Abstract: Human syntactic language has no close parallels in other systems of animal communication. Yet it seems to be an important part of the cultural adaptation that serves to make humans the earth’s dominant organism. Why is language restricted to humans given that communication seems to be so useful? We argue that language is part of human cooperation. We talk because others can normally trust what we say to be useful to them, not just to us. Models of gene-culture coevolution give one plausible explanation for how language, cooperative institutions, and the genetic basis for both could have evolved. Why did the coevolutionary process come to rest leaving a huge space for the cultural evolution of language? We argue that language diversity functions to limit communication between people who cannot freely trust one another or where even truthful communications from others would result in maladaptive behavior on the part of listeners.
Humans are highly unusual animals in depending upon social transmission from others for acquiring most of their adaptations (Boyd and Richerson 2005; Richerson and Boyd 2005), including the specific languages they speak. Language is essential to our complex social life, revolving as it does around institutions that are transmitted by language and operated by oratory. Culture itself evolves by processes that are something like the evolution of genes, but which are different in many important details. In the case of language, the cultural evolutionary processes by which languages change over time are tolerably well understood from the work of sociolinguists (Labov 2001, 1994) and historical linguists (Deutscher 2005). See Progovac (this volume) on how language might have evolved gradually.

The functional hallmark of human language is to combine a large vocabulary of meaningful words into utterances using syntactic rules that convey much information about how the string of words in an utterance modify each other’s meanings. Both of these basic features of language contribute to the ability of hearers to rapidly convey large volumes of information to listeners. The compositionality of language inherent in having simple words that can be used with syntax to build up complex meanings may also make language easier to learn than a communication system that used holistic words to represent complex meanings (Brighton, Kirby, and Smith 2005). We have little to say directly here about the evolution of vocabulary and syntax beyond noting their functional role. Rather, we are interested in a more general question: How might evolutionary processes favor the evolution of a high information volume communication system in humans? To put our objective provocatively, we suggest that syntax and semantics are the easy part of the evolution of language. The hard part is to figure out how humans could make use of language. Put the other way, other animals don’t have language because words and syntax are hard to evolve but because under most circumstances a high information volume communication system is of no use to them. To use language, hearers must trust speakers. But evolutionary biology teaches us that trust between individuals evolves with difficulty. Hence, the difficulty from the evolutionary perspective is not how trust can evolve. If trust evolves complex communication is easy.

Cultural evolution complicates our adaptive analysis because we have to explain many adaptations like language based upon the evolution of both genes and culture. Genes and culture coevolve, as the jargon goes. Since no other animal can be taught to speak, despite many serious attempts to teach them, we know that some sort of genetic evolution was necessary for a language capacity to arise in humans. On the other hand, the vast differences between languages and their manifest change over time indicate that culture and cultural evolution play a large role in language. We assume that some relatively easy gradual path was available such that the evolution of language could proceed from simple vocal and gestural communications to human language by some combination of small cultural and innate steps. We realize that many if not most students of language assume the opposite and this claim is thus highly controversial. We wish to bracket this debate in order to introduce the hypothesis that the evolution of cooperation is also a major, if not the prior and more important issue.

Pinker and Bloom’s (1990) scenario for the evolution of language is much like what we propose, although they were not explicit about the cultural evolutionary mechanisms they imagined operating. We have good models of the coevolutionary process (Richerson and Boyd 1989;
Feldman and Laland 1996) including models tailored specifically to language (Steels In press; Briscoe In press; Brighton, Kirby, and Smith 2005). Nettle (2007) raises the possibility that this sort of coevolution still goes on. Subtle differences in innate capacities may have coevolved with different language families; for example, populations with a long history of using tone to make linguistic distinctions may find it slightly easier to learn and use tones than populations not so exposed. Good evidence to this effect would suggest that the coevolutionary process indeed affects language evolution.

This bare-bones account leaves many unanswered questions. Two seem fundamental to us. First, why is human language unique? Not only primates but also parrots and songbirds have preadaptations for vocal communication and vocal imitation (Pepperberg 2000; Baptista and Trail 1992). Human language is widely counted as our most spectacular adaptation, the very adaptation that made us human, and the adaptation that makes us the earth’s dominant organism (Maynard Smith and Szathmary 1995). Many species would seem to have rudimentary communication systems that could fall under the influence of the coevolutionary mechanism outlined above. Under what scenario might humans, but not other species, come under selection for an advanced inter-individual communication system?

Second, how can we account for the large role of culture in language? Many innatists are themselves minimalists who think that only relatively deep cognitive features of language are innate (Chomsky 1995; Hauser, Chomsky, and Fitch 2002). Minimalists themselves debate about exactly what is innate, what is perhaps innate but not specific to language, and what can be off-loaded to the cultural system (e.g. Arsenijević and Hinzen, this volume). Generative grammarians who pursued the idea that surface diversity of grammars could be rigorously tied to a parsimonious set of principles and parameters seem not to have succeeded (Newmeyer 2005). Agreeing with this assessment, sociolinguists and historical linguists propose mechanisms, based on observations about how language actually changes generation by generation, that seem to account for how a vast diversity of languages can evolve culturally (Guy 1996; Deutscher 2005). Whatever is innate is very deep and hardly constrains the design space of syntax and lexicon Progovac (this volume). The coevolutionary scenario above is silent about the division of cognitive labor between genetically and cultural transmitted elements at the end of period of active evolution of the language capacity. Briscoe (In press) provides a cogent summary of the issues involved in the coevolutionary pressure to evolve an innate language acquisition device. To the extent that culture can adapt language to cope with cognitive constraints (Brighton, Kirby, and Smith 2005), relatively modest responses on the part of genes might have been required and they might be changes related more generally to capacities for culture rather than language specifically. For example children are highly motivated to learn language and other non-linguistic cultural skills. The autism spectrum of disorders seems to include individuals who have a diminished motivation to learn language and other elements of culture even when their cognition is relatively unimpaired. Down’s Syndrome children, by contrast, are cognitively quite limited but acquire language and other skills without motivational problems up to the limits of their cognition. Of course, some genetic changes must have accompanied the evolution of language since even chimpanzees and bonobos when raised in a linguistic environment develop, at best, a rudimentary form of language (Savage-Rumbaugh and Lewin 1994).
We can imagine that language learning and subsequent communication would be much more efficient if much of it was innate. This is seemingly the logic of the much mooted massive modularity hypothesis for cognitive architecture generally (Tooby and Cosmides 1989; Sterelny 2003). For example, we could imagine a communication system that is largely specified innately (Sampson 2002). Such a system might superficially resemble a family of languages like the Romance Languages. Most of the grammar and much of the vocabulary would be quite similar across all the world’s languages, yet a “Castilian” child adopted by “Catalans” might have considerable difficulty learning “Catalan” because of having innate adaptations to speak “Castilian.” Or perhaps everyone would just innately speak “Catalan,” but with minor culturally determined “Valencian” and “Barcelonan” dialects. This scenario is clearly counterfactual. Languages seem to be much more culturally diverse than they need to be for communicative efficiency. Indeed, their balkanization into thousands of mutually unintelligible languages and tens of thousands mutually difficult to understand dialects would seem to favor a much more innately constrained system than we in fact have. We need a scientific account to replace the Tower of Babel story.

Why possibly language is restricted to humans

Language is not the only rich communication system in nature. The cells in our body use nerves and hormones to coordinate our exceedingly complex physiology and behavior. The social insects use pheromones and behavioral cues to coordinate the complex behavior of their colonies. However, it is true to say that most systems of inter-individual communication are quite rudimentary compared to human language and the intra-individual examples of complex communication.

The theory of animal communication was pioneered by John Maynard Smith (1976, 1994) and has recently been applied to human language by Lachmann and Bergstrom (2001; 2004). The basic problem is that individuals do not normally have enough commonality of interests to make honest communication possible. The relationship between individuals is normally competitive, and any communication thus tends to be self interested. However, this self-interestedness means that recipients of communications cannot trust the talk of communicators, so free communication does not evolve. Thus, animal communication is normally extremely impoverished. Often what animal communications exist are restricted to costly signals or in which deceptive signals can be punished cheaply. Peacocks’ tails and the red deer stag’s apparatus for roaring are two stock examples. These features advertise the size and health status of males, the quality of their genes, and their fighting prowess, to others in the vicinity in order to attract females and intimidate competing males. These signals are trustworthy indicators of male quality because only large, healthy males can carry a large display of gaudy feathers or develop the long vocal tract and support the energy expenditure to produce deep, loud roars. A population might start with cheap, accurate signals of genetic quality or fighting prowess, but such communication will be undermined by mutant males that use resources to fake the signal, resulting in an arms race that ends up with an expensive unfakable signal. At an evolutionary equilibrium, the signal will be so costly that inferior males simply have insufficient resources to generate a fake signal of quality. Such a system has such a high overhead that it can produce only a handful of critical messages. Lachmann and Bergstrom show that a combinatorial communication system that relies on rules to interpret a sentence composed of words is especially vulnerable to deceptive messages. The
value of a combinatorial communication system is that it permits limitless flexibility in the
construction of messages using a finite number of words, but deceivers can use this flexibility to
compose artful manipulative messages that recipients will interpret to the advantage of the
signaler but to their own detriment. Thus, the theory tells us that rich combinatorial language-like
systems should not evolve. This prediction is nearly correct; language-like rich communication
systems are of very restricted occurrence.

The exceptional cases of rich, cheap, honest communication systems result from a common
interest of signaler and recipient. This is easiest to see in the case of internal nervous and
hormonal communication systems. All the cells in your body have a common evolutionary
interest in your survival and reproduction. They have no “motive” to deceive each other. When
some of our brain cells tell our fingers to edit this sentence, the muscle cells trust the signal from
the brain implicitly. In the case of social insects, sterile workers have reproductive success only
indirectly via the reproductive castes of the colony. Their common interest in the success of the
colony permits trustworthy communication (Wilson and Wilson 2007; Hölldobler and Wilson
2009). Interestingly, biologists have described many mechanisms by which rich internal
communication systems are “policed” to ensure a community of interest, safeguarding the basis
of communication (Frank 2003). Such policing is necessary because pressure from deceptive
entities, like cancer cells, to reproduce at the expense of the collective always exists.

Human language is part of the extraordinary cooperation of humans with distantly related or
unrelated people. Because humans cooperate in large groups of distantly related and unrelated
individuals, we both have a use for language—to organize our social life—and enough
commonality of interest to support a rich, cheap, and trustworthy communication system. We
police the system—telling a lie is a sin. A number of evolutionary hypotheses are on the table to
explain this cooperation (Krebs 2008; Richerson, Boyd, and Henrich 2003). One of these is the
Tribal Social Instincts hypothesis (Richerson and Boyd 1999), which is a modernization of
hypothesis first developed by Darwin in the Descent of Man. It proposes that human cooperation
arose because human cultural variation is especially susceptible to selection at the level of the
group. The ancestral groups that were relevant during the evolution of our social psychology
were tribes, ethnolinguistic units ranging in size from a few hundred to a few thousand people.
Normally, selection is strongest at the level of individuals, leading to competition between
individuals for reproductive success. Groups of individuals can fall under strong enough
selection to lead to within group cooperation, but the conditions are fairly stringent. They seem
to be met most often when the groups involved are close kin. Kin assortment leads to reduced
genetic variation within groups and exaggerated variation between them, and group selection
requires persistent genetic variation at the group level to work. The reasons that culture is
unusually susceptible to group selection on a much larger scale than kinship groups are outlined
in Richerson and Boyd (2005: 203-13) and more formally by Henrich (2004). We will shortly
consider a linguistic mechanism that limits mixing between groups, helping to preserve between
group variation.

Innate elements of our social psychology—Darwin argued that sympathy and patriotism were
among our “social instincts”—would have evolved by gene culture coevolution following the
same logic outlined above for the language faculty. Language would have been an important
element of the evolution of cooperation. Language is obviously a rather general purpose
communication system that today we use to talk about practically everything. For the purpose of formulating hypotheses about the intertwined evolution of language and cooperation, we need to imagine what the main function(s) of protolanguage were near the beginning of the coevolutionary sequence that eventually gave rise to our advanced capacity for language. Merlin Donald (1991) argues that evolution of innate culture capacities proceeded in two phases. The first phase resulted in advanced capacities for observational learning that he called mimesis. Mimetic abilities would, for example, allow the cultural transmission of manual skills by direct copying unmediated by linguistic instruction. That process might get us hunting skills and stone tools of some sophistication. To get an impression of what mimetic pre-linguistic hominins could do with without language, Donald reviewed the literature on 19th Century deaf-mutes. Deaf-mutes led surprisingly normal lives. Most of them learned skills necessary to support themselves, typically a manual trade or farming, and some became prosperous. Many married and otherwise had relatively normal social lives. This evidence suggests, as Donald argues, that humans might well have had a fairly advanced cultural system before the capacity for language evolved. Thus, cultural group selection might have selected for a measure of extra-familial altruism before the evolution of language even began.

However, we think it unlikely that social systems of the complexity that living hunter-gatherers operate could have functioned without language (see Dunbar (1996) for a slightly different version of this argument). The social life of even the simplest known hunter-gatherer societies, such as the San peoples of the Kalahari and the Shoshone of the North American Great Basin, was regulated by institutions, commonly known rules governing social behavior (Maryanski and Turner 1992). For example, the Kalahari San peoples have a complex tripartite kinship system that in effect makes everyone at least a fictive kin to everyone else, even if no genealogical connection is known (Lee 1986). They also have a ceremonial gift exchange system that cements ties of friendship with people living in distant camps (Wiessner 2002). A large number of these relationships are with distantly related people, especially on the part of successful hunters and their wives. Even these comparatively simple institutions are encoded, transmitted, and reinforced linguistically through stories, gossip, complaining and shaming. In many cases, hunter-gatherer social institutions were not at all simple. The marriage systems of the societies in the Australian central desert were very elaborate, for example (Yengoyan 1968). One might imagine that the first steps toward human cooperation in large groups might take place based on mimetic culture, but it is hard to see complex institutions arising and being maintained across a large group without language.

The general proposition that human intelligence evolved for its social functions is sometimes called the social intelligence or Machiavellian intelligence hypothesis (Whiten and Byrne 1997; Dunbar 1996). It is often contrasted with an ecological intelligence hypothesis. We think that the social intelligence hypothesis as usually stated does not give sufficient attention to the fact that hunter-gatherer subsistence was an intensely social enterprise (Steward 1955). In humans, social intelligence is fundamental to our ecological adaptation. The human hunting and gathering adaptation depended substantially upon cooperative big game, especially during the late Pleistocene (Stiner 2002). Hunting probably provided the nutrients necessary to grow and sustain our very expensive brain (Aiello and Wheeler 1995; Kaplan et al. 2000). Many of the skills involved in hunting could perhaps have been learned by a linguistic mimesis—all those skills for which a picture is worth a thousand words. However, at least one skill, tracking, is practiced in
living hunter-gatherers as a sophisticated collective hypothesis testing enterprise in which verbal discussion is essential (Liebenberg 1990; see also Guthrie 2005). Trackers need to have natural-historical knowledge that often outruns that of the modern naturalists who have interacted with them (Blurton-Jones and Konner 1976) and would no doubt be hard to transmit without language. While modern biologists find many details of their theories dubious, the fact that they have a complex theoretical picture of animal behavior is beyond doubt. A science-like, highly linguistic, collective theory building exercise seems necessary to account for a tracker’s abilities. Bettinger (1991: 204-208) argues that Upper Paleolithic people, but not Neanderthals, exhibited a suite of characters among which language in the modern form was a key component. He sees the development of symbolic culture, of which language is the most important but unfortunately unfossilized element, as the master variable driving the last major modernization of human psychology. Marshak (1971) argued on the basis a close analysis of abstract engraved pieces that Upper Paleolithic hunters reckoned time using a lunar system. This suggests a mathematical symbolic capacity closely tied to language.

The evolution of human intelligence might well have been driven directly by recent climate deterioration. Some theoretical and empirical evidence supports the idea that the evolution of complex, cumulatively improvable culture, including languages, was driven directly by increasingly variable environments over the last few million years (Calvin 2002; Richerson, Bettinger, and Boyd 2005). Simple evolutionary models of culture suggest that the main adaptive advantage of a costly system of social learning is adaptation to environments with rapid, high amplitude variation (Boyd and Richerson 1985). Work on ice and ocean cores has shown that the last ice age had the sort of high amplitude millennial and submillennial scale variation climate variation that the models predict would favor culture (Alley 2000). A recent ocean core suggests that the amount of millennial and submillennial scale variation has increased over the last four age ages (Martrat et al. 2007). If the symbolic artifacts recovered by paleoanthropologists—scarcer and simple before 50,000 year ago, more numerous and more sophisticated after that date—are indicative of the emergence or improvement of languages, they may have evolved in response to the demands of the hypervariable climate of the last ice age. Coincidently or not the hypervariable part of the last ice age started about 70,000 years ago, not long before the paleoanthropological evidence suggests that human symbolic capacities and other indications of behavior likely to be related to language appeared in the record, and shortly before modern humans spread out of Africa.

Why is language weakly constrained by genes?

The number and diversity of human languages is stunning, akin to the species diversity of rain forest trees and coral reef fishes (Nettle 1999). The total number of living languages in the world is currently about 7,000 (http://www.ethnologue.com/), about 1,000 of which are spoken on New Guinea. Unlike species that typically differ in functionally important ways, languages are largely functionally equivalent.

We suggest that the function of language difference is closely related to other symbolic differences. Symbolic differences frequently mark the boundaries of social groups. They serve to identify those who are “like us” and those who belong to other groups. Language differences function to limit communication and hence the spread of ideas from one community to another.
People from other groups are liable to tell you self-serving lies, but, more importantly perhaps, they are likely to tell you things that are useless in the physical and social ecology in which you live. Human cultures are ecologically a lot like species. They typically have different social institutions and often different ways of making a living (Barth 1969). Imitating people from a different culture may often be a bad thing to do (Gil-White n.d.). Language diversity perhaps limits communication adaptively.

Mathematical models support this line of reasoning. Boyd and Richerson (1987) studied a cultural evolutionary model in which two populations lived in different environments. A different value of a quantitative adaptive character was optimal in each environment and there was a generic neutral marker character, also quantitative. (Quantitative characters are behaviors that can be measured on a continuous scale, such as the location of a vowel in formant space). Individuals migrated between the two populations, tending to homogenize them at some intermediate value of the adaptive character not well adapted to either environment. Counteracting mixing, juveniles were assumed to have a tendency to adopt the value of the adaptive character from people whose value of the neutral marker resembled theirs. They also prefer to imitate people who are successful. In the model, the marker characters in the two populations diverged, generating a correlation between the marker and the adaptive trait. At equilibrium, the mean value of the adaptive trait was at the optimum in both environments. The preference for imitating people like you with regard to a neutral marker trait in the presence of a strong correlation between adaptive and marker characters set up an adaptive barrier limiting the flow of wrong ideas from the other environment even in the face of rather strong physical migration. McElreath et al. (2003) studied a conceptually similar model but this time using discrete characters. They also modeled the environment as a social game of coordination rather than a different physical environment. Games of coordination are ones in which high payoffs depend upon matching the behavior of others. Languages are a massive game of coordination. Unless our grammar and lexicon at least roughly match those with whom we wish to communicate, we will fail. Once again migration tends to homogenize the populations but a correlation arises between the neutral marker characters and the move in the game of coordination. Eventually people in the two populations usually play with partners that correctly match their coordination move. Because language can evolve differences so rapidly, it can evolve to calibrate our discrimination against outsiders quite sensitively. I may perfectly well understand someone whose dialect differs only modestly from mine, but I may still distrust them. On the other hand, someone speaking a strange dialect is offering me something that I can judge independently to be highly useful, I might adopt it despite a general suspicion of such folks. Sociolinguists tell us that we make many subtle social decisions based on subtle linguistic differences. Those who study the diffusion of innovations tell us that we can overcome linguistic barriers to the diffusion of ideas and techniques if we are sufficiently motivated. The models are very crude by comparison, but they lend credence to the adaptive importance of linguistic diversity but also to the flexibility of linguistic boundaries.

These model results are in rough accord with the literature on ethnicity (LeVine and Campbell 1972). We hasten to add that ethnic and other symbolically marked boundaries (e.g. caste, class, profession) are very diverse and include such phenomena as intergroup alliances, divisions of labor, and assimilation and merger (Ruttan et al. n.d.). People speaking different languages may not normally communicate much, but they are not like biological species completely
informationally isolated from one another. Ethnographic data from small-scale societies suggest that linguistic and other symbolic boundaries often mark the limits of trust and cooperation. The role of ethnicity in conflict in modern societies is well known (e.g. Horowitz 2001).

The role of language itself as a symbolic marker of group boundaries has been well studied by sociolinguists (Labov and W 1980; Labov 2001). Labov and other sociolinguists discovered that subtle phonological changes grow up rapidly wherever social fault lines arise. For example, on Martha’s Vineyard, a rural island off the coast of Massachusetts, the influx of summer tourists caused Vineyarders’ speech to diverge detectably from the standard New England dialect. People seem to condition their behavior on the dialects that others speak. It is not hard to project the sociolinguistic mechanisms that cause small-scale generation to generation changes in languages to the multigeneration scale on which new dialects, new languages, and eventually new language families evolve (Deutscher 2005; Nettle 1999). Among the fields of language change that strike an evolutionist as well studied include grammaticalization, the process by which conventional lexical constructions come to have syntactic functions (Hopper and Traugott 2003) and the way in which languages in contact influence each other (Thomason 2001).

Conclusion

We think that the theory of gene-culture coevolution provides a useful framework for analyzing the evolution of language. In particular, it can provide a theory for the level of trust and cooperation necessary to make cheap, accurate, and abundant communication between humans possible and an explanation for why so much cultural variation remains in language. In conjunction with data from paleoanthropology, paleoclimatology, and paleoecology we can produce hypotheses about when language probably evolved and why it evolved only in the human lineage. Without doubt, all of these are very difficult questions. We do not flatter ourselves that the specific proposals here will survive challenges of new data, for example data on the history of the genes that underpin language. We are more confident that the form of the coevolutionary analysis is correct.

References


