The gradual emergence of language [1]

Brian MacWhinney

Language is a unique hallmark of the human species. Although many species can communicate about things that are physically present, only humans can use communication to construct a full narrative characterization of events occurring outside of the here and now. The religions of the world have interpreted this unique endowment as a Special Gift bestowed directly by the Creator. Scientists have also been influenced by this view of language, often attributing the emergence of language to some single, pivotal event in human evolution. The idea of linking language evolution to the emergence of a single Special Gift is certainly quite attractive, since it would account for so many facts in linguistics, psychology, and neuroscience through some single, simple mechanism.

The Special Gift view runs into a variety of problems, however, when we glance over the full landscape of six million years of human evolution. During this period, the hominid lineage has undergone a remarkable series of adaptations involving a straightening of upright posture, development of an opposing thumb, changes in the birth process (Hockett & Ascher, 1964), loss of hair (Morgan, 1997), adaptation of the gastrointestinal tract, increased innervation of the intercostal muscles (MacLarnon & Hewitt, 1999), loss of pronounced canine teeth, bending of the vocal tract, refinement of the facial musculature, freeing of the vocal folds, and sharpening of the chin. These morphological changes have been accompanied by a gradual tripling of brain size (Holloway, 1995) which has brought massive changes in the interconnectedness of the frontal lobes, changes in the linkage of vocal production to motor and emotional areas, linkages of the visual areas to motor areas, and expansion of many older areas, including the cerebellum, basal ganglion, and thalamus. Alongside these changes in morphology and neurology, human society has undergone a parallel process of development involving the expansion of social groups, the refinement of warfare, the development of tools, and the emergence of language.

Faced with evidence that evolution has effected so many detailed changes, adherents of the Special Gift approach will argue that one single change or process has been pivotal and the other changes are either unrelated to language emergence or else necessary

---

consequences of the basic change. In most cases, the Special Gift analysis focuses on some morphological or cognitive change that has occurred in the last 100,000 years, after the exodus of the human race from Africa. The changes that were wrought during the previous six million years are considered irrelevant to the emergence of language in homo sapiens.

The second major problem with Special Gift accounts is that they tend to treat language as if it were not grounded on cognition. In particular, they fail to account for the three-fold expansion of brain size that occurred primarily in the period between 2MYA (two million years ago) and .5MYA. The concept of co-evolution (Deacon, 1997; Givón, 1998) serves as a useful way of redressing this imbalance. The theory of coevolution holds that changes in neurological structure have led to advances in planning and communication. These advances in cognitive and communicative capacity then set the stage for further neurological modifications. This notion of an evolutionary ratchet effect is fundamentally attractive. However, without further elaboration, it would seem to depict a species that is slowly and steadily moving toward fuller and fuller control over its environment. In this view, each new advance in ability should be accompanied by a greater control over the environment and a spread of habitat. Unfortunately, matters were not so simple. After a remarkable expansion of habitat by homo erectus after 2MYA, the fortunes of our ancestors went into remarkable decline. The population of our direct ancestors then bounced back a bit in the early Pleistocene to a population of perhaps 100,000, but declined to a mere 10,000 at a point some 70,000 years ago (Stringer & McKie, 1996). If the story of the human race is one of the continual productive coevolution of language and the brain, we have to ask ourselves how it is that we so narrowly escaped extinction twice. The most recent near extinction is particularly problematic. At this point, homo was a highly intelligent species with a well-configured vocal tract and fully articulated social relations. Yet, despite all of its gifts, this talented group of individuals nearly vanished from the planet.

In all fairness to our ancestors, we know that the Pleistocene glaciations put great stress on a human population that had adapted to warmer, lusher conditions. In addition, 70,000 years ago the massive explosion of the Toba Batak volcano in Indonesia brought a sudden drop in global temperatures that worsened an already bad situation. On top of this, one can always imagine that there may have been a plague or some other unrelated evolutionary pressure. My point here is not to criticize early humans for a lack of resourcefulness. My point is only that language, by itself, was not enough to save our ancestors from near extinction. More generally, I think that we have tended to overestimate the extent to which language, in itself, provided a clear evolutionary advantage for early men and women. Instead, the basic advantages that have accrued to individual hominid species are those that relate to their abilities to handle weapons, form tight social groups, navigate across territory, and adapt to new dietary pressures.

My basic thesis is that language evolution operated across the full six million years of human evolution, not just the last 100,000 years. The model of language functioning I am adapting here is one that I have developed elsewhere (MacWhinney, 1999a). This model holds that language use involves the construction of a grounded cognitive simulation in

which we enact the activities of others through processes of perspective-taking and perspective-shifting. I will not repeat those arguments in detail here. Instead, I will refer back to the core pieces of the analysis in terms of the ways that they help us to better understand the processes of co-evolution that have operated across the last six million years.

We can distinguish at least four major co-evolutionary periods. During the first period, which lasted from 8MYA to 4MYA, evolution focused on the linking of morphological and neurological changes to the attainment of bipedal gait and the elaboration of the changes brought about from that new posture. During the second period, from 4MYA to 2MYA, the major co-evolutionary pressures involved the development of methods for solidifying group structure during face-to-face vocal interactions (Dunbar, 2000). This rather labile period is characterized by many failed evolutionary attempts (Wood & Collard, 1999) and a shaky movement forward toward homo erectus. During the third period, from 2MYA to 100,000 years ago, the focus of co-evolution was on the linkage of a broad set of symbolic processes to neural control mechanisms. These linkages involve the growth of prosody, chant, gesture, dance, and a variety of other largely social expressions of language functioning. As Donald (1991) has argued, during most of this period, the use of language was linked to mimetic forms that supported group processes and social structure. During the fourth period, which begins in earnest about 60,000 years ago, we see an increased evolutionary pressure for linking language to the production of a material culture. During this period, language assumed its current dual patterning with organization on the phonological and syntactic levels. However, many of these abilities involve precursors that were initially introduced during the third period for other functions. During each of these four periods, we see a different configuration of co-evolutionary pressures. To help us in thinking about these periods, let us refer to them by the challenges they face. The four challenges are: bipedalism, social cohesion, mimetics, and systematization.

Given this analysis, one might ask why we need to worry about the three earlier periods at all. If we just want to understand the evolution of language, could we just look for the major breakthrough adaptation or emergent process of the last 60,000 years and leave the rest of evolution to the physical anthropologists? The problem is that language itself is deeply rooted in the cognitive advances that were achieved during all four periods. One of the major achievements of cognitive linguistics (Givón, 1979; Langacker, 1989) has been the characterization of the extent which language is grounded on cognition. Much of this cognition is shared with our primate cousins (Savage-Rumbaugh & Taglialatela, 2001), but much is uniquely human. If we focus only on recent evolution, we will have a hard time explaining the sudden emergence of human intelligence. By casting our net more broadly across the whole of the last six million years, we can derive a clearer view of the emergence of language from advances in cognition.

Consider the analysis proposed by Tomasello (1999). According to this analysis, the major cognitive difference between apes and humans is that only humans see the actions of conspecifics as involving intentionality. Tomasello holds that this "intentional stance" and its related theory of mind arose during the last 200,000 years, along with the advent

of spoken language. Because Tomasello links this major leap forward in human cognition to the narrow time period of the last 200,000 years, he is forced to conclude that genetic changes alone were inadequate to support language evolution. In fact, he characterizes the brevity of the period available for adaptation as the major evolutionary puzzle that we have to solve when discussing language evolution. While it is certainly true that cultural organization has led to major advances in the last 40,000 years, there is no reason to argue that the human lineage has not had time to explore and consolidate a wide range of cognitive adaptations that support language use. If we view the relevant period as involving 6 million years and not 100,000 years, it is far more reasonable to imagine that there has been a series of fundamental cognitive changes providing support for the emergence of language. By expanding our timeframe, we can begin to analyze the specific conceptual supports that underly the attribution of intentionality to the actions of both objects and conspecifics postulated by Tomasello. These supports arise from abilities that we see in chimpanzees, but only as they have been refined by a further six million years of evolution.

To flesh out this analysis, let us take a closer look at how each of the four periods of co-evolution forced fundamental cognitive and physical adaptations in our predecessors.

1   Bipedalism

The advent of bipedal gait stands as a remarkably clear evolutionary watershed in the late Eocene. Between 10MYA and 7MYA, Africa experienced a major tectonic event that led to the formation of the Great Rift Valley. This valley runs down the center of the continent, dividing the flat jungle of the West from the more arid plain of the East (Coppens, 1999). The rift valley includes high mountains such as Kilimanjaro and vast lakes such as Lake Tanganyika and Lake Nyasa. Connecting these lakes is the Nile River. The rain shadow created by these mountains also produces a major discontinuity in the two parts of sub-Saharan Africa. Before the emergence of the rift, the great apes of Africa thrived throughout the continent in the warm, lush forest of the Eocene that extended across the continent. With the advent of the rift during the Oligocene, the domain of the apes became restricted to the west of the continent, ending at the rift valley. Those apes that found themselves on the east of the rift were forced to adapt to the loss of the jungle. The major competitors of the apes were the monkeys who were better able to hide in the short bush and scramble for pieces of food, while still maintaining a clear arboreal presence. The challenge to the apes was to make use of the new drier savannah habitat in a way that did not compete with the monkeys. The major adaptation of this period was the shift to bipedalism. Instead of scampering about on four legs or even in a knuckle-walk posture, early hominids began to walk on their two hind legs. Many of the species of australopithecines, such as <u>Australopithecus afarensis</u> (Coppens, 1999), maintained a full ability to climb trees, while still capable of walking on two legs on the ground. Other species, such as <u>Australopithecus anamensis</u>, adapted a more exclusively upright gait. The latter group included our direct ancestors.

1.1 Advantages of bipedalism

Although we know that bipedalism was favored in the sparser habitat of the east, we do not understand exactly why it was such a popular and successful adaptation for so many species of primates. One account focuses on the fact that upright posture allows the animal to expose less body surface to the sun. Having less skin exposed to the sun decreases the impact of radiation and the need for cooling. Another factor must be the ability of the primate to use the forelimbs for other purposes, ranging from tool use to flea picking. There is also the defensive effect of greater height. Hunters in the veldt emphasize the importance of always standing upright to convince lions, rhinos, and other animals of one's size to deter possible aggression.

When we compare the great apes to the monkeys, we can see some other pressures that might have favored the move to an upright posture. When chimps and gorillas are together in social groups, they move away from the four-legged knuckle walk position to a sitting position. This allows them to maintain better eye contact with the others in their group. In the sitting position, they do not have to bend their neck up to maintain eye contact. The shift away from knuckle walking to an upright posture continues this emphasis on body positions that allow full eye contact.

Upright posture and full eye contact also provided room for the emergence of the first gestural signals between early hominids. As many have argued, it is likely that hominids went through a period of relying on some forms of gestural communication. It is clear that upright posture provides room for such a development. However, the evolutionary advantage of early gestures may have been overestimated, since first bipedal primates had cognitive resources that were not yet greatly different from those of today's apes. Although we know that apes can learn and transmit a system of signs (Savage-Rumbaugh & Taglialatela, 2001), there is little evidence that the level of sign use they display in natural contexts (Menzel, 1975) would provide any major evolutionary advantage.

The major evolutionary advantage of bipedalism probably involved the fact that the arms could be used to hold sticks and other objects that could be used as tools and weapons (Coppens, 1995). This ability to wield small weapons and tools would have given early hominids a clear advantage over monkeys in their search for food and defense against predators. Tools could be used to dig for roots, open up termite hills, and even catch fish. To take advantage of these opportunities, early hominids quickly developed an opposing thumb that allowed for a stronger and more precise grasp. Because the arms were no longer needed for tree-climbing, they could be committed in this way to object manipulation. In this account, the first co-evolutionary period focused on the development of new systems for control of the hands.

Bipedalism brought with it a series of evolutionary costs. It placed increased mechanical pressure on the neck, the spine, and all the joints of the legs. The reliance on the feet for walking made them less able to function for climbing. The lungs had to adapt to support the breathing needed for running. However, the most important of these costs involved the narrowing of the hips (Hockett & Ascher, 1964). Because early hominids

did not have a significantly enlarged cranium, this was not a problem during the period before 2MYA. However, after that time, brain expansion ran up against the earlier commitment to bipedalism, forcing a series of adaptations in female anatomy, parturition, and child-rearing.

1.2   Neural modifications

The move to bipedalism opened up major cognitive challenges in terms of the control of the hands. Apes already have good control of reaching and basic object manipulation (Ingmanson, 1996). However, with both hands now always free for motion, humans were able to explore still further uses of their hands. Rizzolatti (1996) has shown that monkeys (and presumably also primates) have "mirror" neurons in the supplementary eye fields of premotor cortex that respond with equal force when an action such as "grabbing" is carried out either by the self or by the other, including a human. This mechanism provides a way of equating actions performed by the self with actions or postures performed by the other. These neurons are part of the dorsal visual system (Goodale, 1993) that allows us to represent the postural perspective of another in a way that is isomorphic with but non-identical to our own postural perspective. This dorsal system for motor matching also connects to systems in motor cortex, cerebellum, and hippocampus that represent various postures and movements of the body. These systems then connect to frontal mechanisms for storage and perspective-shifting that provide a way of using our own full body image to perceive the actions of others. At first this system was only partly elaborated. For early hominids, this system provided support for basic actions such as prying open shells, hitting things with clubs, and digging for roots. As hominids increased their ability to control hand motions and grasping actions, they could use the elaborations of this system to link specific actions to the affordances of different objects, as they are used for different purposes. Early hominids could use this ability to imagine various ways of picking up clubs, prying open termite mounds, and using stones to crack shells.

In the model developed by MacWhinney (1999a), this first level of perspective taking involves the shift from direct perception and action to stored mental representations of perceptions and actions. This is the first step in what I call the "ungrounding" of cognition. The perspective-taking model assumes that cognition is fundamentally grounded in the body. However, the evolution of perspective-taking involves the development of methods for escaping the directness of this grounding. Imagery is a clear example of this, since it is grounded on perception and action, but only through the mediation of additional cognitive storage systems. The additional processes that operate on the direct, grounded system achieve ungrounding. The first cognitive adaptations of the hominid line involved the ungrounding of direct actions and direct perceptions through sensori-motor imagery. When we are directly involved in an activity such as grasping a rock, the actions and perceptions form a direct, integrated whole. However, when we watch other human beings or animals grabbing a rock, the vicarious experience produces a secondary activation of the same pathways used for the direct perception. We can think of this as the social route for the ungrounding of direct perception. This social route is important in allowing us to understand the intentions of others and to learn by

imitating their methods. This social route is particularly important during socialization, since it provides the infant with a way of learning skills such as hunting, carrying, building, and courtship.

The second route for the development of ungrounding involves the planning of motor sequences. Once we have developed a system of imagery to track the actions of others, we can use this same system to plan our own future actions. Unlike social imitation, planning must occur without support from the stimulus. To plan the actions involved in chipping an axe, we must be able to call up an image of the desired product, and we must be able to sequence a long series of specific motions that are needed to locate good stones and devise methods for chipping edges. Unassisted planning is relatively fragile in cognitive terms, whereas planning that occurs in a social context is relatively more stable. It is easier to imitate a plan than it is to build that plan without social support.

The movement to an upright posture and a bipedal gait was crucial in the evolution of motor imitation and planning. The ability to construct a planned sequence of actions appears to be a unique property of hominids, as opposed to monkeys and apes. Students of primate tool use (Anderson, 1996; Visalberghi & Limongelli, 1996) have shown that chimpanzees and capuchin monkeys can use tools in a productive and exploratory way. However, they do not appear to make planful use of mental imagery to limit their search through possible methods of tool use. Instead, they apply all directly perceptible methods in hopes that one may succeed.

The full development of this system for perspective-switching was achieved gradually across the full 6 million years of human evolution. During the first stage of this evolution, hominid abilities to imitate motor sequences were probably not much more advanced than the abilities we see in chimpanzees. However, because of the new centrality of the hands during this period, there was continual evolutionary pressure to improve abilities to imitate motor gestures. As we noted above, the first development were probably those that supported the direct social imitation of the manual gestures, body postures, and activities of others.

Early hominids did not simply reuse the neural areas that had previously been committed to manual activities in the arboreal environment. Because the move to a terrestrial environment was quite gradual (Corballis, 1999), hominids needed to provide neural control for the use of the hands in both the arboreal and terrestrial environments. The arboreal environment favors the development of a specific type of motor imagery. Povinelli and Cant (1995) have noted that increases in body weight for larger apes such as orangutans make it important to be able to plan motions through the trees. To do this, the animal needs a map of the self as it executes possible motor actions. The reflexes of this penchant for postural adaptation are still evident in the human enjoyment of dance, dress, exercise, and sport. Donald (1999) makes a similar point in terms of his explication of earlier mimetic plans in hominids. The pressures in the arboreal environment that had favored some limited form of brain lateralization were then carried over to the terrestrial environment (McManus, 1999). This ability to shift quickly between alternative environments required neural support for competing postural and affordance systems.

This postural flexibility may also have allowed some early hominids to adapt partially to an aquatic environment (Morgan, 1997).

Bipedalism also put some pressure on another set of neural mechanisms. Because hominids ceased relying on trees for refuge, and because they were now ranging over a wider territory, they needed to develop improved means of representing spaces and distances. All species must have some way of representing their territory. However, hominids faced the task of representing a rather large territory in which they were both the hunters and the hunted. To do this, they needed to further elaborate earlier mechanisms for spatial encoding. The basic neural mechanisms for this are already well-developed in many mammalian species, including primates (Menzel, 1973), canines, and felines. By linking newly developed systems for body imagery to earlier systems for spatial localization, hominids could begin to plan for the motions of their own bodies through space and time. In addition, hominids could use the primate system of mirror neurons to track not only the posture of conspecifics, but also their movements through space and time. Given the structured nature of primate societies, and the increased mobility brought about from bipedalism, it is likely that some form of spatial perspective-taking developed in early hominids. In terms of the model of MacWhinney (1999a), the second level of degrounding from direct perception involved the development of systems for spatial imagery.

Holloway (1995) has presented evidence from endocasts indicating that there was, in fact, a major reorganization of parietal cortex after about 4MYA. This reorganization involved the reduction of primary visual striate cortex and the enlargement of extrastriate parietal cortex, angular gyrus, and supramarginal gyrus. Much of the evidence for Holloway's analysis comes from traces of the changing positions of the lunate sulcus and the intraparietal sulcus over time. According to Holloway, the areas that were expanded during these changes in the parietal support three basic cognitive functions:
1. Processing in the dorsal (parietal) stream of the visual field is important for representing actions of the other in terms of one's own body image.
2. The association areas of parietal maintain a map of the environment for navigation in the new bipedal mode.
3. The supramarginal gyrus is involved in face perception. Expansion of this area would facilitate the development of social patterns and memory for social relations.

The first two of these functions are directly related to the developments that resulted from the adoption of bipedal gait. The third function may have played a large role during the next period in which the focus of evolution was on social cohesion.

2 Social Cohesion

The shift to bipedalism is clearly documented in the fossil record. However, fossils speak only indirectly about the evolution of primate social structures. Even facts about hominid group size are difficult to determine from the fossil record. We do know, from studies of primate groups, that increases in group size have positive consequences for

protection, particularly of the young. We also know that maintaining larger groups also requires the refinement of methods for social communication and for gathering food.

2.1   Advantages of vocal support for social cohesion

As group size increases, there is increasing conflict for resources such as food, rank, control, and access to females. Dunbar (2000) argues that primates developed a larger neocortex to deal with these tensions. However, it is likely that there was marked pressure on early hominids to maintain large social groups well before we find evidence of a significant expansion in brain size. In particular, between 4.5MYA and 3.5MYA the hominids in East Africa went through an expansion of their range and a proliferation of species. This proliferation was then followed by a period of tight competition with range contraction (Foley, 1999). During this period of range expansion and contraction, our ancestors did not improve their social organization simply by growing larger brains, although brain size did increase a bit, mostly allometrically (Holloway, 1995). Instead, as Nettle and Dunbar (1997) have argued, it is likely that the hominids of the Pliocene consolidated their group structure by a set of targeted neural-behavioral adaptations. Chief among these, I would argue, is the subordination of the vocal system to cortical control.

It is likely that our Pliocene ancestors possessed the same set of vocal calls available to other primates. By linking communication to group solidarity and membership, our ancestors achieved a comparative advantage over their competitors. Cortical control of vocalization also allowed our ancestors to communicate more intentionally over a distance. More importantly, it allowed them to use chatter to gain the attention of their compatriots to negotiate the basics of group relations. The fact that the great apes did not go through a parallel evolutionary process in West Africa can be attributed to the different requirements on group size in their arboreal habitat.

It is likely that one of the side effects of an increase of cortical control over vocalization was the ability of hominid groups to lock in patterns of vocal behavior that characterize the local group, as opposed to the wider hominid community. At first, these local forms of communication would not be sharply defined. However, as the mechanics for vocalization come increasingly under cortical control, it would become easier for a group to differentiate itself from others by unique vocal features. Songbirds achieve this effect through species-specific processes of vocal learning (Konishi, 1995). In birds, dialects allow individuals and groups to maintain their territory against competitors. Hominids could achieve the same effect through differentiation of local patterns for speech and gesture.

2.2   Neural modifications

Achieving neocortical control over the vocal-auditory channel required neuronal reorganization without a major increase in brain size. Ploog (1992) has shown that humans have more direct pyramidal connections between motor cortex and the speech

and vocalization areas of the brain stem than do monkeys. Certain areas of the limbic system, such as the anterior thalamic limbic nuclei, have grown disproportionately large in humans. These nuclei serve the supplementary motor area and premotor and orbital frontal cortex. The expansion of these structures points to increased limbic input to the cortex as well as input from the cortex to the limbic structures. Tucker (this volume) shows that the basic adaptation here involved the absorption of the primate external striatum by the neocortex (Nauta & Karten, 1970).

In macaques (Jürgens, 1979), control of the vocal system relies on the periaqueductal gray matter of the lower midbrain. Additional midbrain regions can stimulate the periaqueductal gray, but the neocortex does not control or initiate primate vocalizations. In man, on the other hand, electrical stimulation of both the supplemental motor area and the anterior cingulate of the frontal cortex can reliably produce vocalization. Primates make few attempts to structure local dialects or otherwise structure their call system through learning (Seyfarth & Cheney, 1999). Yerkes and Learned (1925) and others have tried to condition chimpanzee vocalizations in the laboratory and have failed. Human infants, on the other hand, rely at least in part on highly plastic cortical mechanisms to control vocalization. This allows them to pick up the sound patterns of their community through mere exposure. As a result, each hominid group can build up a local vocal accent that is passed on to the next generation through mere exposure. Other aspects of communication, such as conversational sequencing (Trevarthen, 1984) may be more linked to modeling and imitation. However, this learning of conversational functions only occurs because the child is locked into the interaction by motivational forces that reward face-to-face vocalization (Locke, 1995).

Although primate vocalization is not under cortical control, it has an extremely direct connection to midbrain motivational areas (Pandya, Seltzer, & Barbas, 1988). Human language continues to rely on this underlying limbic architecture to provide emotional coloring to vocalization. As Tucker (this volume) argues, the linkage of the vocal system to limbic mechanisms provides grounding in terms of arousal (brainstem and amygdala), motivation (basal ganglion), patterning (striatal-thalamic ciruits), and memory (limbic circuits). Man also retains some direct links between audition and these limbic circuits, as evidenced in the directness of our responses to sounds such as infant cries or the growls of predators.

The linkage of vocalizations to cortical control allowed our ancestors to distinguish themselves from other hominids. It also allowed them to build up a system of face-to-face social interactions. MacNeilage (1998) has argued that the primate gesture of lip smacking is the source of the core CV syllabic structure of human language. The CV syllable has the same motoric structure as lip smacking. Moreover, it is produced in an area of inferior frontal cortex close to that used for lip smacking and other vocal gestures. Primates use lip smacks as one form of social interaction during face-to-face encounters. However, even bonobos, the most social of all primates, do not maintain face-to-face conversations for the long periods that we find in human interactions. Obviously, one must go beyond a boring repetition of lip smacking to maintain a reasonable level of sustained face-to-face vocal contact. Increased cortical control of vocalization allowed

our ancestors to begin the process of developing these elaborations. By linking its members into tight affiliative relations through face-to-face interaction, our ancestors achieved a form of social organization that allowed them to maintain large social groups for defense against other hominid groups. Other primates have also responded to these pressures by developing a variety of social support mechanisms (de Waal & Aureli, 1996). Other primates have also developed systems for attending to face-to-face interactions and pointing behavior (Gomez, 1996). To maximize the effectiveness of face-to-face interactions, hominids then brought the production of facial gestures under cortical control. As in the case of the control of tool use through motor imagery, humans differ from monkeys in the extent to which the cortex can produce gestures upon demand (Myers, 1976).

In considering the role of face-to-face vocalization in hominid groups, we must not forget the possible divisive role played by aggressive males (Anders, 1994; Goodall, 1979). Hominid groups relied on aggressive males for their skills as hunters and their ability to defend the group against attack. However, groups also needed to provide ways to avoid the direction of male aggression toward other members of the group, particularly other males. We know that primates had already developed various methods for handling these conflicts, including exile for problematic males, the formation of master-apprentice relations, and development of male social groups. Within this already established social framework, males could also benefit from ongoing reaffirmation of their social status through face-to-face chat. By socializing young males into this productive use of language for social cohesion, mothers could also contribute to the stability of the group. Breakdowns in these processes could threaten the survival of the group and even the species.

It is easy to assume that the acquisition of a local pattern of vocalization and gesture would lead to conflict between hominid groups. However, studies of kinship patterns (Fox, 1976) show that societies often place a premium on the practice of outbreeding. In modern societies, this is formalized through rules of cross-cousin marriage or other forms of exogamy. It is possible that early humans developed similar patterns of controlled intermarriage as a way of solidifying relations between groups. However, the fuller articulation of this system probably occurred in later periods.

This account has emphasized the importance of cortical control over the vocal apparatus. I have made no proposals regarding the evolution of auditory abilities in man. The reason for this is that it appears that primates have already achieved a level of auditory processing ability sufficient to support analysis of all vocal communications (Hauser, Newport, & Aslin, 2001; Kuhl & Miller, 1978). There is currently no reason to believe that the human auditory system underwent any major adaptation in the last six million years. The linkage of vocal and facial expression to cortical control may seem like a fairly trivial neurological adaptation. However, it helped our ancestors through this period of intense competition between groups and set the stage for the major changes that were to come in the next period.

3   Mimesis

By 2MYA, <u>homo</u> <u>erectus</u> emerged victorious from the period of intense competition. Recent analysis points to <u>homo</u> <u>ergaster</u>, rather than <u>homo</u> <u>erectus</u> as the direct ancestor of <u>homo</u> <u>sapiens</u>. However, the details of the relations between <u>homo</u> <u>erectus</u> and <u>homo</u> <u>ergaster</u> are not yet clear. In this paper, I will refer to these two related populations as <u>homo</u> <u>erectus</u>, understanding that the details of this particular lineage may soon be revised.

During this period, the species had achieved some level of group solidarity through the social use of vocalization. Beginning sometime before 2MYA, our ancestors were confronted with a third major evolutionary opportunity. Having committed themselves to face-to-face communication, and having elaborated their basic systems of social identification and imitation, the first groups of <u>homo erectus</u> were then able to elaborate new forms of symbolic communication in both vocal and gestural modalities.[1] These new systems involved the conventionalization of vocalizations, postures, and gestures in specific social and pragmatic contexts. This conventionalization must be sufficient to allow for social transmission of newly-minted forms from one generation to the next.

Darwin (1877) thought it unlikely that a system such as language could have emerged from gesture. He believed that this would require the shifting of a function from one organ to another and then back again, as if flying had moved from the wing to the stomach and then back to the wing. However there is no reason to think that Darwin would have excluded the possibility that gesture and vocalization underwent a parallel co-evolution. Today, speech and gesture complement each other during communication (McNeill, 1985). Some messages are conveyed through speech, but others are conveyed through the hands and the posture. Still other messages are conveyed by changes in the tone of voice. It is likely that each of these channels of communication were also available to <u>homo</u> <u>erectus</u> and that they functioned in an interactive and complementary fashion, much as they do now.

3.1   Advantages of Mimesis

Mimetic processes (Donald, 1991) include gesture, signing, chant, song, drama, ritual, and basic forms of costume. Mimesis relies primarily on iconicity and aesthetics to establish communication. It is often difficult for us to imagine a species that relies on this type of communication. Our first intuition would be that these forms are created on-the-fly and vanish after each communication. However, a system of this sort would provide little evolutionary advantage. Instead, mimetic systems must rely on the construction of socially accepted forms. In this way, the early mimetic systems could support group solidarity, planning, and socialization in an increasingly structured way.

The operation of mimetic processes provides strong linkage of the individual to the group. Through mimetic elaborations of chants, songs, whistles, and yells, the group can contrast itself with its competitors. Through chant and dance, the male society can plan

hunts and battles and reenact past struggles. At the same time, mothers can use song, sound play, and chant to hold the attention of their young children while socializing them into the practices of the community. Mimetic processes are linked to the emotional use of language that was elaborated during the previous evolutionary period. As a result, mimetic communications can be used to move people emotionally, preparing them for war, migration, or other major group activities.

It is important to understand the limits of mimesis. Mimesis is not a necessary condition for the execution of cooperative hunting. Felines (lions, tigers), canines (wolves, dogs), and chimpanzees (Goodall, 1979) are able to organize sustained cooperative hunting without chant, dance, or other forms of mimesis. It is also misleading to equate mimesis in homo erectus with the advanced forms of chant and dance found in modern-day tribal societies. These observations suggest that the major contribution of mimesis is its ability to elaborate the details of a shared social plan. It could remind group members of the ways in which animals would be trapped by herding into a closed canyon or even over a cliff. It could help organize the preparation and distribution of food. It could also be used to mark social relations of dominance and allegiance. In general, it would be open to much more concrete and immediate uses than we see in modern-day forms of mimesis.

Mimesis was enormously successful. It allowed homo erectus to expand its territory to all of Africa, eventually leading to the extinction of all other hominid species. It also allowed homo erectus to migrate successfully out of Africa to all parts of Asia, including the Middle East, China, and Indonesia. We should attribute this successful expansion to two basic processes. The first was an ability to eliminate competitors, particularly in Africa. To do this, erectus must have relied on group solidarity as a support for warfare. Second, to support the migration to new territories, erectus must have been extremely adaptive. This adaptivity could not have been grounded in some simple physical change. Instead, it must have resulted from a general improvement in cognitive capacity, particularly as reflected in group problem-solving and adaptation.

MacWhinney (1999a) and Tomasello (1999) have developed parallel accounts of these changes that emphasize the importance of being able to take the perspective of another human being as an intentional agent. In both accounts, this ability is a precondition to the full successful use of language. Tomasello locates the emergence of this ability in the period after 200,000 years ago. However, MacWhinney, like Donald (1991), sees the ability arising beginning nearly 2MYA. Malle (this volume) has asked whether it is possible that theory of mind predates language during evolution. Mimetic communication certainly satisfies Malle's criteria for a level 1 theory of mind, but so might many forms of primate communication. Given that mimesis implies capacities for imitation, joint attention, inferential sensitivity, perspective-taking, and the negotiation of unclear meanings, it would appear to satisfy much of Malle's requirements for a level 2 theory of mind. Thus, during this period, theory of mind and mimetic language underwent an intimate co-evolution. However, language as we currently know it was not yet invented.

3.2   Neuronal Adaptations

The brain size of <u>homo erectus</u> tripled in size during the period between 2MYA and 100,000 years ago. Some of this increase is allometrically related to the overall growth in stature during this period. Thus, it might be better to speak of a doubling in size, rather than a tripling in size. The brain expanded in size because the preconditions for a successful expansion were now all in place. In particular, previous evolution had already produced at least these five abilities:

1. The shift to bipedalism had freed the hands for gestural communication.
2. The freeing of the hands for tool use had led to further elaboration of primate mechanisms, such as "mirror" neurons, for representation of the actions of others.
3. During several millennia of migration and population movements across open ground and through the bush, our ancestors had developed systems for tracking their own spatial positions in the present, past, and future, as well as the spatial positions and perspectives of others.
4. Our ancestors could use their visual system to generate images of past actions and spatial configurations involving themselves and others.
5. During the period before 2MYA, our ancestors had developed a tight linkage of attention to vocalization processes during face-to-face interaction.

These are the preconditions for the evolution of mimesis. To support mimesis, the brain needed to provide complete episodic storage for combinations of gesture and vocalization. It had to store whole chants or gestures as they expressed particular events of importance to individuals and the group. Some of these chants and gestures might be rather extended. For example, there might be a dance that represented the time of ripening of the mongovo nut or one that outlined the components of the hunt for eland. At the same time, these larger mimetic sequences would be composed of smaller pieces that had achieved other cognitive-social grounding. For example, by pointing to a place in the area around the group's encampment, the dancer could signal the direction of a hunt. By holding his arm in a particular stance, the dancer could mime the attack on the game. These component gestures, movements of the body, and vocal chants could then be reused with other components in a relatively unsystematic fashion.

To fully process and interpret a mimetic communication such as a chant with dance, the onlookers would need to represent the perspective of the chanter, much as they now assume the perspective of the speaker of current human language (MacWhinney, 1999a). However, even with the support of perspective-taking, communication still depends on the ability of group members to learn the meaning of new mimetic forms. Thus, mimesis confronts the neural system with a new set of cognitive challenges:

1. The production system must link up stored visual representations to the output processes of chant, gesture, and dance. This linkage of vision to gestural and vocal output requires not only the expansion of both central and peripheral (MacLarnon & Hewitt, 1999) output control areas, but also the expansion of their connections to basic visual areas (Givón, this volume). More generally, control of

this system requires the construction of a cognitive simulation of the human body (MacWhinney, 1999a).
2. Mothers must be able to socialize their children into an understanding of the core mimetic sequences of their own social group.
3. The episodic memory system must store mimetic sequences and their components.
4. As mimetic sequences become elaborated, the brain will need to provide methods for storing whole perspectives, such as that of the hunter, to allow for a switching of perspective, as well as traditional reenactment of these shifts.

Of course, evolution does not respond directly to the changes in behavior that underly these new pressures. Instead, its response is indirect and operates in terms of the principles of population genetics. However, as Dobzhansky (1972) has noted, major evolutionary changes typically arise in response to opportunities opened up by behavioral shifts. In this case, the selective pressure would be for individuals who were able to store the multitude of representations made possible by mimetic communication.

Unlike the evolutionary pressures of earlier periods, the pressures imposed by mimesis cannot be solved simply by linking up older areas or by reusing earlier connections. Instead, the brain must add new computational space to store the multitude of new visual and auditory images (Li, this volume). In addition, the brain needs to expand the role of the frontal areas for storing and switching between perspectives. Because this system grew up in a haphazard way from earlier pieces of lip smacking, pointing, gesture, and rhythm, it would be difficult to extract a core set of elements from mimetic communications, thereby reducing requirements for storage space. Instead, many patterns and forms must be learned and stored as holistic unanalyzed sequences. This Gestalt-like shape of early mimetic patterns corresponds well with the Gestalt-like cognitions that we develop through our interactions with objects. For example, when we chop wood, there is a complete interpenetration of muscle actions, visual experiences, hand positions, and sounds. We can think of this as a single merged form such as I-hands-back-lift-axe-drop-split-chips-wood-cut. Mimetic forms have this same unanalyzed quality. This lack of analysis is not the result of chunking or automatization, since the Gestalt is not constructed by a system of combinatorial semantics. Instead, each chunk is a raw, unanalyzed whole that is fully grounded on direct action and perception. Because they are highly grounded on our direct perceptions and actions, they communicate in a basic way. However, they provide little support for cognitive organization.

The growth of the brain in response to these pressures was so rapid that it is typically assumed that it involves a single genetic mechanism. One such mechanism might involve regulatory genes (Allman, 1999) that control the overall size of the cortex. Changes in the timing of the expression of these genes can lead to the observed across-the-board increase in size for the cortex and cerebellum that we see in homo erectus. However, the expansion of the cortex placed additional adaptive pressures on homo erectus. The bigger brain required a much greater level of caloric intake. This pressure could be met through changes in diet and modifications to the digestive system. A more fundamental pressure was the fact that increases in the size of the infant brain produce problems for the birth process. The width of the hips had narrowed in both men and women as a response to

bipedalism. As long as the skull was not much larger than that found in the primates, this did not cause major problems for the birth process. However, the expansion of the skull in homo erectus ran directly into this evolutionary barrier. To deal with this, the infant is born at a time when it is still fairly immature and the skull is relatively pliable. The increasingly organized shape of the society guarantees the survival of the child. In addition, women have had to sacrifice their ability to run quickly so that the hips could widen, permitting births with larger infant heads. The slowing of infant development not only helps in the birth process, but also helps the child maintain cortical plasticity (Elman, Bates, Plunkett, Johnson, & Karmiloff-Smith, 1996; Julesz & Kovacs, 1995) even into adolescence, thereby further enhancing the ability of the group to construct accepted mimetic patterns.

4   Systematization

By the end of the Pliocene, homo erectus had achieved dominance over its hominid competitors. There were no remaining hominid species in Africa. Presumably homo erectus had eliminated its competitors either through warfare or competition for resources. Homo erectus had also migrated to Eastern Europe, Asia, and Indonesia. However, with the onset of the glaciations of the Pleistocene, our ancestors came under increasing pressure to adapt to the colder, drier environment. It is remarkable to think that this smart primate with its large brain was unable to deal effectively with these new pressures. The contraction of territory back to Africa and then the near extermination of the species at about 70,000 years ago suggest three possibilities. First, it appears that much of the earlier success of homo erectus was achieved at the expense of other hominids. In other words, it was not an expansion of an ecological niche as much as the seizure of a niche that been occupied by various species. Second, in addition to the evolutionary pressure of extreme climate change homo erectus may have had to deal with pressures from diseases that targeted this particular species. Third, the mere expansion of the brain by itself did not achieve a decisive evolutionary advantage.

The fact that the brain expansion was not enough by itself helps us to understand the shape of recent processes in human evolution. The analysis of mitochondrial DNA (Cann, 1995; Cavalli-Sforza, 1991; Templeton, 1992) allows us to trace the lineage of all current humans back to a single population and perhaps a single mother that lived in Africa about 200,000 years ago. By this time, the brain size of homo sapiens had nearly reached its full proportions. Mere increases in brain size were unlikely to achieve any further evolutionary advantage. Instead, humans needed some way to systematize the profuse growth in mimetic processes that had occurred during the Pliocene. We now have a set of fairly well-articulated proposals regarding the nature of this change. The core of the new system involved the introduction of a set of phonological contrasts (Hockett & Altmann, 1973). To achieve accurate articulation of these contrasts, a further set of adaptions were needed for the serial ordering of actions and the precise articulation of sounds. This phonetic revolution occurred gradually over the period from 200,000 to 50,000 years ago.

4.1   Advantages of systematization

Archeological and linguistic research has provided a clear picture of the advantages of phonological and lexical systematization. By coding words into a compact set of contrastive features, homo sapiens was able to build up a lexicon that can name and encode all of the important objects, properties, and actions in our environment. Having a full inventory of the physical world allowed early humans to use different animal and plant species for increasingly refined purposes. Language was used to pinpoint the location of quarry sites for flints and other materials for stone weapons and tools. Master craftsmen used language to teach apprentices how to chip points, prepare hides for tanning, carve out wooden bowls, bind axes to poles with sinew, sew hides into tents, and tame pets. As tool-making progressed, language was used to explain how, where, and when to plant and water seeds. By codifying these times and practices in verse and religion, the first agricultural systems were developed and shortly thereafter the first settled villages. From these roots, came the civilizations of the ancient Middle East, Egypt, and China with their invention of writing, organized warfare, chariots, metal-making, government, and increasingly formalized religion. Modern society has elaborated on this foundation with the creation of cities, books, bombs, law, medicine, and the Internet. All of these developments are consequences of the introduction of systematization for phonology and lexicon. However, this recent expansion would not have been possible without the major cognitive modifications of the full 6 million years of human evolution.

4.2   Neuronal adaptations

Some of the adaptations required for smooth vocal production are quite peripheral. (Lieberman, 1973), involving changes to the vocal tract, the structure of the larynx, muscle innervation, tongue support, and facial musculature. Some of these changes were underway before the Pleistocene; others have been more recent. To control this additional external hardware, the brain has needed to fine-tune its mechanisms for motor control. This fine-tuning does not require the type of brain expansion that occurred in homo erectus. Instead, it involves the linking of inferior frontal areas for motor control to temporal areas (Gabrieli, Brewer, Desmond, & Glover, 1997) for sequence storage. These linkages (Damasio & Damasio, 1988) involve pathways that lie under the central sulcus. They constitute a functional neural circuit that implements a phonological loop for learning new words (Gupta & MacWhinney, 1997). The auditory shapes of words are stored in topological maps (Miikkulainen, 1990) in superior temporal auditory cortex and can be associated to visual images in inferior temporal areas. This linkage of the vocal-auditory channel to the visual channel further develops binding the entrainment of the vocal-auditory channel by the visual channel (Givón, this volume).

Once homo sapiens had achieved an ability to produce, store, and learn a large vocabulary of phonologically organized forms, the remaining steps in the evolution of language were comparatively easy. Humans had already achieved a mimetic system for perspective-taking and perspective-switching. This system allowed listeners to mentally reenact the motions, rhythms, and chants of the speaker as they depicted movement

between places and actions on objects. Once words became available, speakers and listeners could parse these single-package gestalt-like communications into their components. With words to name specific objects and participants, it was possible to separate out nouns from verbs. This adaptation to grammar required no particular new cognitive skill for nouns. However, for predicates such as verbs, it was important to store linkages between the overall configuration of the action and the specific uses with participants. In other words, children had to learn how to manage language in terms of item-based syntactic constructions (MacWhinney, 1975, 1982), including "verb islands" (Tomasello, 2000). Neuronal processes for this level of control involve little in the way of new evolution. However, they place storage demands on the pathways between the temporal lexical areas and the frontal planning and sequencing areas.

As speakers build up longer and longer strings of propositions, they rely increasingly on frontal areas, such as dorsolateral prefrontal cortex (DLPFC) for the storage of one perspective that allows shifting to a secondary perspective. Shifts of this type are central in the processing of anaphors and gaps in argument structure. As MacWhinney (2000) has shown, these various syntactic processes are grounded not on the construction of abstract syntactic trees, but on the direct processing of embodied perspectives of the type that were also important during the period of mimetic communication.

4.3   Cultural evolution

Given the minimal nature of the additional adaptations needed to support human language, why did the human species suddenly become so successful? The reason for this great success is that these minimal changes were crucial in triggering a reuse of previous adaptations. With the onset of phonological systematization, humans were able make a much fuller use of the massive expansion in brain size. They did this by constructing a linguistic system that uses the entire brain to represent experience. This was a social construction of the last 60,000 years, not a process grounded on the slow timetable of biological evolution.

Language builds systematically on each of the achievements of the six million years of evolution.
1.   It uses the basic primate sensorimotor systems of posterior cortex to encode objects in terms of direct perceptions and properties.
2.   It uses the dorsal systems developed in the late Miocene for storing mental representations of action sequences on objects. This system supports the representation of concrete nouns, adjectives, and intransitive verbs.
3.   It uses the hippocampal and parietal systems developed in the late Miocene for encoding navigation paths and the spatial location of others. This system supports the representation of deictic terms, prepositions, and locative adverbs.
4.   It uses the system developed during the Pliocene for the planning of mimetic sequences. MacWhinney (2000) views this ability as forming the core of the grammatical system of transitivity.

5.   It uses the perspective-taking involved in mimetic communication as the core of perspective-taking in language. Systems for expressing these shifts include complementation, embedding, binding and anaphora, factivity, evidentiality, and marking for politeness and kinship.

Language is grounded entirely on these achievements of the past 6 million years. However, once language is available as a method of cognitive representation, cognition itself is fundamentally altered. The most remarkable property of human language is that it has smooth and nearly immediate access to the entire brain. Through this ability to integrate across modalities and modules, language is able to overcome modularity and open up the mind to fully human consciousness. Language relies on the entire brain to achieve its complete cognitive simulation of experience in terms of objects, space, action, and social relations. Because it integrates these separate modules so thoroughly, it allows us to fully escape the modularity that is present in primates (Russon & Bard, 1996) and young children (Hermer-Vazquez, Moffet, & Munkholm, 2001). Without language, it may be possible to focus directly on the position of an object without regard to earlier orientations or the orientations of others. Without language, we can focus on an action without breaking it apart into its component participants. In order to achieve lexicalization, language forces us to analyze experience into its components. Although language forces us to break our experiences into pieces, it provides ways of then recombining these pieces into much larger edifices. Moreover, narrative and discourse allow us to integrate our own experiences more fully with those of others. In this way, language bundles the whole of mental life into a single, more fully conscious, but relatively less grounded whole.

Because language is a human invention, the brain provides us with a great deal of latitude in the way it can be represented. As Wittgenstein (1953) observed, language is a like a well-pruned hedge. Externally, each hedge must look like the others. However, internally each hedge can have a very different shape. This means that we should not be surprised to find large individual differences in the neuronal basis of higher-level dynamic control for language. For example, Children with large focal lesions to the left hemisphere areas that typically control language are able to achieve normal language functioning by using parallel areas in the right hemisphere (Booth et al., 2000). Several patterns of reorganization to damage have been identified for young children (MacWhinney, Feldman, Sacco, & Valdes-Perez, 2000) and similar reorganization certainly occurs in adults (Holland et al., 1985). In addition to the flexibility found in patients with brain lesions, there are important individual differences in the way that the brain organizes for language in normally-developing speakers. In particular, it is likely that the process of cerebral lateralization operates in very different ways in different children, with some making a sharp separation between the functions of the two hemispheres (Beeman et al., 1994) and others allowing for more redundancy.

4.4   Social-developmental consequences

Although the emergence of language in humans places only modest demands on additional neuronal evolution, it had major consequences for the way in which we use our

brains. Vygotsky (1962) showed how language begins as an external social phenomenon and then is internalized to provide the backbone of human cognition. Vygotsky attributed this internalization to the emergence of "inner speech" and compressed forms of reference and topicalization, but provided little additional detail regarding his proposal. Tomasello (1999) extended Vygotsky's analysis by linking language learning to the child's ability to treat others as intentional agents. This assumption of an intentional stance helps the child follow the meaning of the parent when learning new words. Tomasello also shows how intentionality allows the child to construct the representations underlying causal predications. Infants will extend intentionality even to inanimate objects, when they have eyes that represent a face (Gergely, Nádasdy, Csibra, & Bíró, 1997) or even when they do not but appear to have intention (Csibra, Gergely, Bíró, Koos, & Brockbank, 1999). Later, through fictive extensional processes (Lakoff, 1987; Talmy, 1988), children treat even inanimate objects as acting intentionally and causally. This system of causal representation forms the basis for grammatical processes such as case marking, conflation, passivization, inverse, and causitivization.

MacWhinney (1999a) shows how the construction of the intentional perspective is a preliminary to a wider expansion of language and perspective-taking to incorporate the specific linguistic propositions needed for an articulated theory of mind. In computing the details of social relations, we often need to take the perspective of the other and imagine how that person will react to a variety of possible actions that we might perform. Typically, we select the exact shape of a given utterance or action out of a much larger field of potential actions based largely on this act of social perspective-taking. Would our utterances offend some particular individual? Would they violate some moral precept? Would we fail to properly articulate our position and therefore leave ourselves open to later misinterpretation? As we compute these various possible scenarios, we use the three lower levels of perspective-taking (direct perception, space-time, causal action) to move about in a conceptual space that was constructed through linguistic interactions embedded in a social context. These processes allow us to construct enormously complex systematic (Levi-Strauss, 1963) views of social structure. Yet the cognitive resources we are using are little more extensive than those available to homo erectus. Instead, through the construction of a method for creating new words, we have been able to reuse our mimetic resources to support the full articulation of human culture. In this sense, the appearance of modern language is a paradigm case of an emergent behavior (MacWhinney, 1999b).

4.5   Disorders of Communication

A good understanding of language evolution should be able to illuminate issues in communicative disorders. In particular, we might expect that the most recent adaptations would be the ones that are least fully integrated into the human genome. One clear prediction of the current analysis is that the most recent adaptations for language involve motor control for phonology. This view is highly compatible with the fact that we see a series of language disorders such as:
1. stuttering: a process that involves problems in the activation and coordination of output processes based on individual lexical items,

2. dysphagia: problems with chewing and swallowing that may reflect some instability in the rewiring of innervation of the vocal tract,
3. articulation disorders: between the ages of 4 and 8, many children have problems articulating the exact shapes of words. Disorders such as lisping sometimes continue into adulthood.
4. tongue curling: about 30% of the adult population cannot perform the type of tongue curling or the type of tongue bending needed to properly articulate sounds such as the retroflex stops of Hindi. There are similar problems with the articulation of one of the three Czech r-sounds.

There is good reason to believe that these disabilities represent incomplete consolidation of recent evolutionary changes. If we then further parcel out cases of mental retardation[2], autism[3], fetal damage, and chromosomal abnormalities[4], we are left with a group of children who are said to have specific language impairment (SLI). The incidence of some form of specific language impairment in the population is often estimated to be about 7%. In the clinic, language disorders are nearly four times more frequent in males than in females (Bishop, 1997). However, epidemiological studies have shown that the actual balance in the population is nearly equally distributed between the sexes (Leonard, 1998). We can divide the overall category of specific language impairment into three subtypes:
1. expressive language disorders,
2. pragmatic disorders,
3. grammatical disorders

The category of expressive language disorders is probably just a further project of the group of motor speech disorders we mentioned above. However, this group has even deeper problems with motor output and formulation. The instability of these skills is, again, evidence for the recency of these evolutionary developments. The existence of a class of children with pragmatic disorders indicates that some aspects of perspective-taking are still variable in the human species. The account offered above suggests that these abilities were consolidated long ago during evolution. It is likely that children with pragmatic disorders are actually reflecting incomplete consolidation of the more recent evolutionary changes in frontal lobe memory processes used to store and switch perspectives.

Finally, we turn to a consideration of the evolutionary status of children with both expressive and receptive disorders, including grammatical disorders[5]. Recent studies (Franks & Connell, 1996; van Der Lely & Stollwerk, 1997; van der Lely & Stollwerk, 1996) have suggested that children with grammatical disorders have a specific problem with argument chains. However, it may be that these problems are based on representational deficits, rather than any specifically grammatical problem. Some recent accounts (Merzenich et al., 1996; Tallal et al., 1996) have attempted to link specific language impairment to an underlying auditory deficit. Given the fact that it is difficult to find any clear structural or functional differences between human and primate auditory processing or learning (Hauser et al., 2001), it would seem that raw auditory processing is an unlikely candidate for a weak link in the evolutionary chain. Grammatical SLI may reflect difficulties in the representation of lexical items (Chiat, 2001; Gupta &

MacWhinney, 1997) or problems in the storage of chains of lexical items for syntactic combination.

The patterns of population dispersion that occurred during the last 100,000 years might well have an impact on the relative frequency of language disorders in different human populations. For example, there is now evidence that Cro-Magnon groups intermarried with Neanderthal groups in both Europe and the Middle East. It is generally accepted that Neanderthal people, although derivative from the basic human stock, had a relatively low level of cultural innovation and perhaps a low level of vocal control (Lieberman, 1973). Any intermarriage between groups with greater linguistic ability and groups with less linguistic ability could lead to a population that included individuals with some form of language impairment or limitation. Currently, our information on the actual distribution of language disorders in different parts of the world is so incomplete that we cannot evaluate these patterns. Moreover, languages with different structures would put different strains on children's ability to learn and produce language (Leonard, 1998).

If language is such an important aspect of human existence, then how can it be that such a sizeable portion of the population has language impairment? We know that adults with language impairments are less successful in finding mates (Leonard, 1998). Given this, it is likely that the genes selecting for specific language impairment are linked chromosomally to other genes that provide compensatory adaptive advantages. However, we have currently no ideas about what these additional characteristics might be.

5 Conclusion

The study of language evolution has advanced enormously in recent years. New evidence from the fossil record, paleoclimatology, genetic analysis, neuroscience, infancy research, and cognitive grammar have fueled these advances. As the database of evidence regarding man's last six million years continues to grow, we will be able to articulate increasingly precise ideas about the co-evolution of language, brain, and social processes. In this paper, I have presented one fairly detailed account that takes into account the complete set of adaptations that occurred across the full range of six million years. By casting a broad net across this whole period, we can see more clearly how modern language relies on a set of adaptations of which some are shared with our primate cousins, some are ancient hominid inventions, and some have emerged in our recent past.

References

Alcock, K. J., Passingham, R., Watkins, K. E., & Vargha-Khadem, F. (2000). Oral dyspraxia in inherited speech and language impairment and acquired dysphasia. *Brain and Language, 75*, 17-33.

Anders, T. (1994). *The origins of evil: An inquiry into the ultimate oringins of human suffering*. Chicago: Open Court.

Anderson, J. (1996). Chimpanzees and capuchin monkeys: Comparative cognition. In A. E. Russon & K. A. Bard & S. T. Parker (Eds.), *Reaching into thought: The minds of the great apes* (pp. 23-56). Cambridge: Cambridge University Press.

Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker's direction of gaze strategy to crack the code of language? *Child Development, 68*, 48-57.

Bates, E., & Goodman, J. C. (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (Ed.), *The emergence of language* (pp. 29-80). Mahwah, NJ: Lawrence Erlbaum Associates.

Beeman, M., Friedman, R. B., Grafman, J., Perez, E., Diamond, S., & Lindsay, M. B. (1994). Summation priming and coarse coding in the right hemisphere. *Journal of Cognitive Neuroscience, xx*, 26-45.

Bishop, D. (1997). *Uncommon understanding*. Hove, UK: Psychology Press.

Booth, J. R., MacWhinney, B., Thulborn, K. R., Sacco, K., Voyvodic, J. T., & Feldman, H. M. (2000). Developmental and lesion effects during brain activation for sentence comprehension and mental rotation. *Developmental Neuropsychology, 18*, 139-169.

Cann, R., L. (1995). Mitochondrial DNA and human evolution. In J.-P. Changeux & J. Chavaillon (Eds.), *Origins of the human brain*. Oxford: Clarendon.

Cavalli-Sforza, L. (1991). Genes, people, and languages. *Scientific American, November*, 104-110.

Chiat, S. (2001). Mapping theories of developmental language impairment: Premises, predictions, and evidence. *Language and Cognitive Processes, 16*, 113-142.

Coppens, Y. (1995). Brain, locomotion, diet, and culture: how a primate, by chance, became a man. In J.-P. Changeux & J. Chavaillon (Eds.), *Origins of the human brain*. Oxford: Clarendon Press.

Coppens, Y. (1999). Introduction. In T. G. Bromage & F. Schrenk (Eds.), *African biogeography, climate change, and human evolution*. New York: Oxford University Press.

Corballis, M. C. (1999). Phylogeny from apes to humans. In M. C. Corballis & S. E. G. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution* (pp. 40-70). Oxford: Oxford University Press.

Csibra, G., Gergely, G., Bíró, S., Koos, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of pure reason in infancy. *Cognition, 72*, 237-267.

Damasio, A. R., & Damasio, H. (1988). The neuroanatomical correlates of aphasia and the understanding of the neural substrates of language. In L. M. Hyman & C. N. Li (Eds.), *Language, Speech and Mind - Studies in honor of Victoria A. Fromkin*. London: Routledge & Kegan Paul.

Darwin, C. (1877). A biographical sketch of an infant. *Mind, 2*, 292-294.

de Waal, F. B. M., & Aureli, F. (1996). Consolation, reconciliation, and a possible cognitive difference between macaques and chimpanzees. In A. E. Russon & K. A. Bard & S. T. Parker (Eds.), *Reaching into thought: The minds of the great apes* (pp. 80-110). Cambridge: Cambridge University Press.

Deacon, T. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: Norton.

Dobzhansky, T. (1972). Genetics and the diversity of behavior. *American Psychologist, 27*, 523-530.

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

Donald, M. (1999). Preconditons for the evolution of protolanguages. In M. C. Corballis & S. E. G. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution* (pp. 138-154). Oxford: Oxford University Press.

Dunbar, R. (2000). Causal reasoning, mental rehearsal, and the evolution of primate cognition. In C. Heyes & L. Huber (Eds.), *The evolution of cognition*. Cambridge, MA: MIT Press.

Elman, J., Bates, E., Plunkett, K., Johnson, M., & Karmiloff-Smith, A. (1996). *Rethinking innateness*. Cambridge, MA: MIT Press.

Foley, R. (1999). Evolutionary geography of Pliocene African hominids. In T. G. Bromage & F. Schrenk (Eds.), *African biogeography, climate change, and human evolution*. New York: Oxford University Press.

Fox, R. (1976). *Kinship and marriage: An anthropological perspective*. Cambridge: Cambridge University Press.

Franks, S. L., & Connell, P. J. (1996). Knowledge of binding in normal and SLI children. *Journal of Child Language, 23*, 431-464.

Gabrieli, J. D. E., Brewer, J. B., Desmond, J. E., & Glover, G. H. (1997). Separate neural bases of two fundamental memory processes in the human medial temporal lobe. *Science, 276*, 264-266.

Gauger, L. M., Lombardino, L. J., & Leonard, C. M. (1997). Brain morphology in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*, 1272-1284.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1997). Taking the intentional stance at 12 months of age. *Cognition, 56*, 165-193.

Givón, T. (1979). *On understanding grammar*. New York: Academic Press.

Givón, T. (1998). On the co-evolution of language, mind and brain. *Evolution of Communication, 2*, 45-116.

Gomez, J. C. (1996). Ostensive behavior in great apes: The role of eye contact. In A. E. Russon & K. A. Bard & S. T. Parker (Eds.), *Reaching into thought: The minds of the great apes* (pp. 131-151). Cambridge: Cambridge University Press.

Goodale, M. A. (1993). Visual pathways supporting perception and action in the primate cerebral cortex. *Current Opinion in Neurobiology, 3*, 578-585.

Goodall, J. (1979). Life and death at Gombe. *National Geographic, 155*, 592-620.

Gopnik, M. (1990). Feature blindness: A case study. *Language Acquisition, 1*, 139-164.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59*, 267-333.

Hauser, M., Newport, E., & Aslin, R. (2001). Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition, 78*, B53-B64.

Hermer-Vazquez, L., Moffet, A., & Munkholm, P. (2001). Language, space, and the development of cognitive flexibility in humans: The case of two spatial memory tasks. *Cognition, 79*, 263-299.

Hewes, G. W. (1973). Primate communication and the gestural origin of language. *Current Anthropology, 14*.

Hockett, C., & Ascher, R. (1964). The human revolution. *Current Anthropology, 5*, 135-167.

Hockett, C. F., & Altmann, S. A. (1973). A note on design features. In T. A. Sebeok (Ed.), *Animal communication*. Bloomington, Indiana: Indiana University Press.

Holland, A., Miller, J., Reinmuth, O., Bartlett, C., Fromm, D., Pashek, G., Stein, D., & Swindell, C. (1985). Rapid recovery from aphasia: A detailed language analysis. *Brain and Language, 24*, 156-173.

Holloway, R. (1995). Toward a synthetic theory of human brain evolution. In J.-P. Changeux & J. Chavaillon (Eds.), *Origins of the human brain*. Oxford: Clarendon Press.

Ingmanson, E. J. (1996). Tool-using behavior in wold Pan paniscus: Social and ecological considerations. In A. E. Russon & K. A. Bard & S. T. Parker (Eds.), *Reaching into thought: The minds of the great apes*. New York: Cambridge University Press.

Julesz, B., & Kovacs, I. (Eds.). (1995). *Maturational windows and adult cortical plasticity*. New York: Addison-Wesley.

Jürgens, U. (1979). Neural control of vocalization in nonhuman primates. In H. D. Steklis & M. J. Raleigh (Eds.), *Neurobiology of social communication in primates*. New York: Academic Press.

Konishi, M. (1995). A sensitive period for birdsong learning. In B. Julesz & I. Kovacs (Eds.), *Maturational windows and adult cortical plasticity* (pp. 87-92). New York: Addison-Wesley.

Kuhl, P. K., & Miller, J. D. (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *Journal of the Acoustical Society of America, 63*, 905-917.

Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago: Chicago University Press.

Langacker, R. (1989). *Foundations of cognitive grammar. Vol. 2: Applications*. Stanford: Stanford University Press.

Leonard, L. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Levi-Strauss, C. (1963). *Structural anthropology*. New York: Basic Books.

Lieberman, P. (1973). On the evolution of language: A unified view. *Cognition, 2*, 59-94.

Locke, J. L. (1995). Development of the capacity for spoken language. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language*. Oxford: Basil Blackwell.

MacLarnon, A., & Hewitt, G. (1999). The evolution of human speech. *American Journal of Physical Anthropology, 109*, 341-363.

MacNeilage, P. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences, 21*, 499-546.

MacWhinney, B. (1975). Pragmatic patterns in child syntax. *Stanford Papers And Reports on Child Language Development, 10*, 153-165.

MacWhinney, B. (1982). Basic syntactic processes. In S. Kuczaj (Ed.), *Language acquisition: Vol. 1. Syntax and semantics* (pp. 73-136). Hillsdale, NJ: Lawrence Erlbaum.

MacWhinney, B. (1999a). The emergence of language from embodiment. In B. MacWhinney (Ed.), *The emergence of language* (pp. 213-256). Mahwah, NJ: Lawrence Erlbaum.

MacWhinney, B. (2000). Perspective-taking and grammar. *Japanese Society for the Language Sciences, 1*, 1-25.

MacWhinney, B. (Ed.). (1999b). *The emergence of language*. Mahwah, NJ: Lawrence Erlbaum Associates.

MacWhinney, B., Feldman, H. M., Sacco, K., & Valdes-Perez, R. (2000). Online measures of basic language skills in children with early focal brain lesions. *Brain and Language, 71*, 400-431.

McManus, I. C. (1999). Handedness, cerebral lateralization, and the evolution of language. In M. C. Corballis & S. E. G. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution* (pp. 194-217). Oxford: Oxford University Press.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review, 92*, 350-371.

Menzel, E. (1973). Chimpanzee spatial memory organization. *Science, 182*, 943-945.

Menzel, E. (1975). Purposive behavior as a basis for objective communication between chimpanzee. *Science, 189*, 652-654.

Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science, 271*, 77-81.

Miikkulainen, R. (1990). A distributed feature map model of the lexicon, *Proceedings of the 12th Annual Conference of the Cognitive Science Society* (pp. 447-454). Hillsdale, NJ: Lawrence Erlbaum Associates.

Morgan, E. (1997). *The aquatic ape hypothesis*. London: Souvenir Press.

Myers, R. E. (1976). Origins and evolution of language and speech. *Annals of the New York Academy of Sciences, 280*, 745-757.

Nauta, W. J. H., & Karten, H. J. (1970). A general profile of the vertebrate brain, with sidelights on the

ancestry of cerebral cortex. In G. C. Quarton & T. Melnechuck & G. Adelman (Eds.), *The Neurosciences* (pp. 7-26). New York: Rockefeller University Press.

Nettle, D., & Dunbar, R. (1997). Social markers and the evolution of reciprocal exchange. *Current Anthropology, 38*, 93-99.

Pandya, D. P., Seltzer, B., & Barbas, H. (1988). Input-out organization of the primate cerebral cortex. In H. Steklis & J. Irwin (Eds.), *Comparative primate biology: Neurosciences*. New York: Liss.

Plante, E. (1996). Phenotypic variability in brain-behavior studies of specific language impairment. In M. L. Rice (Ed.), *Toward a genetics of language*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ploog, D. W. (1992). Neuroethological perspectives on the human brain: From

the expression of emotions to intentional signing and speech. In A. Harrington (Ed.), *So human a brain: Knowledge and values in the neurosciences* (pp. 3-13). Boston: Birkhauser.

Povinelli, D. J., & Cant, J. G. H. (1995). Arboreal clambering and the evolution of self-conception. *Quarterly Journal of Biology, 70*, 393-421.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*, 131-141.

Russon, A. E., & Bard, K. A. (1996). Exploring the minds of the great apes: Issues and controversies. In A. E. Russon & K. A. Bard (Eds.), *Reaching into thought: the minds of the great apes*. New York: Cambridge University Press.

Savage-Rumbaugh, E., & Taglialatela, J. (2001). Language, apes, and understanding speech. In T. Givón (Ed.), *The evolution of language*.

Seyfarth, R., & Cheney, D. (1999). Production, usage, and resonse in honhuman primate vocal development. In M. Hauser & M. Konishi (Eds.), *The design of animal communication*. Cambridge, MA: MIT Press.

Stringer, C., & McKie, R. (1996). *African exodus*. London: Pimlico.

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., Schreiner, C., Jenkins, W. M., & Merzenich, M. M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science, 271*, 81-84.

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science, 12*, 59-100.

Templeton, A. R. (1992). Human origins and analysis of mitochondrial DNA sequences. *Science, 255*, 737.

Tomasello, M. (1999). *The cultural origins of human communication*. New York: Cambridge University Press.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*, 209-253.

Trevarthen, C. (1984). Biodynamic structures, cognitive correlates of motive sets and the development of motives in infants. In W. Prinz & A. F. Sanders (Eds.), *Cognition and motor processes*. Berlin: Springer.

van Der Lely, H., & Stollwerk, L. (1997). Binding theory and grammatical specific language impairment in children. *Cognition, 62*, 245-290.

van der Lely, H. K. J., & Stollwerk, L. (1996). A grammatical specific language impairment in children: An autosomal dominant inheritance? *Brain and Language, 52*, 484-504.

Visalberghi, E., & Limongelli, L. (1996). Acting and undertanding: Tool use revisited through the minds of capuchin monkeys. In A. E. Russon & K. A. Bard & S. T. Parker (Eds.), *Reaching into thought: The minds of the great apes* (pp. 57-79). Cambridge: Cambridge University Press.

Vygotsky, L. (1962). *Thought and language*. Cambridge: MIT Press.

Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Blackwell.

Wood, B., & Collard, M. (1999). The human genus. *Science, 284*, 65-71.

Yerkes, R. M., & Learned, B. W. (1925). *Chimpanzee intelligence and its vocal expressions*. Baltimore: Williams and Wilkins.

---

[1] As Hewes (1973) and many others have argued, there is reason to believe that early humans went through a period of heavy use of gestural communication. Some have suggested that the freeing of the hands by the shift to bipedalism should lead directly to the use of gesture. However, in cognitive terms, our ancestors from 6MYA were still close to the ancestors of the great apes. Their systems of natural gestural communication might well have been parallel to those observed in today's chimpanzees (Menzel, 1975). Thus, it seems to me that the real flowering of gestural communication probably did not occur until this third period, when systems for perspective-taking, sustained attention, imitation, and planning could make better use of conventionalized gestures.

[2] Many children with language impairments also have some organic brain problem. These difficulties can include smaller brain sizes (Gauger, Lombardino, & Leonard, 1997), lower neural activation (Plante, 1996), early brain lesions (MacWhinney et al., 2000), traumatic brain injury, hydrocephalus, and epilepsy.

[3] Some authors have claimed that autism and Landau-Kleffner's syndrome represent disorders in the biology of a cognitive module supporting theory of mind (Baron-Cohen, Baldwin, & Crowson, 1997). Although some linkage of this type seems likely and would fit in well with the current theory, the embryological and genetic facts about the origin of autism cast some doubt on any theory that views it as a unitary disorder with a single consequence.

[4] Another set of developmental problems is based not on specific motor disabilities, but on malformations that occur during fetal development from various imbalances. These include cleft palate and cerebral palsy. We also know that the processes of meiosis and mitosis are vulnerable to basic errors in translocation and nondisjunction in all species. These processes might also impact the use of language, but would not represent any pattern with a clear status in recent evolution. The most important syndromes here are Down Syndrome, Williams Syndrome, Turner's syndrome, and Kleinfelter's syndrome. These genetic abnormalities have important impacts on language development, but they cannot be linked in any direct way to the evolution of language. For example, Bates and Goodman (1999) have shown how we can compare normal language development with language development in children with Down and Williams syndrome. This comparison underscores the linkage of grammar to the lexicon in the first stages of language learning. However, this type of comparison does not tell us whether grammar and the lexicon developed in parallel across evolution.

[5] One study of a family in London (Gopnik, 1990) has shown a pattern of disorders consistent with autosomal dominant inheritance. Affected family members have problems forming regular past tense forms such as "jumped," but can produce irregulars such as "fell." However, this disorder involves not only language, but also chewing, swallowing, and a variety of other behaviors (Alcock, Passingham, Watkins, & Vargha-Khadem, 2000).