Akita Prefectural University RECCS Bulletin,No.6 (2005) : 1 − 8

Symbols and Syntax; Emergence in Language Evolution— An Overview

Stephen Shucart

INTRODUCTION

Understanding the evolution of human language is arguably the 'hardest problem in science' (Christiansen and Kirby, 2003). In the year 1866 this subject was considered so controversial that the Societe de Linguistique de Paris, the prime authority on language in its time, banned all discussion of language evolution. Scientific interest wasn't rekindled until 1975, when the New York Academy of Science sponsored a conference entitled 'Origins and Evolution of Language and Speech.' The reasons for this hiatus are manifold, but one of the principle, nonsectarian, causes was that the study of Language Evolution must include multidisciplinary research, and this has only recently become practical. One of the most significant new advances comes from the use of modern, highpowered computers for mathematical and computational modeling. The various fields that offer perspectives on language evolution include: psycholinguistics, linguistics, psychology, primatatology, philosophy, anthropology, archeology, biology, neuroscience, neuropsychology, neurophysiology, cognitive science, and computational linguistics.

The modern mind was coemergent with complex human language (Merlin, 1991). Communication exists between members of many animal species, but what separates human language from that of the other animals? Davidson (2003) says "There is a broad measure of agreement that there are two features of language … that are generally not present among the communication of other animals: symbols and syntax." More specifically, the question addressed in this paper is whether syntax is a genetic adaptation or a result of cultural constraints. Christiansen and Kirby state:

> "…there seems to be agreement that prior to the emergence of language some pre-adaptations occurred in the hominid lineage. There is less agreement about what these may have been, but one candidate that seems to be put forwards by many is the ability for using symbols.
> Most also see grammatical structure as emerging during a later stage in language evolution, though opinions differ as to whether this was a consequence of an evolved innate grammar or the emergence of grammar through cultural transmission.
>
> (Christiansen and Kirby, 2003, p. 13-14)

I have divided this paper into two parts, Syntax and Symbols. Syntax is 'the traditional term for the study of the rules governing the way words are combined to form sentences in language.'(Crystal, 1991). In the context of Language Evolution, Syntax must be thought of in the terms of Chomsky's (1981, 1986) innate Universal Grammar, and that a 'Language Acquisition Device (LAD) is a complex biological adaptation evolved through Darwinian natural section.
I shall base the first part of this paper on research by such distinguished scientists as MIT's Steven Pinker, the Linguist Derek Bickerton, and others.

In part two I shall present research that places the human ability for complex symbolic communication at the center of the evolution of language. I shall focus of the works of the cognitive scientist Terrence Deacon, the computational linguist Simon Kirby and others who see language universals as an emergent product of the cultural process rather than a set of evolved innate constraints.

I - SYNTAX

No one argues the evolutionary primacy of Symbols over Syntax, rather the debate centers on whether or not syntax was a biological adaptation that allowed for the emergence of the modern mind and human languages. Though scientific interest in language evolution was rekindled in 1975, it wasn't until 1991, when Steven Pinker and Paul Bloom published their landmark paper 'Natural Language and Natural Selection' in the respected journal Behavioral and Brain Sciences, that the scientific study of Language Evolution began in earnest.

Is language an adaptation? An adaptation is a trait whose genetic basis is shaped by Darwinian natural selection, so the more precise question should be 'Is Language a Distinct Part of the Human Phenotype?' (Pinker, 2003) The alternative to this question would consider language to be a manifestation of a more general cognitive ability, for instance 'general intelligence'; 'a symbolic capacity'; 'cultural learning'; 'mimesis'; or 'hierarchically organized behavior'.

If language is an adaptation, the next key to unraveling the puzzle is to decide what is it an adaptation for? This is a question related to the engineering design of language, and the selection pressures that shaped it. Pinker (2003) echoes Dell Hymes' famous definition of 'Communicative Competence' when he states:

> "What is the machinery of language trying to accomplish?
> The system appears to have been put together to encode propositional information - who did what to whom, what is true of what, when, where and why - into a signal that can be conveyed from one person to another. It is not hard to see why it might have been adaptive for a species with the rest of our characteristics to evolve such an ability."
> (Pinker, 2003, p. 27)

Gathering and exchanging information is integral to the larger cognitive niche that Homo sapiens occupies. Information is what economists refer to as 'a non-rival good', and human hypersociality comes about because information is a particularly important commodity of exchange. The Biologist Maynard Smith (1982) used Evolutionary Game Theory to predict how organisms ought to intera-ct and co-evolve their own communication strategies.

Syntax has a cost requirement in attending to the order of words, and its benefits only exceed the costs when the number of events worth communication exceeds a certain threshold. The 'Syntax Threshold' can only be crossed in a complex, combinational environment when syntax becomes invaluable to an analytical mind. Maynard Smith and Szathmary (1985) identified eight major transiti-ons in evolution, and the last transition is the emergence of human society with language.

The distinguished linguist, Derek Bickerton, sees the evolution of language as a two-stage process. The first stage took place 1.6 million years ago when Homo erectus developed a symbol based 'protolanguage', but 'protolanguage' lacked syntax. Rather than subscribe to Pinker's theory of the gradual evolution of language through the means of natural selection, Bickerton based his theory of innate syntax on the idea of 'Punctuated Equilibrium' (Gould and Eldredge, 1972), i.e. long periods of stasis in the evolutionary record punctuated by sharp jumps in complexity, then a leveling out on a higher plateau. Thus the second stage in the evolution of language came about through the same series of mutations that triggered the rapid evolution of Homo sapiens and the modern mind about 120,000 years ago.

In order to bridge the evolutionary gap between the symbolic representations of 'protolanguage' and modern syntactic speech Dickerton (2003) states ' the most crucial thing to grasp about the emergence of symbolic representation is that it must have been primarily a cultural rather than a biological event.' He believes the potential for symbolism to exist in many animals, including such phylogenetically distant species as the African Grey Parrot. In order to posit the rapid, genetic mutation and adaptation of syntax in modern homo sapiens, Dickerton uses 'surface minimalism' to reduce syntax to only three components:

> " (a) Conditions on the attachment of words to one another.
> (b) Cycles of attachment yielding domains that consist of heads and their modifiers (phrases and clause).
> (c) Principles derived from the order in which constituents are attached to one another."
> (Dickerton, 2003, p.89)

To summarize Dickerton's hypothesis, at some time in pre-history hominids were driven by climatic changes into habitats where food was scarce and predators were common so they began to exchange basic information about the environment in order to survive. A protolanguage of symbols multiplied, but structure, because of an inadequate number of neurons and connectivity, did not. A genetic mutation in a small community catalyzed the development of syntax, and this

mutation was so advantageous that it spread rapidly through the hominid population. The rest is, as they say, history.

Proponents of the theory of Syntax as an innate biological adaptation posit a genetic adaptation for language, but what proof do they have? Some interesting evidence for this genetic adaptation comes from recent studies of the KE family, an extended family known to possess a genetic disorder of speech and language. The disorder is transmitted as an autosomal-dominent monogenic trait encoded by a mutation on chromosome 7, known as FOXP2 (Corballis, 2003). A team of geneticists reported in an article in Nature (Enard et al. 2002) that the FOX2P protein occurs in other higher mammals, but that it is highly conservative. Thus Pinker believes that its association with genetic speech disorders in Homo sapiens is proof that it was a target for natural selection, rather than a statistically random mutation (Pinker, 2003). Whether this genetic adaptation was gradual or sudden is still open to debate.

II - SYMBOLS

Cognitive science has no more troublesome term than the word 'symbol'. Terrence Deacon (2003) says 'the standard caricature of the symbolic reference relationship is a code in which a random correspondence is set up between elements of two kinds, symbol tokens and objects.' This definition breaks down because it assumes sets of discrete referents. In reality the realms of words and physical world phenomena must be linked via a complex and multilayered intervening web of semiotic relationships. Much of the confusion over the term 'symbol' stems from the fact that the technical use of the expression has diverged along at least two distinctive disciplinary lines; the humanities and social sciences on one side and mathematics and cognitive science on the other. Deacon summarizes them as follows:

> "Humanities: A symbol is one of a conventional set of tokens that marks a node in a complex web of interdependent referential relationships and whose specific reference is not obviously discernible from its token features. Its reference is often obscure, abstract, multifaceted, and cryptic, and tends to require considerable experience or training to interpret.
> Computation: A symbol is one of a conventional set of tokens manipulated with respect to certain of its physical characteristics by a set of substitution, elimination, and combination rules, and which is arbitrarily correlated with some referent."
>
> (Deacon, 2003, p. 118)

It would seem that the old 'Nature versus Nurture' debate is the paradigm of the question of Symbols or Syntax as the prime mover of language evolution, but is it? An innate Universal Grammar evolved through random mutation and Darwinian natural selection sits on the Nature side of the spectrum, and the 'Symbolic' connectionist, neural network approach which

spontaneously self-organizes in a social context sits on the Nurture side.

Major advances in computing power have made it possible to simulate the complex interactive dynamics of language learning. In these simulations a population of agents interacts to produce self-organized emergent behavior that can reflect aspects of language evolution. Languages are neither natural nor artifacts, but rather they are a third kind of a phenomenon that results from the summed independent actions of individual agents (Hurford, 2003). One of the earliest proponents of computer simulated NK fitness landscapes to explore alternatives to Darwinian natural selection was the Biologist Stuart Kaufman of Dartmouth College and the Sante Fe Institute (Kaufman, 1993). Both exaptation (Gould, 1991) of pre-existing traits and saltation, the inheritance of acquired characteristics, are compatible with modern evolutionary theory, and can provide models for the cultural evolution of syntax without the need to hypothesize an innate UG and LAD.

Simon Kirby's Iterated Learning Model (ILM) is one of the best computer simulations of language evolution. ILM is a multi-agent model that falls within the general framework of situated cognition (Kirby and Christiansen, 2003). What makes this model unique is that the data that makes up the input to learning are also the output of the same process. It treats populations as consisting of individual agents, each learning by observing the behavior of the other agents. Kirby and Christiansen say:

> "Iterated learning is an evolutionary approach to the link between bias and structure, in that it looks at the dynamical system that arises from the transmission of information over time. It no longer makes sense to talk about language structure as being purely innately coded. Instead, learning bias is only one of the forces that influence the evolutionary trajectory of language as it is passed from generation to generation. Other factors include number of utterances the learner hears, the structure of the environment, social networks, and population dynamics."
>
> (Kirby and Christiansen, 2003, p.290)

The cognitive scientist Terrence Deacon does not dispute the fact that human languages contain Universal Grammar, but he believed that the nature/nurture debate has diverted attention from the more important question of the ultimate origins of the design principles of language. He thinks that there are major aspects of UG that are neither cultural nor biological in origin. Instead he believes that core language universals reflect semiotic constraints inherent in the requirements for producing symbolic reference (Deacon, 2003). His position is:

> "With respect to both so-called 'symbolic' approaches (which merely invoke features of UG by fiat and assume their pre-formation in the human mind) and various inductive, functional and 'neural network' approaches (that attempt to demonstrate how they can spontaneously emerge as optimizing solutions in

response to communicative or learning constraints) this represents a third paradigm of explanation, with affinities with both sides of the debate. It posits universality of a kind that is prior to language experience and yet also argues that this is only expressed functionally. As these constraints shape the self-organization and evolution of communication in a social context."

(Deacon, 2003, p. 112)

Deacon's 'semiotic constraints' are prior to language and thus are, strictly speaking, not innate, yet they influence the emergence of universals of language structure. Mathematics provides one of the foremost exemplars of universals that are neither biological nor cultural in origin. Because of its deep arbitrariness mathematics done anywhere in the universe will have the same form, though the medium of notation will be radically different. For example, a 'prime number' is a generalization about the limitation of division. Primeness is universal in the most literal sense, and it is implicit in any sufficiently complex mathematical system. Recognizing the possibility of this type of universal is a prerequisite in escaping the meaningless debate over the nature or nurture of language universals. Both numbers and words are symbolic references that require rule systems in order to be manipulated, so this is where the semiotic constraints connect. Mirroring Pinker, Deacon (2003) says ' similar to mathematical constraints on quantitative equivocation, avoiding equivocation about who did what to whom, when, with what, in which way, is a critical requirement in language communication.'

Thus the universals of grammar come for free, as it were, required by the very nature of symbolic communication itself. Semiotic constraints comprise selection pressures to speed language evolution towards forms that effectively communicate by not violating them. These semiotic constraints are self-organizing and emergent, and therefore they limit the space of possibilities in which language can evolve.

CONCLUSION

Complexity Science is a relatively new field, but its multi-disciplinary approach to research and its reliance on high-powered computers to construct models and simulations makes it especially useful for understand language origins and evolution. Scientists on both sides of the Symbols versus Syntax debate on the origins of language have made good use of it. This paper merely provides an overview of one of the main controversies in the field. There are many others, including the timing for the emergence of complex human language; some say as recently as 30,000 years, other put it back as far as 200,000. There is also the debate as to whether vocalization or manual gestures provided the main pre-adaptation. Where does the origin of art and complex tool manufacture fit into the timeline? Did other hominids, most notably the Neanderthals, also possess language? The field is both wide and deep, and as new research is undertaken and more powerful computers are created, perhaps some of the answers will become clear. Until then, they must remain in the realm of speculation.

REFERENCES

Bickerton, D. (1995) Language and Human Behavior. Seattle: University of Washington Press

Bickerton, D. (2003) 'Symbol and Structure: A Comprehensive Framework' in Christiansen, M. H. and Kirby, S. (eds) Language Evolution, Oxford: OUP

Christiansen, M. H. and Kirby, S. (eds) (2003) Language Evolution, Oxford: OUP

Christiansen, M. H. and Kirby, S. (2003) 'Language Evolution: The Hardest Problem in Science?' in Christiansen, M. H. and Kirby, S. (eds) Language Evolution, Oxford: OUP

Chomsky, N. (1981) Lectures On Government and Binding. Dordrecht: Foris.

Chomsky, N. (1986) Knowledge of Language: Its Nature, Order, and Use. New York: Praeger.

Corballis, M. (2003) 'From Hand to Mouth: The Gestural Origins of Language' in Christiansen, M. H. and Kirby, S. (eds) Language Evolution, Oxford: OUP

Crystal, D. (1991) A Dictionary of Linguistics and Phonetics, Oxford: BLACKWELL

Davidson, I. (2003) 'The Archaeological Evidence of Language Origins: States of Art' in Christiansen, M. H. and Kirby, S. (eds) Language Evolution, Oxford: OUP

Deacon, T. (2003) 'Universal Grammar and Semiotic Constraints' in Christiansen, M. H. and Kirby, S. (eds) Language Evolution, Oxford: OUP

Deacon, T. (1992) The Symbolic Species. New York: W.W.Norton.

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

Enard, W.,M. Przeworski, S. E. Fisher, C. S. Lai, V. Wiebe, T. Kitano, A. P. Monaco, and S. Paabo (2002), Molecular evolution of FOXP2, a gene involved in speech and language. Nature 418:869-72.

Gould, S. J., and Eldredge, N. 'Punctuated Equilibria: The tempo and mode of evolution reconsidered. Paleobiology, 3:115-251, 1977

Gould, S. J. (1991) Exaptation. Journal of Social Issues, 47, 43-65

Hurford, J. (2003), 'The Language Mosaic and its Evolution' in Christiansen, M. H. and Kirby, S. (eds) Language Evolution, Oxford: OUP

Kauffman, S. (1995) At Home In The Universe: The Search for the Laws of Self-Organization and Complexity. Oxford: OUP

Kirby, S. and Christiansen, M. H. (2003) ' From Language Learning to Language Evolution' in Christiansen, M. H. and Kirby, S. (eds) Language Evolution, Oxford: OUP

Noble, W., & Davidson, I. (1996) Human Evolution, Language, And Mind. Cambridge: CUP.

Pinker, S. (2003) 'Language as an Adaptation to the Cognitive Niche' in Christiansen, M. H. and Kirby, S. (eds) Language Evolution, Oxford: OUP

Pinker, S. and Bloom, P. (1990), 'Natural Language and Natural Selection' Behavioral and Brain Sciences 13:707-84.

Smith, J. M. and Szathmary, E. (1985) The Major Transitions in Evolution, Oxford:OUP