution of this first set of utterances; the complexity of this set was still constrained by the physical system with which it interacted. Biological evolution is a much slower process than memetic evolution, thus protolinguistic evolution didn't simply take off upon the emergence of intentional vocalization; vocabulary did not skyrocket. Protolanguage was relatively stabilized as a holistic system precisely because of the slow pace of the evolution of the human biological equipment, especially the brain and the vocal tract. Protolanguage, during this time was constantly adjusting itself to the evolving human cognitive, cultural and biological niches, which, in turn, seem to have been adjusting to the constraints and demands of the communicative system itself. The expansion of a holistic lexicon was likely a slow process, which would have been taking place in the 500,000 years between the emergence of *Homo habilis* and *Homo erectus*. Furthermore, this transition would also have been a transition from a communicative system dominated by gesture with some vocalization, to a system primarily of holistic utterances, expressed in concert with gesture. With the emergence of *Homo erectus*, the next transition began.

Considering Hockett's (1960) 13 design features of language, the proposed protolanguage of *erectus* was only lacking the feature of *productivity*, which can be understood as compositional meaning, and may or may not have lacked the property of *dual patterning;*

while the property *displacement* (i.e., communicating about something displaced in time and space) may have not initially been a property of vocal communication, it was at least an aspect of mimetic representation and as a general cognitive skill could have translated easily. A compositional semantic system differs from a holistic one in that (a) meaning in a holistic semantic system is not derived from its parts, while in a compositional system it is, and, therefore, (b) a compositional system is an open system such that new meanings can be expressed by different configurations of conventional, discrete categorical (i.e. lexical) utterances, a property lacking in a system of holistic utterances. As to whether this hypothesized protolanguage had the property of dual patterning, it seems that a system that already has the properties of *arbitrariness*, *discreteness* (i.e. *distinctiveness* in Zuidema and Westerman's (2003) terms) and *semanticity* (i.e. *specificity* and *coherence* in Zuidema and Westerman's terms) would only be constrained by the semantic domains available for communication. Duality of patterning, which is characterized by the derivation of semantic meaning from the arbitrary configuration of discrete, meaningless phonological elements, is only functional and lends to communicative efficiency when the semantic domains available for communication reach a critical mass, such that either (a) the complexity of discourse must reach a plateau or (b) the communicative system must itself complexify to accommodate the pressure for complex meanings. Clearly, the complexity of discourse did not plateau with protolanguage, and if it did, this plateau was quickly destabilized by the emergence of dual patterning in the holistic system. I propose that the emergence of dual patterning, itself having emerged from the communicative pressure for increasingly complex discourse, triggered the emergence of a compositional (i.e. productive, open) semantic system, which in turn set the groundwork, as will be illustrated later, for grammatical structures to emerge in language. I suggest that the property of dual patterning did not emerge before the ascent of *Homo erectus*, and probably happened shortly before the emergence of archaic *sapiens*.

What the emergence of dual patterning allowed for is the rapid expansion of vocabulary. However, a holistic system is limited in terms of the efficiency with which one can communicate complex meaning. A rapidly expanding inventory of utterances available for use in the communication of semantic meaning is only functional if the domain of discourse is broad enough to require such complexity. A holistic system is clearly less complex and thus more limited in term of communicative efficiency than a compositional system. The question arises, then, of how this transition from a holistic to a compositional system occurred.

Wray (2000) proposes that compositionality emerges from the use of non-compositional utterances in a phonologically analytic transition. For instance, if two holistic utterances share phonological features, and are uttered often enough, language users will parse out particular shared features of the utterances and attach meaning to the sub-units. For example, consider the following strings:

tebima – 'give that to her'

kumapi – 'share this with her'

According to Wray, if these strings are used often, language users will parse out the segment *ma* and attach the meaning 'her' to it.

For a compositionality to emerge from a holistic protolanguage, at the very least semantic fractionation, or decomposition, is necessary. A holistic utterance does not necessarily refer to a particular object or motion, but rather refers to a complete event; meaning from a compositional sentence is attained from the sum of its parts. Wray's view assumes that semantic fractionation implies phonological fractionation. This is not necessarily the case.

As the vocabulary of a holistic protolanguage increases, there emerges a pressure for holistic utterances to go through semantic decomposition. It seems unlikely that protolanguage users would parse holistic utterances phonologically. However, as words emerge with competing semantic elements, there would be pressure for the holistic semanticity of utterances to be pared down. Holistic utterances tend to be bound by context, i.e. they tend to use indexical or deictic reference. Using modified versions of Wray's examples, let us assume that the holistic utterances above have the following meaning:

tebima – 'give [that object] to [that individual]'

kumapi – 'share [this object] with [that individual]'

The specifics of exactly which object and which individual would most likely depend on the context of the utterance, and would be supported by gesture and other non-linguistic cues. A few scenarios could explain the gradual emergence of compositionality without gradual phonological analysis of holistic utterances.

First, if the existence of proper names for individuals were to emerge, compositionality would likely follow from semantic fractionation. For instance, is a group begins to name individuals, then suddenly the holistic utterances cited above suddenly contain less semantic information, while at the same time expressing more meaning compositionally. For example:

tebima Mary – 'give [that object] to Mary'

kumapi John – 'share [this object] with John'

Furthermore, suppose the following utterances emerged to talk about specific objects in general ways:

Rakiku – 'game animal [existing somewhere] [doing something]'

Tipita – 'tool [being used, made, requested, etc.]

Again, the particular variable implied in these holistic utterances would depend on context, on would most likely be expressed with non-linguistic cues. When linguistic cues are available, even in separate holistic forms, it is likely that these utterances would be combined as such:

tebima tipita – 'give that tool to [that individual]'

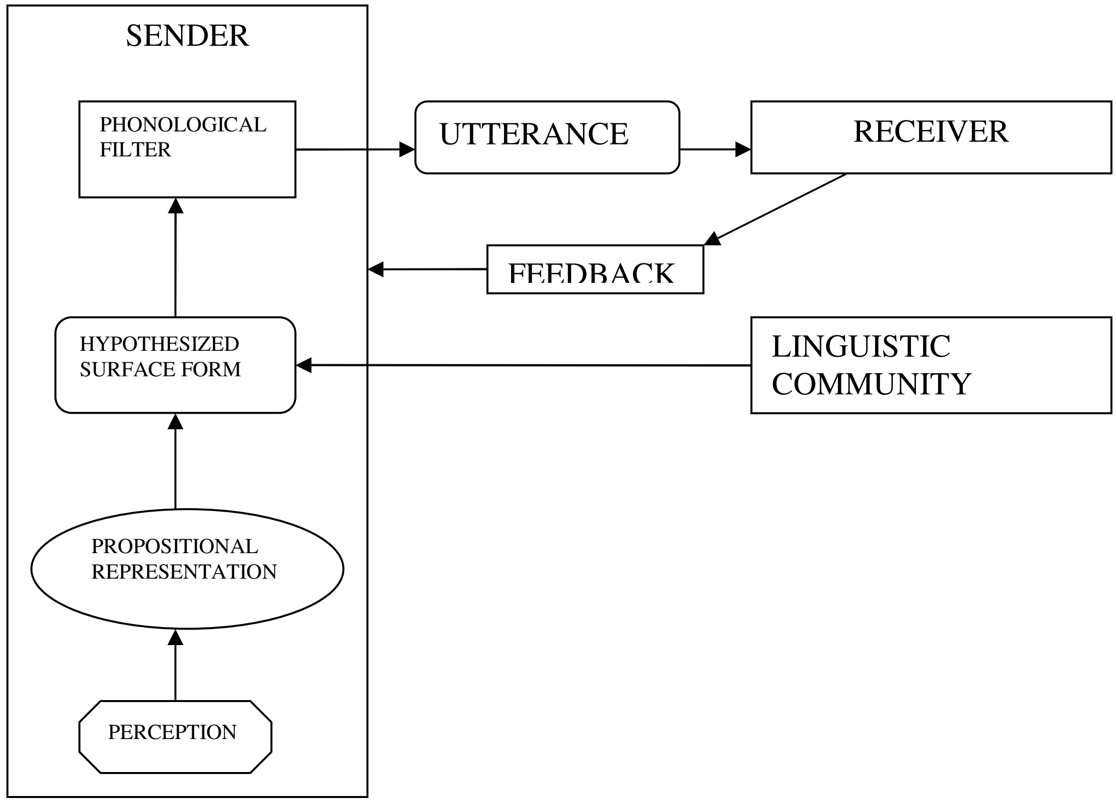
kumapi rakiku – 'share this game animal with [that individual]'

What this view supposes is that a core meaning exists in each holistic utterance, and that with an increase in vocabulary, meaning gets pared down to the core semantic feature of the utterance. Ultimately, the only pressure for the emergence of compositionality is the increase in vocabulary of a holistic protolanguage. While this is consistent with Wray's view, the particular ways in which the transition to compositionality takes place is not. Wray assumes a phonologically and semantically analytical transition to compositionality, while this view supposes a synthetic phonological transition and an analytical semantic transition from holistic utterances to compositional utterances.

Thus while the surface forms of utterances were undergoing synthesis, a new level of semantic representation was developing, or had developed, in the minds of speakers. As seen in Figure 2, propositional representations had begun to decompose and be re-represented as discrete categorical lexical units composed in order to represent a proposition. In other words, propositional representation began to decompose, began to fractionate internally. Rather than utterances representing propositions holistically, our representations began to be filtered lexically; the meaning of utterances began to fractionate. This fractionation or decomposition creates multiple hypothesized surface forms that compositionally represent a propositional representation. In order to express a whole proposition that is decomposed into different discrete units of lexemes, each lexeme must have its own hypothesized surface form. Each lexeme must be synchronized to its own external standard, informed by the linguistic community and tweaked by feedback with other individuals. Each discrete aspect of the decomposed propositional representation has its own hypothesized surface form, its own feedback, its own input, etc. This new compositional way of representing propositions vocally, one that is fractionated or decomposed into discrete lexical units, is certainly more efficient than a holistic system. If the hypothesized surface form of a holistic utterance does not match community phonological or semantic standards, it is wrong and it must be adjusted appropriately. Such an error would transmit a message that at best would be misinterpreted and at worst would be meaningless. With a compositional utterance, a propositional representation is composed of discrete lexical units. If one lexical unit likewise doesn't match the phonological or semantic conventions of the community, then depending on the complexity of the utterance, it is necessarily more felicitous to the propositional representation than a misinterpreted or meaningless holistic representation. A compositional utterance of 5 words with one word differing from the community standard still faithfully transmits some aspects of the proposition, even if the only part of the message that the receiver interprets correctly is a subject or an object or relationships between misinterpreted semantic content. In a holistic system, if the utterance doesn't match the standard, it's 100% wrong. A compositional system, then, is much more efficient than a holistic system. Should a holistic system be a memetic system, i.e. a system of imitation, a system of selection and a system of variation, as is the protolanguage that has been proposed, compositionality will likely emerge as that system evolves with its users.

What has not been addressed, however, is a necessary feature of a system based on compositional lexical units representing propositions: grammatical relations between such units. The emergence of compositional semantic representation represents a great leap in the evolution of a communicative system like human language precisely because of the subsequent, if not synchronous, emergence of grammatical relations. With the emergence of compositionality, characterized by a system in which categorical lexical units are configured in different ways to represent different propositions, communication moved from a closed system, limited qualitatively and quantitatively in what can be communicated, to an open system, a system limited not by syllables in configuration but by concepts in configuration, thus unlimited in what messages are available for communication. Just as the fractionation of the phonological filter, i.e. the lowering of the larynx, allowed for a greater number of holistic utterances, and thus a broader conceptual space available for expression, the decomposition of propositional representation introduced an increasingly fractionating lexical filter, allowing for a greater number of increasingly complex compositional utterances. This crossover from holistic to compositional representations, furthermore, seems to be, fundamentally, a tradeoff of cognitive resources. The demands of a holistic system on memory, given limited storage space for holistic form/meaning pairs, is such that, at some point, either a ceiling in the number of transmissible meanings is reached, or that ceiling is breached by adopting a compositional system, opening the communicative system and allowing for an unlimited number of meanings to be transmitted. Thus, compositional meaning emerged not because of an expansion of cognitive resources, but because of embodied limits on those resources, coupled with the tendency of external language to self-organize into optimal communicative states (see, e.g. Kirby 1998; Batali 1998).

The impetus behind this fractionation of propositional representations and the emergence of compositional utterances could very well have been the pressure for increasingly complex communication as the social structure of Homo erectus complexified. Furthermore, a complex communication system is clearly adaptive in that it allows for more complex planning, for example, in group hunting and gathering. If a group is to work together to perform complex tasks, a communication system at least of greater complexity than a holistic system is not only adaptive, but seems to be necessary. If dual patterning began to emerge in the communication system of *Homo erectus*, then this pressure for greater communicative complexity might have also been a pressure in the continuing lowering of the vocal tract as well as the greater encephalization of the species following *erectus*, *Homo sapiens*.



(Figure 1) Sender has representation of intended message. Sender hypothesizes a surface form in preparation for utterance. Hypothesis is based on prior input from the linguistic community. Hypothesized surface form is sent to the phonological filter and expressed as an utterance. Receiver gives feedback, intentional or otherwise, on the fidelity of the received message. Sender adjusts hypotheses based on feedback and information from the linguistic community.

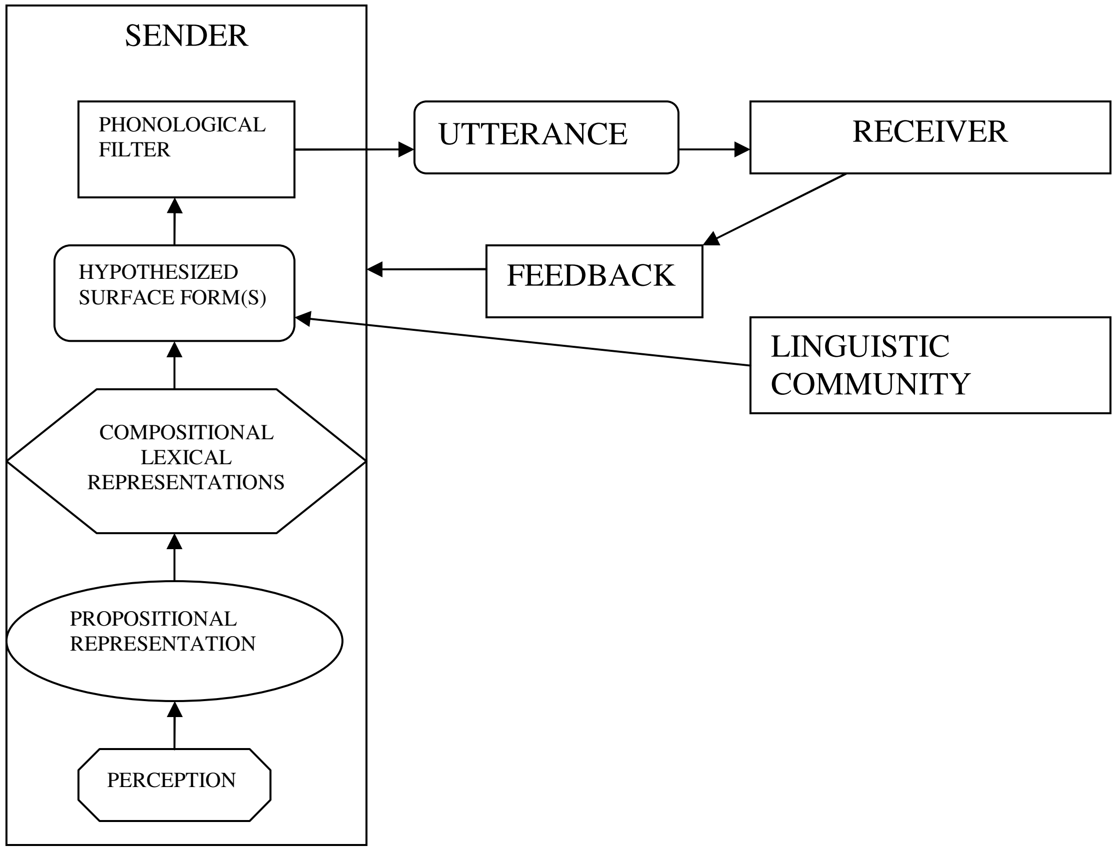


Figure 2. Propositional representation has decomposed into discrete lexical representations of propositional elements. Each lexical representation has its own hypothesized surface form, each of which is informed by the linguistic community.

CHAPTER 5

ACCOUNTING FOR INNATIST EVIDENCE

It is assumed that a rudimentary syntax (e.g. loose rules for word order) emerged with compositionality; indeed, there must be a distinction between "the child burns the log" and "the log burns the child" for these utterances to even be considered compositional. Syntax, however, as most linguists will testify, is much more complex than the order of subject, object and verb in an utterance. It is this very complexity that creates that illusion that language is unlearnable without syntactically specific, biologically innate specifications. Before the emergence of this complexity is addressed, an account of other evidence for an innate grammar will be considered. To reiterate the premises of the innatist claims, there are five pieces of evidence to account for:

1) Rate of language acquisition – children acquire language completely, at a very high rate
2) Age-dependence – Adults have a more difficult time acquiring a second language than do children. While children acquire the language of their surrounding community fluently, even if the language is not their native one, adults who learn a second language rarely become fluent. There seems to be a critical age at which one can fully acquire a language.
3) Poverty of stimulus – A child seems to acquire general syntactic rules of a language from inadequate linguistic stimuli.
4) Convergence of grammars within a linguistic community – language users within a linguistic community show little grammatical variance.
5) Language universals – Unrelated languages seem to exhibit common structural features

Each of these will be dealt with in light of recent discussions, drawing on evidence from computational simulations of language evolution. First, it will be shown how grammars can converge in a linguistic community in absence of innate grammatical specifications (Premise 4). Furthermore, species-wide grammatical convergence in the form of cross-linguistic universals can occur (Premise 5), not directed by a universal grammar module, but rather shaped by constraints in the form of cognitive biases in learning, parsing and processing grammatical language. Premise 2 will be discussed in terms of the evolutionary advantages and implications of a critical period for language acquisition. Premises 1 and 3 will be subsumed in a single discussion regarding what many regard as the core feature of language's complexity: recursion. This feature will be discussed both in terms of external language and internal language, i.e. how did recursion emerge in language, and what kind of cognitive organization could have evolved to handle it or produce it?

While the flaws in the innatist argument have already been discussed earlier in this paper, the purpose of this discussion is to take each of these premises for granted and to offer an explanatory model that differs from the nativist explanation for these premises, i.e. that an innate, genetically determined, hardwired neurological module dictates the external structure of language. What this paper proposes is that each of these premises can be accounted for without such a module, by considering language as an external entity, evolving autonomously, in a sense, from its human users, but nevertheless constrained and shaped by cognitive biases inherent in the learning, parsing and processing of language by its users. In other words, while the nativist viewpoint purports to deduce an innate module as the cause of these premises, this paper will show that an alternate cause for these premises is plausible, thus weakening the nativist argument.

While the argument for an innate module from convergence of community grammatical standards is spurious (see, e.g. Sampson 1997: 45-48), grammatical convergence is still explicable in terms of the tendency for language to converge at an optimized communicative state. Such an optimum state would include not only lexical but grammatical conventions. As optimized lexical conventions can emerge in a population with biases against homonymy and synonymy, as illustrated by Smith (2004) and Zuidema and Westerman (2003), grammatical conventions can occur in similar ways. Communication is based fundamentally on an intent to transmit a message that, when received and decoded by an interlocutor, retains a high degree of fidelity to the original message (i.e., the sender's propositional representation). This intent to transmit felicitous messages is itself enough to explain the relative grammatical uniformity of idiolects across a population: we prefer to speak in such a way that others understand us (the opacity of academic discourse notwithstanding), and this is optimally and most efficiently achieved through a communicative system with structural conventions. The emergence of these conventions in external language is understood to be a consequence of the individual's intent to communicate effectively; the real problem behind the premise of grammatical convergence is not why individual grammars are similar across a linguistic community, but how individuals can learn and use similar grammatical rules given different input in initial stages of acquisition. This problem relates directly to the problem of the poverty of stimulus (Premise 3), which will be discussed later.

The account of the existence of language universals builds off of the argument for grammatical conventions as an emergent property of interactive agents wishing to communicate effectively. However, while standards within a given linguistic population are necessary for communication with other speakers of the surrounding community, cross-linguistic universals cannot be explained in the same way; there is no reason in terms of effective communication for there to be shared structural characteristics between mutually unintelligible languages. Such universals can either be explained, then, as vestiges of some Original Language spoken by our ancestors hundreds of thousands, if not millions, of years ago (i.e. as homologous characteristics), or are dependent on innate constraints, with models of cognition ranging from biases in learning, parsing and processing to a full Chomskyan Universal Grammar module. The first explanation, that universal features are homologous, is in fact subsumed by the second explanation; if there are innate predispositions, whether they be ‘soft’ biases or ‘hard’ modules, that shape the structure of external language, then they must have been with us since the initial emergence of compositional syntax or else must have evolved subsequently as a response to the increasing complexity of language. A third explanation for language universals, that these shared characteristics are mere analogous coincidences unrelated to cognitive constraints, is rejected on the grounds that these features represent only a small subset of all possible communication structures and therefore are not necessary for communication. Thus, the only way to explain language universals is to turn to species-universal perceptual and cognitive bases of these features.

The only question that remains, then, is of the nature of the constraints placed on language by the physical and cognitive processors involved in its use. A parsimonious model of these constraints will be preferable to one that posits more than is necessary to explain language universals. As such, it is best to start ‘soft,’ and consider whether general processing biases are sufficient to explain universals. Kirby and Christiansen (2003; also see Christiansen and Devlin 1997; Christiansen and Chater 2001; Christiansen and Ellefson 2002) make such a claim, suggesting that many universals emerge based on cognitive limitations on sequential learning. Two universal features accepted by many as evidence of an innate Universal Grammar are:

1) Consistent branching direction
2) The Subjacency Constraint

Feature 1 characterizes the branching direction of constituent nodes within a sentence. Languages universally, for example, implicate the use prepositions, rather than postpositions with all verb-object (VO) languages, and the use postpositions with all OV languages. The Subjacency Constraint (Feature 2) places restrictions on the number of binding nodes that can be crossed by *wh*-words, e.g. in the formation of questions. In various studies, Christiansen (with Devlin 1997; with Chater 2001; with Ellefson 2002) has shown how these features can emerge based on learning biases inherent in a Simple Recurrent Network (SRNs), connectionist neural network models trained to learn sequential input. Christiansen's SRNs are fed input generated from an array of possible grammars, and thus each grammar can be analyzed in terms of its learnability by the network. Christiansen finds that these two universals are consistently more easily learnable by the SRNs in comparison to grammars with inconsistent branching direction or no constraint on subjacency. Kirby (1998) shows, using Hawkin's (1994) theories of processing complexities, how universal branching direction consistency, with regional preferences for branching directions, can emerge in a population of interacting agents with only a bias towards minimal attachment to the mother node of immediate constituents (i.e. keeping related constituents close together). Kirby uses the following examples:

1) John **looked** the number **up**
2) John **looked** the number of the pub on Rose Street **up**.
3) John **looked up** the number of the pub on Rose Street.

All of these examples are grammatical, but there is a preference of the processor to use (3) instead of (2), in order to minimize the scope of the bolded constituent. Kirby (1998; 1999) shows how these general (i.e. not specific to grammar) cognitive biases place pressure on language to fit the demands of the parser (i.e., our cognitive systems). Kirby states:

> As long as the only way for a language to be transmitted from generation to generation is through being repeatedly produced and parsed, then the form of the language that survives this process will naturally be adapted to being produced and parsed. This is what I mean by saying that languages adapt to aid their own survival. (1998:371)

What Kirby and Christiansen have shown is twofold: first, that the universal structures we find in language happen to be the easiest for a connectionist neural network to learn and, second, that these biases, which are inherent in the process of sequential learning, are enough to explain the existence of language universals in absence of any further innate predispositions for these universal features. As Kirby (2002; with Christiansen 2003) has suggested, the initial acquisition of language by a population of learners creates a bottleneck for language evolution. When faced with this bottleneck, those varieties of grammatical structure that are more difficult to learn are statistically selected against. Language thus converges on varying levels, with some universal features and some that seem to vary according to certain principles or parameters as implicational features (i.e., related necessarily to other features). Critically, these universal features can be understood as reflecting biases originating in the process of learning sequential, hierarchical, compositional forms. It remains unnecessary to suppose that these universals are genetically encoded. Thus, it seems likely that these universals would have evolved over time in language, externally, as a reflection of these inherent biases.

The apparent existence of a critical period for language acquisition does, as nativists claim, seem to have a biological basis. However, this critical period may, as has been claimed, be part of a critical period for learning in general, implying that that it need not be based on a genetically encoded syntactic module. Kirby and Hurford (1997) show that the critical period for language acquisition, occurring and ending before sexual maturity, seems to have an evolutionary explanation. Individuals who have not completely acquired a language by the time of sexual maturity would, over time, be selected against, if facility with linguistic communication is a factor in reproductive success (e.g. as an element of sexual attractiveness), which is clearly likely given a population of cooperative communicators. While this might help explain the timing of critical periods for language, it does not explain why there would be a critical period for language in the first place. Kirby and Hurford's results on critical periods are couched in what they describe as an incremental learning model. Kirby and Hurford suggest that language is acquired incrementally, utilizing an innately constrained language acquisition device which deteriorates over time according to a genetically encoded schedule. They find, furthermore, that this deterioration corresponds to phases in which, initially, resources for learning are increased over time in response to input data (a constructivist phase), then become relatively stable as maturation is reached, a non-constructivist phase in which "control of development is insensitive to input" (1997: 25). Kirby and Hurford test this hypothesis against two others, one of which is a purely constructivist learning strategy, the other is a purely non-constructivist (i.e. innate) learning strategy. While a purely constructivist learning strategy seems to be better suited to learning language than a purely non-constructivist learning strategy, neither model alone offers an accurate analogy to human proficiency in learning, within a critical period of time (i.e., before puberty). Kirby and Hurford show how such an incremental learning strategy might emerge in a population of agents evolving over time, with no initial overall biases for one strategy or the other. What these results suggest is that while a language learning system that remains flexible over the lifespan of the organism (i.e. a constructivist strategy) may be more proficient at learning a language than a system that is innately constrained throughout the lifespan of the organism, it is evolutionarily advantageous to have a critical period of learning characterized by a shift from constructivist to non-constructivist strategies, thus striking a balance between flexibility and inflexibility of cognitive resources. In a sense, language is initially acquired at a high rate *because* of the flexibility of the cognitive resources, not in spite of them, as an insistence on an inflexible innate grammar module might suggest. Thus the *process* of language learning may be genetically constrained, or at least constrained by the genetically determined maturational schedule of the organism, while specific aspects of language such as grammar are initially not encoded neurally, but over time become 'soft-wired' and thus appear biologically innate. The evolutionary implications of such a proposal suggest that the neural plasticity of infants is not necessarily a mere consequence of being born large-brained and helpless, but rather could be considered evolutionarily advantageous to an individual in a population that is required to learn, among other things, a complex communication system in order to be considered a fully-functioning adult.

The question that remains, then, is whether language is learnable at all without prior innate constraints; the poverty of stimulus argument posed by nativists suggests that it is not. This argument revolves around the notion that language is a system that can construct infinite output from finite input. In other words, the question nativists seek to answer by proposing a Universal Grammar module is how the brain can learn a generative, hierarchical, recursive system. Their answer, appropriately, is to propose prior innate constraints. The nativist model of these constraints, however, goes beyond the minimum required to account for the evidence. Perhaps the centerpiece of all linguistic features that nativists refer to as evidence for a genetically determined linguistic module is the characteristic of digital infinity, which is allowed for by language's recursive property. A nativist accounts for recursion as a feature of external language by suggesting that an innate recursive processor specifically designed for linguistic use gradually or suddenly emerged as a biological trait in our species, allowing individuals to wield an infinitely productive communication system. The account proposed here, that recursion can emerge in language without genetic modification of the species, therefore requires not only an account of how that emergence might proceed, but also a refutation of the notion that language is unlearnable without a genetically determined recursive processor. What this account will show is that 'poverty of stimulus' is a misnomer: input for language acquisition is sufficient to the point of appearing to have been designed to fit the cognitive mechanisms and learning strategies that have emerged evolutionarily in the human species and its relationship with its systems of communication.

According to Pinker and Bloom:

> All you need for recursion is an ability to embed a phrase containing a noun phrase within another noun phrase or a clause within another clause, which falls out of pairs of rules as simple as NP → det N PP and PP → P NP. (1990: 724)

Kirby (2002) treats the emergence of compositionality and recursion as a necessary consequence of communicating about a broad range of semantic domains with a limited system of semantic storage. Just as a syntactically compositional system is more efficient in terms of communication than a holistic system, a recursive system further allows for a greater range of expressible meanings within the limited cognitive capacities of the human organism. Recursion, like compositionality, emerged as a consequence of cognitive constraints. Furthermore, when these structural characteristics are viewed as replicators, as memes in their own right, they seem to be more likely to survive the bottleneck created by individual language acquisition, given their greater expressive powers and learnability, than a system lacking compositional, recursive syntax (Kirby 2002). Kirby, using a computational evolutionary model of a population with no prior bias toward a recursive or compositional system, shows how this would occur in any population with only a bias for communicating effectively and, as such, a bias toward preserving semantic mappings of utterances across a population (i.e., *specificit*y and *coherence* in Zuidema and Westerman's (2003) terms). What Kirby's results suggest is that compositionality and recursion are the optimal solutions for the problem of communicating about a virtually unlimited range of meanings with only limited biological resources. Batali (2002) uses similar methodology and arrives at similar results. Both Kirby's and Batali's models presuppose, however, the *ability* to generalize from data and learn recursive systems. While recursion might very well be advantageous to an open-ended communication system with limited resources, it must be learnable by its users in order for it to have emerged in the first place.

The concept of digital infinity, as has been mentioned, is only applicable to human language in a formal sense, and not in a practical sense. No human can in actuality speak or understand an infinite sentence, due primarily to limits on cognitive processing, not to mention lifespan. Pinker seems to support the notion of computational models of language evolution, and computational models of language learning: "A good theory uses some independently established finding of engineering or mathematics to show that some mechanisms can efficiently attain some goal in some environment" (2003: 31); elsewhere, however, Pinker rejects the notion that computational models can serve as valid models of cognition in most cases (see, e.g. Pinker and Prince's (1988) discussion of parallel distributed processing/connectionist models of language learning). Christiansen and Chater (2001) propose and test connectionist neural network models that seem to be able to learn sets of recursive grammars, to a degree, using associative statistical learning strategies. A key point proposed by Christiansen and Chater is that these models need not be able to process unbounded recursion, because humans can only process recursive structures to a limited degree. As such, these connectionist models need only to be able to process recursion to a degree similar to that of humans. Christiansen and Chater find that Simple Recurrent Networks can learn to handle recursive input to a degree similar to that of humans. Furthermore, they find that certain kinds of recursion rarely found in human language, such as unbounded center-embedded recursion, are more difficult for their networks to learn, as they are for humans. What Christiansen and Chater conclude is that a connectionist neural network model can learn recursive forms actually found in human languages to a degree similar to that set by the human benchmark, and find it as difficult as do humans to learn recursive structures that are rarely found in human languages. While innatists claim a 'poverty of input' as a justification for an innate grammar module, with the concept of digital infinity as a centerpiece of their arguments, they ignore a certain 'poverty of output' in actual human use: language is not infinitely productive in a practical sense. This suggests that a connectionist model of cognitive organization, requiring a minimum of innate constraints, could be a valid one as it applies to the human cognitive system.

To generalize and summarize from the above alternate accounts of nativist premises, each of these pieces of evidence (i.e., age-dependence, poverty of stimulus, grammatical convergence and language universals) can be explained evolutionarily without presupposing innate, genetically determined constraints on language processing. The only prerequisites for these explanations are:

1) Mimetic skill and imitation
2) Inherent non-linguistic constraints stemming from limitations on learning structures and strategies
3) An initial period of neurocognitive plasticity followed by a period of relative stability and inflexibility, i.e. learned structures are soft-wired over time
4) A desire on the part of individuals in a population for effective communication, coupled with a need for efficient use of limited physical and cognitive resources, leading to optimized forms of communication in the form of compositional, recursive syntax

The account of language processing offered as an alternate explanation to the innatist account is thus conceived of as a connectionist model of learning. Furthermore, innate drives to imitate and to communicate propositional representations effectively and efficiently, within the limitations of such a connectionist learning mechanism, are understood to shape the external structure of language.

Such a model is consistent with models of language evolution offered by various scholars. Deacon (1997) suggests that the emergence of the ability to manipulate and use symbols was a major shift in the evolution of language and cognition. This ability is instantiated in cognition in terms of abstracting and re-representing percepts and propositions symbolically. Deacon proposes that the emergence of symbol use was, at least in part, a consequence of the ritualization of social relationships (e.g. pair bonding).

Karmiloff-Smith (1992) offers a connectionist account of development based on her theory of representational redescription (RR). She suggests that human development entails going through a series of phase changes, each of which involves a higher level of abstraction and re-representation of previous levels, beginning with implicit, procedural levels, followed by a series of explicit levels increasingly available to conscious access and verbal report. These phase changes are characterized as states of relative equilibrium which are made unstable as learning proceeds, followed again by a new state of relative equilibrium. While Karmiloff-Smith's model of cognitive development presupposes no innate linguistic knowledge, it allows for the modularization of knowledge over the course of development, suggesting that, indeed, it is plausible for linguistic knowledge to be 'soft-wired' as the individual matures.

These accounts lend support for the models of language and cognition offered above as alternatives to the innatist models. Karmiloff-Smith's account of ontogenetic development, in particular, while not explicitly suggesting an evolutionary account, seems to be lend a useful analogy to phylogenetic development. It is possible that the human species, in the course of its evolution into a language-using organism, initially used vocal communication in a purely procedural way, relying on implicit representations inaccessible to conscious manipulation. Only with the emergence of the capacity for symbol use did this level of implicit representation become available for abstraction and re-representation, and thus become available for conscious manipulation and intentional communication. There were perhaps two or three of these phase changes, corresponding first to the emergence of a symbolic, holistic protolanguage, then to the emergence of a compositional syntactic system, replete with dual patterning, productivity, etc. Finally, utterances themselves were able to be accessed as representations, allowing for embedded, recursive structures beyond the simple hierarchies of compositional sentences. While this is not necessarily the full story of phylogenetic development of the cognitive mechanisms required for language, it al least offers a plausible alternate framework that requires no innate magic language box in the brain and no miraculous macromutations, as suggested by Pinker and Chomsky, and no sudden shifts in the use of any pre-existing cognitive mechanisms, as is suggested by various theories involving exaptations (e.g. Bickerton 1998; 2002), co-optations (e.g. Carstairs-McCarthy 1999; 2000) or pre-adaptations (e.g. Lieberman 1985; 1998; 2000) for a biological basis of syntactic structure.

CHAPTER 6

SUMMARY AND CONCLUSIONS

What I have attempted to offer here is a gradual, continuous, parsimonious account of the evolution of the kinds of syntactic features and their cognitive counterparts that are commonly proposed by Chomsky et al. as evidence of an innate, genetically determined, domain-specific Universal Grammar, without presupposing that these features are explained only by such a biological module. To put this account in perspective, a summary of the whole evolutionary story is in order.

Human language has its roots in non-human primate communication. In terms of Hockett's 13 design features of language, primate communication lacks only traditional transmission, displacement, productivity and duality of patterning (1960: 93-95). Importantly, other primates communicate with systems, vocal and/or gestural, that possess the features of semanticity, arbitrariness, and discreteness. It is suggested that in early pre-hominids, a communication system emerged which combined previously distinct communication systems, namely an existing iconic, intentional gestural system with an arbitrary, reflexive vocal alarm system. As such, an arbitrary, intentional vocal communicative system emerged as an augmentation to a pre-existing gestural system. Over time, the communicative support system reversed, and gesture began to augment vocalizations. It is supposed that this system was emerging and evolving in the australopithecines. Communication would have been heavily reliant on holistic utterances, in concert with iconic gestures. Communication was likely limited to propositions situated in the immediate context. Furthermore, communication began to serve the function of social cohesion, allowing for family structures, longer weaning periods and division of labor.

The emergence of *Homo habilis* marked a period of relative instability. While australopithecines wandered the African savannah for 2 million years, *habilis* emerged and disappeared within one quarter of that time period, giving way to Homo erectus 1.5 million years ago. During this period, our ancestors' brains began to enlarge, and they began to use rudimentary stone tools. I suggest that this period is marked by the emergence of mimetic skill, of imitative skill, and/or of symbolic skill. These descriptions seem to be of a common concept, characterized by a change in the type of communication which began to emerge in *habilis* and came to fruition in *erectus*. While australopithecine communication was largely limited by anatomical and cognitive factors, the emergence of mimesis, imitation or symbol use relaxed the cognitive constraints to a degree, as the brain expanded to accommodate this new way of storing and communicating information. Accordingly, communication began to be transmitted by tradition, through imitation, and constraints on learning and memory began to shape the structure of vocal communication. Furthermore, the use of mimesis and symbolic communication allowed for displacement, in Hockett's sense, the ability to communicate about objects or events outside of the immediate perceptual field. This step in the emergence of language cannot be underestimated. The emergence of this skill, by whatever name it is called, initiated a domino-effect of complexification of communication that continues even in our species, albeit externally in our communication technologies. *Homo habilis* is likely representative of a transitional set of species, given the apparent rapidity with which *Homo erectus* emerged thereafter. The use of imitation and symbols set the stage for the increasing complexification of communication in *erectus*.

While *habilis* probably still used holistic vocalizations, the seed for compositionality was planted with the emergence of symbolic understanding. With erectus, this seed had grown, and holistic vocal representations decomposed into combinatorial lexical representations. While utterances in this phase of the evolution of language were syntactically simple, they were enough to stabilize the species as the brain gradually increased in size, the larynx began to lower gradually, and *erectus* began to leave Africa for the far reaches of Europe and Asia. With *erectus*, a state of relative stability was reached in terms of communication and cognition. While this new form of communication did not necessarily replace prior forms of communication, but merely augmented them, as arbitrary vocalizations augmented iconic gestures, this new form of compositional representations began to dominate communication, and hence further complexified as erectus evolved in *Homo sapiens*.

As the external structure of language strayed more and more from instinctive communicative behaviors of other primates, and thus began to diversify across the emerging cultures of the erectus species, language began to take shape as an autonomous, evolving entity unto itself. The memetic nature of language, the algorithmic characteristics of selection, retention and heredity, ensured that language would evolve beyond the genetic programs of prior species. In its evolution, language was, and is still, only constrained by the embodied physical and cognitive characteristics of its users. Rudimentary grammatical rules emerged with erectus, as reflections of the demands placed on the communicative system by the limitations of its populations of users. Grammatical rules are, in essence, symbolic (Deacon 1997). The emergence of compositionality was essentially an increase in the complexity of the symbol-using faculty. Grammatical and lexical standards, in terms of the conventionalized, symbolic mapping of meaning to form, evolved in language as optimal communicative strategies as a result of the desire on the part of individuals to communicative effectively and efficiently.

After a long period of relative stability with *erectus*, the complexity of communication approached a plateau, and pressure began again for the communication system to either remain stable within the constraints of erectus biology and cognition, or to breach the ceiling and drive the further evolution of the brain and vocal tract to its present state in our species. The emergence of archaic Homo sapiens represents this unstable transitional phase.

On this model, the emergence of recursion has no definite place on the timeline. The only constraints on its evolution in language are based on the limits of learning and cognition in a communicating population. It is likely that recursion did not emerge until a few hundred thousand years ago, with the transition to *Homo sapiens*, but it could have emerged sooner if the cognitive capacities existed in *erectus*. Given the relative stability of *erectus*, however, it is likely that the emergence of complex grammatical features did not occur in language until this transition, and in part drove the transition from *erectus* to the more complex cultural and social systems of *sapiens*.

As such, as this paper is intended to illustrate, there remains no need to posit a genetically determined biological Universal Grammar module as an explanation for the kinds of complex features we find in human language. While it seems to be suggested here that language is only different from prior primate communication systems in terms of quantity, rather than quality, I would instead suggest that the externalization and autonomization of language, the freeing of language from the leash of the genes, allowed for, if not directly effected, the emergence of qualitatively different forms of communication. While only the biological structures of our species have changed quantitatively (i.e., lower vocal tract, bigger brains), the product of such changes allow for wholly new patterns to emerge in our systems of cognition and communication. The evidence proposed in support of a biological language module can be accounted for as the products of the interaction of language with our flexible, but inherently constrained, systems of learning and cognitive processing. Grammar seems modular because of the developmental soft-wiring of communicative structures, phonological, semantic and grammatical, as we approach a mature state. Grammatical rules and features, including recursion, emerged as a product of both our neural plasticity and our cognitive limitations. There was no miraculous appearance of a magic language brain box, and so there is no need to argue over how such a genetically determined, modular syntactic structure might have evolved: put simply, it never did.

BIBLIOGRAPHY

Batali, J. (1998). Computational simulations of the emergence of grammar. In Hurford, J.R., Studdert-Kennedy, M. and Knight, C., 405-426.

-------- (2002). The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In Briscoe, T., ed., *Linguistic Evolution through Language Acquisition*. Cambridge: Cambridge University Press.

Bickerton, D. (1998). Catastrophic evolution: the case for a single step from protolanguage to full human language. In Hurford et al. 341-358.

--------(2002). Foraging versus social intelligence in the evolution of language. In Wray, 207-225.

Blackmore, S. (1999). *The Meme Machine*. Oxford: Oxford University Press.

Bohannon, N., & Stanowicz, L. (1989). Bidirectional effects of imitation and repetition in conversation: adult responses to children's language errors. *Developmental Psychology*, 24, 684-689.

Botha, R. (2003). *Unravelling the Evolution of Language*. Oxford: Eslevier.

Briscoe, T., ed. (2002). *Linguistic Evolution through Language Acquisition: Formal and Computational Models*. Cambridge: Cambridge University Press.

Bybee, J. (2002). Sequentiality as the basis of constituent structure. In Givon and Malle 109-134.

Carstairs-McCarthy, A. (1999). *The Origins of Complex Language: An Inquiry into the Evolutionary Beginnings of Sentences, Syllables and Truth*. Oxford: Oxford University Press.

--------(2000). The distinction between sentences and noun phrases: an impediment to language evolution? In Knight et al., 248-263.

Cheney, D.L. and Seyfarth, R.M. (1990). *How monkeys see the world: inside the mind of another species*. Chicago: University of Chicago Press.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

-------- (1972). *Language and mind*, 2nd ed. New York: Holt, Rinehart and Winston.

-------- (1976). *Reflections on language*. Temple Smith.

-------- (1988). *Language and problems of knowledge: the Managua lectures*. Cambridge, MA: The MIT Press.

-------- (1998). Language and mind: current thoughts on ancient problems. Part I & Part II. Lectures presented at Universidad de Brasilia, published in *Pesquisa Linguistica* 3,4. Page references are to the English manuscript.

Christiansen, M.H. and Chater, N. (2001). Finite models of infinite language: a connectionist approach to recursion. In Christiansen and Chater, 138-176.

-------- and Chater, N., eds. (2001). *Connectionist Psycholinguistics*. Westport, Conn.: Ablex.

-------- and Devlin, J. (1997). Recursive inconsistencies are hard to learn: a connectionist perspective on universal word order correlations. In Shafto, M. and Langley, P., eds. *Proceedings of the 19th Annual Cognitive Science Society Conference*. 113-118. Mahwah, NJ: Erlbaum

-------- and Ellefson, M. (2002) Linguistic adaptation without linguistic constraints: the role of sequential learning in language evolution. In Wray, 335-358.

-------- and Kirby, S., eds. (2003). *Language Evolution*. Oxford: Oxford University Press.

Corballis, M.C. (2002). *From Hand to Mouth; The Origins of Language*. Princeton: Princeton University Press.

-------- (2003).

Dawkins, R. (1976*). The Selfish Gene*. Oxford: Oxford University Press.

Deacon, T. (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*. New York: Norton.

Dennett, D. (1995). *Darwin's Dangerous Idea*. London: Penguin.

Donald, M. (1991). *Origins of the Modern Mind*. Cambridge, Mass.: Harvard University Press.

--------(1998). Mimesis and the Executive Suite: missing links in language evolution. In Hurford et al. 44-67.

Farrar, J. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology*, 28, 90-98.

Givon, T. and Malle, B.F., eds. (2002). *The Evolution of Language out of Pre-language*. Amsterdam: John Benjamins

Gopnik, M., ed. (1997). *The Inheritance and Innateness of Grammars*. Oxford: Oxford University Press.

Hauser, M., Chomsky, N. and Fitch, W.T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science*

Hawkins, J.A. (1994). *A Performance Theory of Order and Constituency*. Cambride: Cambridge University Press.

Hockett, C. D. (1960). The origin of speech. *Scientific American*, 203 (3): 88-96.

Hurford, J.R., Studdert-Kennedy, M., and Knight, C., eds. (1998). *Approaches to the Evolution of Language: Social and Cognitive Bases*. New York: Cambridge University Press.

Karmiloff-Smith, A. (1992). *Beyond Modularity*. Cambridge, MA: MIT Press.

Kirby, S. (1998). Language evolution without natural selection: from vocabulary to syntax in a population of learners. *Edinburgh Occasional Papers in Linguistics*.

-------- (1999). *Function, Selection and Innateness: The Emergence of Language Universals.* Oxford: Oxford University Press.

-------- (2002). Learning, bottlenecks and the evolution of recursive syntax. In Briscoe, T., ed. *Linguistic Evolution through Language Acquisition*. Cambridge: Cambridge University Press.

-------- and Christiansen, M.H. (2003). From language learning to language evolution. In Christiansen, M.H. and Kirby, S., eds. *Language Evolution*. Oxford: Oxford University Press.

-------- and Hurford, J.R.. (1997). The evolution of incremental learning: language, development and critical periods. *Edinburgh Occasional Papers in Linguistics*.

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

Lenneburg, E. (1967). *Biological foundations of language*. Wiley

Lieberman, P. (1985). On the evolution of human syntactic ability: its pre-adaptive bases, motor control and speech. *Journal of Human Evolution* 14: 657-668.

--------(1998). *Eve Spoke: Human Language and Human Evolution*. New York: Norton.

--------(2000). *Human Language and Our Reptilian Brain: The Subcortical Bases of Speech, Syntax and Thought*. Cambridge, Mass.: Harvard University Press.

Loritz, D. (1999). *How the Brain Evolved Language*. Oxford: Oxford University Press.

Okanoya, K. (2002). Sexual display as a syntactical vehicle: the evolution of syntax in birdsong and human language through sexual selection. In Wray, A., ed. 46-63.

Pinker, S. (1994). *The Language Instinct*. New York: HarperCollins.

-------- (1997a). *How the Mind Works*. New York: Norton.

-------- (1997b). Evolutionary biology and the evolution of language. In Gopnik, M., ed. *The inheritance and innateness of grammars*. Oxford: Oxford University Press.

-------- (2003). Language as an adaptation to the cognitive niche. in Christiansen & Kirby, eds. *Language Evolution*. Oxford: Oxford University Press.

-------- and Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. In Pinker, S. and Mehler, J., eds. *Connections and Symbols*. Cambridge, MA: MIT Press.

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

Sampson, G. (1997). *Educating Eve: The 'language instinct" debate*. London: Cassell.

Smith, K. (2004). The evolution of vocabulary. *Journal of Theoretical Biology*, 228: 127-142.

Tomasello, M. (1990). Grammar yes, generative grammar no. Peer commentary on Pinker and Bloom (1990), in *Behavioral and Brain Sciences* 13: 759-760.

Tomasello, M. (1995). Language is not an instinct. *Cognitive Development* 10: 131-156.

Tooby and Cosmides (1989). On the universality of human nature and the uniqueness of the individual: The role of genetics and adaptation. Journal of Personality.

Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P. and Passingham, R. (1995). Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Procedures of the National Academy of Science USA* 92, 930-933.

Wray, A. (2000). Holistic utterances in protolanguage: the link from primates to humans. In Knight et al. 285-302.

-------- ed. (2002). *The Transition to Language*. Oxford: Oxford University Press.

Zuberbühler, K., Cheney, D.L. and Seyfarth, R.M. (1999). Conceptual semantics in a nonhuman primate. *Journal of Comparative Psychology* 113, (1), 33-42.

Zuidema, W. and Westerman, G. (2003). Evolution of an optimal lexicon under constraints from embodiment. *Artificial Life* (9), 387-402.