1

Introduction: The Beginning of Words

If alien visitors were to come to earth, there is little doubt that we would be the first species they noticed as they approached from space. After all, we are the only animals that have managed to leave our planet and adventure into outer space, sending robots to other planets and placing a probe on the surface of a nearby comet. Few if any changes in the history of life have been as radical as the ones we are imposing with our technological capacity, changing the shape of the biosphere in a geological instant. Thus, there is no doubt that our species is vastly different from all others in the ability to alter the environment for our immediate benefit. Furthermore, this is largely due to our unique ability to communicate through language. Language enables a mental or semantic space that we share with others and helps us to coordinate our behavior, anticipate the future, describe the world around us, imagine utopic scenarios, and manipulate our surroundings. Language is expressed in a variety of forms, the most obvious being speech, but we also use language for reading and writing, and some use sign language. Some claim that language first arose in the form of hand gestures that were later overtaken by the elaboration of speech, while others, including myself, are more comfortable with the notion that speech was the first

F. Aboitiz, *A Brain for Speech*,
DOI 10.1057/978-1-137-54060-7_1

way to express modern language. In any case, our capacity to communicate has made us perhaps the most successful animal species, which is the signature of biological adaptation. Our present nature is inseparable from our language. How our ancestors came to acquire language is therefore a fundamental evolutionary and social question that touches on our very nature and, as I believe, can give us useful information on how to survive as a species.

This book is concerned with the key neurobiological steps that allowed us to start the language explosion that changed our lives forever, which for all we know does not show signs of having ended. It is commonly said that we are genetically 98.6% similar to our closest relative, the chimpanzee, so this extraordinary impulse must have been caused by only a few mutations that reorganized our brains and our capacity to understand the world, to communicate with others, and to manipulate our environment. These changes may have been caused by classical genetic mutations, possibly in so-called regulatory genes that work as master organizers of large-scale developmental processes. Mutations in these genes may have been important in producing rapid changes in the overall structure of neural networks or in rapidly increasing brain size. One example of these genes may be FOXP2, whose mutations have been associated with certain forms of speech disorders. In addition, recent research has called attention to epigenetic modifications that are acquired but lasting alterations in the patterns of expression of some genes. At the behavioral level there are cultural modifications that can influence the plastic development of the brain, producing connectional rearrangements in ways we do not yet completely understand, and which may have contributed to rewiring our brains for language. One intriguing possibility is whether epigenetic mechanisms are influenced by cultural transmission. Although research on these lines is fundamental, one problem with genetic studies of language capacity is that they do not tell us what the phenotype is, or precisely which anatomical and functional characteristics allowed us to develop speech, and eventually language. In my opinion, this is the most critical question of language origins, and all others, including genetic, cultural or linguistic accounts, will eventually have to be subordinated to an explanation of how our brains construct language.

Before continuing this discussion, it will be useful to clarify some basic definitions, so that we can agree about the terms I will use. First, what is language? Human language differs from other forms of animal and human communication in its internal structure, which is organized in several components. Human language has syntax or grammar, a lexicon or set of words, and semantics or meaning. For modern linguists and other specialists, language is not equivalent to speech. Speech is a particular way to express language, as there are also sign and written languages. Furthermore, spoken language has phonology, which refers to the articulation of different sounds (phonemes) to make up larger meaningful units (morphemes, that are words or parts of words), and prosody, which relates to the intonation patterns and emotional contents we transmit during speech. Other forms of language, like sign language, have an equivalent of phonology and prosody, while written language relies on auditory representations that contain phonological and some prosodic features. Finally, associated with language are other elements like its pragmatics, or the social context in which language operates, and related cognitive abilities like mathematics and of course music, which Charles Darwin said was closely related to speech in its origins.

While the more abstract notion of language has been the subject of interest for most linguists and some biologists, in this book I will focus rather on speech. The latter is an observable behavior that includes clear functional and morphological adaptations, making it more amenable to a biological approach. Furthermore, in recent years there has been a tendency (with some notable exceptions) to downplay the importance of speech in language origins in favor of hypotheses that consider speech as a secondary achievement. Throughout this book, I will speak of language when referring to more abstract aspects of human communication, and to speech when referring to the specific sensorimotor system involved in speech production. Sometimes, I will refer to language and speech to emphasize that I am speaking of both the vocal-sensorimotor and the abstract components of language. My interest in this book is to highlight speech as a fundamental element in the origin of modern language and to depict an educated scenario for the evolution of the neural circuits involved in its generation. For this purpose, I will explore the evolutionary history of our brain to understand from where these networks originate.

Darwin, Broca and the Human Brain

Perhaps the most basic assumption we need to make before continuing is that the human brain is the product of biological evolution and that the origin of language is inevitably related to the evolution of the human brain. In the mid to late nineteenth century, Charles Darwin's ideas on evolution sparked intense debate, not only with the church, but also within science. Darwin first published his *Origin of Species* in 1859, and in 1871 he published the *Descent of Man*, in which he argued that humans and apes had a common ancestry, and also proposed a biological origin of human language (Darwin 1859, 1871). On this point he dissented from the co-founder of the theory of evolution, Alfred Russell Wallace, and with the main linguistic tendency of the time, which viewed language and the human mind as attributes unique to the human species. Darwin proposed three stages in the acquisition of human speech or language, first a general increase in intelligence that permitted more complex social behavior and a primitive kind of thought, which was followed by the development of complex vocal control in the form of primitive melodies, a "musical protolanguage" as termed by Tecumseh Fitch (Fitch 2010). Subsequently, these melodies achieved meaning by imitating natural sounds, aided by signs and gestures.

However, Darwin never directly addressed the issue of brain evolution, or whether the human brain evolved gradually from that of non-human primates. On the other hand, Richard Owen (Owen 1837; Desmond 1984), Darwin's main scientific opponent and arguably the most brilliant anatomist of his time, pointed to apparent key differences between the human brain and that of non-human primates. Instead of believing in a historical transformation of species, Owen believed in an ideal, abstract world where forms or species existed immutably. Owen was also a strong defender of man's privileged place in nature and published several works highlighting the uniqueness of the human brain. In 1857, two years before Darwin's *Origin*, Owen published an article indicating characters apparently unique to the human brain, namely the so-called calcar avis or hippocampus minor, which is a small groove in the posterior horn of the lateral ventricle (now known to be associated with the development of the calcarine fissure, which contains the primary visual area) (Gross 1993). In addition, the development of the posterior lobe of the brain, associated with the enlargement of the posterior horn of the lateral ventricle, was also considered to be a uniquely human attribute. Both features endowed humans with their supreme power over all other animals and creation. Owen illustrated this point by comparing the brains of a gorilla and a "negro", the former supposedly lacking these characteristics. While Darwin made only mild commentaries on Owen's statements, his close follower Thomas Henry Huxley (nicknamed "Mr. Darwin's bulldog") defended him against Owen and the church's opposing creationism. On this particular issue, Huxley demolished Owen by quoting other authors that showed the presence of a calcar avis in several animals including non-human primates (Gross 1993). This pointed to a gradual transformation of the brain from apes to humans. Still, the question remained, what was in our brains, and not in those of apes, that made us speak?

Another character of relevance for our purposes is Pierre Paul Broca, a brilliant physician, surgeon, and anthropologist contemporaneous to Huxley (Broca was only 1 year older), who embraced the notion of human evolution (Schiller 1992). He has been quoted as saying that he preferred being an evolved ape than Adam's degenerate descendant. However, although he agreed that it was a splendid hypothesis, he insisted on the lack of evidence to support the theory of evolution by natural selection. In 1859, Broca founded the Society of Anthropology of Paris, which was dedicated to the study of the human races, their origins, and their evolution. Two years later, Broca presented the case of a patient with a localized lesion in the brain that had lost the ability to speak, a crucial finding that initiated the study of the neural basis of language and established that speech was the product of specialized structures in the human brain (see Chapter 2). In addition, Broca made important contributions to physical anthropology, analyzing the cranial shapes of different European groups and attempting to trace European – and French – ancestry. The first formal paper from the Society of Anthropology was of his authorship, in which he argued that the French were in fact a mixture of peoples, at odds with the widespread notion that the French were a single race derived from the Celts. An important part of his anthropological work consisted of a craniometric analysis of Basque skulls, as compared with skulls from northern France. Although Broca believed in a separate origin of the human races, he also firmly argued that the mixing of races was not detrimental to their vigor or intelligence, which challenged the contemporaneous notion of the superiority of "pure races" like the Celts or Aryans. Broca also did some paleoanthropological work, notably describing the skull and skeleton of the Cro-Magnon man found in the region of Les Eyzies in southern France, and the study of ancient trephined Peruvian skulls.

Did Language Evolve?

While the notion of human evolution was slow to be accepted, ideas about the origin of language and speech were debated long before Darwin's evolutionary theory. However, it was only after him that these traits began to be considered a biological issue. In *The Descent of Man*, Darwin asserted that speech owes its origin to the capacity to imitate and modify natural sounds, as well as the voices of other humans and other species (Richards 1989). He emphasized the coevolution of music and language, arguing that the earliest languages were musical, and claimed that language evolution was closely aided by the development of communicative gestures. With his characteristic insight, he went on to propose that the process of speech acquisition was not much different from the mechanism of song learning in songbirds. However, Darwin was not alone in his interest in the genesis of speech. Nearly contemporaneous to the publication of Darwin's work, there was an avalanche of ideas on the origin of language, including proposals of imitation capacity, emotional exclamations, rhythmic behavior, and gestural communication. Among the most influential of these theories was that of Gottfried Herder (Herder 1800), who proposed that the first words were imitations of natural sounds, like the onomatopoeias of modern languages (words that recall natural sounds like “oink”, “meow”, “buzz”, etc.). Another theory was that innate calls like crying or laughing served as the substrate for the origin of words, as these calls convey socially important emotional information. Grunts and other calls gradually transformed into primitive words, or proto-words as they are called. In a similar line, James Burnet proposed that innate cries became varied by musical tones before becoming articulated words. This hypothesis implied that speech derived from music, which was considered to be a more primitive form of expression. However, the neurological findings at the time, showing dissociation between speech capacity and emotional expression, were used as firm evidence against this hypothesis. Finally, the prestigious philosopher Étienne de Condillac supported the notion that language originated as gestural communication, akin to the sign communication used by the hearing impaired (Richards 1989; Fitch 2010).

All these theories had one thing common: none had a single bit of evidence in their support. They were all speculations about our early history. In 1866, the *Societé de Linguistique de Paris* decided to ban this sterile discussion from academic contexts, producing a long eclipse in research about language evolution. Fitch asserts that the linguist Friederich Max Müller was perhaps the most radical opponent to theories of speech origins (Fitch 2010). He rejected the most well-known theory of onomatopoeia on the grounds that most words are not strict imitations of sounds. However, this imperfect imitation can be sufficient for others to match the vocalized sound to the natural sound it refers to. It does not need to be perfect to communicate its meaning. Müller acknowledged that humans could have evolved from other species, but in agreement with the Book of Genesis, he believed language to be a gift from God, who gave humans a single language that diversified into all extant languages. Müller’s research focused on the reconstruction of the original human language, a subject on which he had been a pioneer, and in this respect considered himself to have been a “Darwinian before Darwin”. However, the original language was to him an abstract entity, a machine for thought, not a concrete system of sounds. On this, he joined the idealist tradition that has continued into modern linguistics.

Even if at first sight the arguments proposed at that time sound naïve, all these proposals remain important in the literature of language acquisition and evolution. To be fair, although we have acquired tremendous knowledge of linguistics, biology, anthropology, and psychology, the main question of why we were the only species to acquire language, and the specific process underlying this transformation, remains unsolved. This may not be anyone's fault, given that language and behavior, unlike other biological characters, do not leave fossil traces, and we cannot know directly how our ancestors communicated. Moreover, despite our genetic similarity, non-human primates show nothing remotely similar to language or speech, and there are no living human-like species using a primitive form of communication that would help in tracing the history of language acquisition.

Deep Structures

In the second half of the last century, the extraordinary linguist Noam Chomsky (well known to the general public for his radical anarchistic declarations) and biologists like Richard Lewontin (Chomsky 1957; Lewontin 1975) further contributed to dismissing the evolutionary origins of language by boldly claiming that language was so unique that it was not explainable by evolutionary theory. Excluding notable attempts by a few twentieth-century neurologists and psychologists, scientific enquiry into the origins of language and speech only re-emerged in the last 20 years by virtue of the advent of neural imaging techniques to assess language processing and brain anatomy, and the development of comparative approaches to non-human species that provided insightful models of the development of communication and other behaviors. There has been much research recently on neural networks and the mechanisms underlying language, memory, and motor control, together with exciting comparative studies on the brains of non-human primates and animals like songbirds that are able to learn new vocalizations. All these studies have yielded important evidence that, although still fragmentary, provide a new avenue to thinking about language and its origins.

Apart from this recent influence, linguistics has traditionally been an issue of paper-and-pencil work, attempting to unveil the logical organization of linguistic utterances. Chomsky's revolutionary theory emerged in this context, claiming that despite their superficial differences, all languages share a deep grammatical organization based on the hierarchical organization of phrase structure (broadly referred to as generative grammar). The acquisition of grammar is considered to be innate, as all humans share the capacity (or competence) to master language. Thus, language has a universal structure and we are endowed with the ability to learn it from birth. Furthermore, the structure of language is considered to be unique, having no parallel in either other human cognitive functions or any animal cognitive or communication system. Chomsky strongly emphasized phrase structure as the key element of language, downplaying other elements like lexicon, phonology, or semantics as less relevant to the essence of language. For Chomsky, language consists of a core computational system that, although useful for communication, represents the fundamental structure of the human mind.

Chomsky made a titanic contribution to formal linguistics. He is probably the best syntactician that has ever lived, and imposed a tough, logical approach to the study of grammar. His influence began with the publication of the book *Syntactic Structures* (Chomsky 1957), which was based on his doctoral dissertation. There, he attacked and practically destroyed the then prevalent view that language was acquired by behavioral mechanisms of learning and associativity, proposing instead that language was the result of an innate capacity. In subsequent works, Chomsky engaged in in-depth analysis of syntactic organization, for which he developed a highly intricate logical system that, although clear to him, became increasingly obscure for non-linguists and even many linguists as well (Chomsky 1965). This whole analysis revealed that the grammatical structure of language could become extremely complex, so much that it required a sort of Copernican revolution to make more sense of it. Attempting to simplify his theoretical construct, Chomsky argued that the essential feature of language is its recursive grammatical structure, which in simple terms is the process of inserting sentences into other sentences, generating an embedded organization where phrase components are hierarchically nested into longer phrases in a potentially infinite branching tree. Recursion, he claimed, is unique to language and is not originally intended for communication, but makes up the architecture of the mind. Going further in this search for simplicity, Chomsky offered "The Minimalist Program", in which he pointed to a minimal operation he called Merge, which consists basically of joining different elements (be they words or phrases) iteratively in a binary tree that is able to generate all syntactic structures (Chomsky 1981). It is interesting to note that in the opinion of the evolutionary linguist Derek Bickerton, Merge may more accurately represent the brain mechanisms involved in language than does recursion, as it refers to the binding and connectivity of lexical items in terms of their semantic significance (Bickerton 1990, 2009, 2014). Furthermore, Bickerton boldly asserted that by creating Merge, Chomsky "assassinated" his own child, recursion. He went on to argue that all recursive structures can be achieved solely by using Merge, with no need for recursion. Thus, it is the lexical properties of words that determine the binding rules and the resulting hierarchies of phrase structure. In Bickerton's view, what is critical for the initial emergence of language is not syntactic structure but a lexicon and its associated semantic representation. He said that this of course is only the basics. Much more is needed to develop modern language, including grammatical elements of inflection, case marking, etc. that are not accounted for by this model.

Chomsky's perspective has been also criticized by linguists like Steven Pinker and Ray Jackendoff, who argued that phonology also shows a unique syntactic organization (but different from phrase structure) and highlighted the relevance of many other aspects of language, including semantics, the lexicon and large-level discourse structure (Pinker 1994; Jackendoff 1999). Pinker and Jackendoff claimed that syntax is actually a mechanism to represent hierarchical cognitive mechanisms in a phonological dimension. In particular, Steve Pinker advocated a more biologically based perspective on language. In line with Charles

Darwin, he saw language evolution as a highly complex adaptive process at all levels, from the remodeling of peripheral vocal structure to the elaboration of instinctive learning mechanisms, very much like the acquisition of birdsong or the development of flight in birds. According to Pinker, language is in fact unique, but so are the elephant's trunk and many other specialized organs in the animal kingdom.

Shared or Unique?

Considering the above, it is not surprising that in relation to language evolution, Chomsky has always been skeptical. In his early years Chomsky claimed that the complexity of language was such that it was impossible to find an organization so intricate in general cognition, even less in animal communication. Furthermore, his view that language is perfect does not admit the possibility of a gradual acquisition of distinct elements. It is either complete, or it is not. However, in later years, coincident with his strategy of simplifying his syntactical theory, Chomsky has become closer to biology and evolutionary theory. He teamed up with evolutionary psychologist Marc Hauser, who up to that point had strongly advocated a gradual Darwinian evolution of language. After what was probably very intense conversations at the beginning of this century, Chomsky and Hauser reached an agreement in which they parceled the study of language into two territories: one amenable to comparative and evolutionary studies (Hauser's domain) and the other reflecting the core elements of language and impenetrable to evolutionary analyses (Chomsky's domain). In 2002, they published a now highly cited paper, together with co-author Tecumseh Fitch, whom we met above, in which they made a clear distinction between what they termed the faculty for language in the broad sense (FLB, Hauser and Fitch's expertise), and the faculty for language in the narrow sense (FLN, Chomsky's expertise) (Hauser et al. 2002). FLB includes all biological traits shared with other species or with non-linguistic cognitive mechanisms, while FLN is a single monolithic and indivisible trait that must have appeared only once (recursion stands out as a prime candidate to be included in FLN).

As an example of FLB, Hauser, Chomsky, and Fitch discussed the case of categorical perception, which consists of the capacity of the auditory system to perceive discrete phonological categories like /ba/ as distinct from /pa/, while in fact there is a continuous transformation of the spectral composition as one goes from one phoneme to another (more on this in Chapter 8). This was first observed in the laboratory of Alvin Liberman in the 1950s and was presumed to be the basis for speech perception (Liberman et al. 1957, 1967; Liberman and Mattingly 1989). Theoretically, Lieberman's hypothesis considered that the acoustic system was somehow framed by the phonological motor apparatus by mapping continuous perceptual categories into discrete motor programs involved in the execution of these phonemes. This hypothesis was based on two assumptions: first that categorical perception is unique to humans and is an adaptation for speech perception; and second that speech perception is highly dependent on motor programs. Neither of these statements survived for long in science. In the seventies and eighties, several articles appeared showing categorical perception in chinchillas (a cute furry South American rodent), monkeys and even birds, none of which are able to speak (see Chapter 10) (Hauser et al. 2002). Thus, the trait is not uniquely human, and categorical perception does not depend on the existence of phonological motor programs, because these are clearly absent in chinchillas, monkeys, and birds. This theory has re-emerged as the now fashionable mirror neuron hypothesis, which I will discuss at more length in Chapter 8. For Hauser, Chomsky, and Fitch, categorical perception is a paradigmatic instance of a trait that fits into FLB. Furthermore, they say that the claim of human uniqueness is difficult to demonstrate as it "must be based on data indicating an absence of the trait in nonhuman animals and, to be taken seriously, requires a substantial body of relevant comparative data" (p. 1572). This example also illustrates some ambiguity in defining FLB, because although they worded their arguments in such a way to give space for evolutionary processes, they also implicitly consider the traits included in FLB as uninteresting and uninformative about language origins. Another arguable point is that Hauser, Chomsky, and Fitch provide examples of species distantly related to humans as instances of shared FLB, denying that this character is absent in non-human primates and must have been acquired during human evolution, independently of whether distant lineages also acquired it in parallel.

FLN should include features that are present only in human language and are not shared with other animals or with other general cognitive abilities. In the article's abstract, they "hypothesize that FLN only includes recursion and is the only uniquely human component" (p. 1569). In the article, however, they emphasize that recursion maps into the sensorimotor (phonological) and conceptual-intentional (semantic) interfaces, and extend their hypothesis for FLN as consisting of "the core computational mechanisms of recursion ... and the mappings of the interfaces" (p. 1573). Thus, FLN may not only consist of recursion but also of any link that connects recursive processes with phonological and cognitive mechanisms. Besides lacking precision, this sentence leaves much in the air. In this extended version, FLN could consist of multiple mechanisms ("the mappings of the interfaces") besides recursion itself, which would make FLN a complex evolutionary acquisition. If this is so, FLN might not be the single monolithic element that was proposed above. Finally, they proposed that FLN (recursion) originated from a domain other than communication (in line with Chomsky's assumption that recursion is not primarily for communication), but then, in what appears to be a concession from Chomsky to his evolutionist partners, they proposed that it may have emerged from computations involved in number processing, navigation skills, or social relations. This is perfectly possible but there is no evidence suggesting that this is the case, as opposed to the alternative of an origin directly related to communication. Furthermore, it contrasts with Chomsky's earlier argument that recursion is produced by a modular, encapsulated computational system that is distinct from those involved in general cognitive mechanisms. Attempts by other authors to observe recursive-like processing in animals have been severely criticized by these authors, who seem to deny this possibility *a priori* while opening the possibility of recursion having its origins in animal capacities. More recently, Chomsky has associated with computational linguist Robert Berwick, insisting on the separation between externalized, sensorimotor elements, and the central properties of language (Berwick and Chomsky 2016). Berwick and Chomsky make special emphasis on stochastic effects in evolution, where few mutations can account for dramatic evolutionary changes (as seen in the descending neural control of vocalization in humans and songbirds), and on the fact that contingent factors like having a very small population size (as it may have happened in human evolution) can yield non-adaptive changes in gene frequencies, driven by chance. While I will get into those issues through the book, now I will focus on the central argument of FLN as the core element of language.

Experiments on Recursion

Although Hauser, Fitch, and Chomsky sometimes defended their notion of FLN as a hypothesis, in subsequent articles, these authors and others assumed that recursion actually represents the core aspect of language and neglected any other approach to language evolution as being totally irrelevant (Hauser et al. 2014). While they asserted that a claim of the uniqueness of a human trait requires “a substantial body of relevant comparative data”, they did not apply this criterion to the capacity for recursion. There have been only a handful of experiments assessing the capacity of recursion in non-human animals, and by no means have these been exhaustive. Perhaps the most important experiment in this line was in 2004, when Fitch and Hauser assessed the capacity of tamarin monkeys (a cute, little South American monkey with a complex social life) to process recursive auditory patterns (Fitch and Hauser 2004). To do this, they implemented what is known as an artificial grammar, or a laboratory invented grammar. Two grammatical versions were included, using acoustic stimuli consisting of consonant-vowel syllables like “ba”, “pa”, “du”, etc. These syllables were separated into two classes, those voiced by a female (class “A” syllables) and those voiced by a male (class “B” syllables). They then combined these stimuli into two different patterns, one that ordered the syllables in sequence, like “AB”, “ABAB”, “ABABAB”, representing what they termed a finite-state grammar and dubbed $(AB)^n$. The other pattern consisted of recursively inserting a pair of AB syllables into another AB syllable ("AB", "AABB", "AAABBB"), termed phrase structure grammar and dubbed A^nB^n. Two groups of tamarin monkeys were exposed to repeated playbacks of either the $(AB)^n$ or the A^nB^n grammar, to familiarize them with these patterns. The next day, the monkeys were briefly re-familiarized with the same stimuli and then presented novel stimuli, some that fitted the "grammar" that they had been habituated to, and some that violated this "grammar", for example, a sequence "AABA". Previous studies in human infants and animals had shown that a novel, unexpected stimulus raises the subject's attention and makes him/her look at the stimulus source (the speaker in this case). Fitch and Hauser found that monkeys that had been trained with the non-recursive grammar $(AB)^n$ were surprised when a deviant pattern was presented, while monkeys that were habituated to the recursive structure A^nB^n did not distinguish violations from grammatically correct sentences. For adult humans, it is very easy to distinguish the deviants of both recursive and non-recursive grammar types. Thus, they concluded that tamarin monkeys were unable to process recursive patterns. In 2006, Timothy Gentner and colleagues published a highly controversial paper in which European starlings (a species of songbird) were able to learn recursive A^nB^n grammars (Gentner et al. 2006), but this has been dismissed by several authors that argued the birds were able to predict surface regularities of the different grammars based on probability and memory, and not by making an abstraction of a recursive pattern. We will go further into this discussion in Chapter 9, when discussing songbirds as animal models for speech acquisition.

In 2005, Steven Pinker and Ray Jackendoff published a strong and lengthy critique of Hauser, Chomsky, and Fitch's original article, claiming that language is the result of highly complex and interrelated adaptations (Pinker and Jackendoff 2005). Furthermore, Pinker and Jackendoff highlighted phonological and lexical processes and many other elements of syntax like case marking, agreement, and other grammatical subtleties. These are essential for the main function of language, which is to transmit information of who did what to whom, what is where, and other semantic relations. Pinker and Jackendoff claimed that these issues had been severely neglected by Hauser, Chomsky, and Fitch, in the same way as Chomsky's minimalist program de-emphasizes them. Furthermore, they criticized Fitch and Hauser's experiment on tamarin monkeys on the basis that although language is recursive, the artificial grammar A^nB^n is not a possible human language, and violates basic principles of Chomsky's universal grammar. Finally, they ended by agreeing with Hauser, Chomsky, and Fitch in their statement that recursion may not be unique to language, and assert that the only reason language needs to be recursive is for expressing recursive thoughts. There was a reply from Fitch, Hauser, and Chomsky, and a counter reply from Jackendoff and Pinker, both claiming that the others had not touched their essential points, reaching then a point at which no one cared about or considered the others' arguments (Fitch et al. 2005; Jackendoff and Pinker 2005). Perhaps one issue that deserves further comment is Jackendoff and Pinker's claim that at least human visual processing has an A^nB^n recursive structure. Consider a semi-regular array of pairs of dots, nested in groups of two, four, sixteen, two hundred and fifty six, and so on. This combinatorial process can go on indefinitely and may play a role in visual categorization. However, this is not exactly the same as linguistic recursion, as there are no labels or rules to know which patterns combine with others.

Another less publicized but very important point was raised by Pierre Perruchet and Arnaud Rey in 2005, who showed that humans learning the A^nB^n recursive pattern did so "without exploiting in any way the embedded structure" (Perruchet and Rey 2005). Furthermore, when they modified the conditions and made recursive processing mandatory, human subjects were unable to learn these patterns, just like monkeys. More recently, in 2012, Rey, Perruchet and Joël Fagot showed that baboons were able to process visual stimuli in a center-embedded or recursive pattern (Rey et al. 2012). They trained baboons in an intensive task where the animals had to sequentially touch pairs of visual shapes presented in a screen as in (first, touch figure A1 → then, touch figure B1); (A2 → B2); (A3 → B3), etc. The pairs were sometimes presented together with distractors to be ignored. In the second part of the experiment, (A1) appeared on the screen together with a distractor, and the animal had to press on (A1). Then, (A2) appeared with another distractor and the animal had to press (A2). Afterwards, (B1), (B2) and a third stimulus (that could be a distractor or an unrelated target figure like (B3)) appeared together on the screen, and the baboon had to touch on either of the target stimuli (but not on the distractor) in whatever sequence. Subjects preferred to touch (B2) and then (B1), as (A2) had previously appeared more recently than (A1), making up the embedded sequence (A1 → A2 → B2 → B1). The authors concluded not only that human recursive abilities may be tracked to non-human animals, but more strongly that they rely more on short-term, or working memory capacities than on abstract representation of recursive rules. As in Pinker and Jackendoff's example of visual recursion, the kind of embedding reported by Rey, Perruchet and Fagot is rather different from the recursive mechanisms observed in language, but it does fit the artificial A^nB^n pattern used by Fitch and Hauser with tamarin monkeys.

Regardless of discussions about their linguistic validity, Fitch and Hauser's experiments raise important questions about the real psychological processes used by subjects to solve these tasks. Although the researchers had the abstract property of recursion in mind when they designed the tasks described above, this may not have been the mechanism by which subjects resolved the problem. Rather, they may have relied on clever, but perhaps lower-level tricks based on associativity and short-term memory. A notable example of this is Clever Hans, not the Brothers Grimm fairy tale character, but an actual horse that lived in the early 1900s (Richards 1989; Candland 1995). The horse's owner claimed it could count, add, subtract, and provide correct answers to complex questions. While different answers were presented to the horse, it responded only to the correct one by tapping its forelimb. Wilhelm von Osten, the horse's owner, was a mathematics teacher and was impressed with Darwin's ideas about the evolution of behavior and intelligence. Clever Hans received so much attention that Oskar Pfungst, a psychologist interested in this phenomenon, went to observe the horse's performance. First, he found that the horse responded correctly to anyone, not only to its owner. But it responded correctly only when the questioner also knew the answer. Pfungst went on to conduct many tests, including isolating the animal from the public and blocking its visual contact with the questioner. He found that the body posture and facial expressions of the questioner changed before offering the correct answer. The horse perceived these subtle cues and hoofed its response right away. The extent to which we are dealing with "Clever Hans" phenomena in language or psychology is a major issue today. Are we really grasping at abstract representations of complex stimuli, or are we using subtle unconscious cues to make up an illusion about reality? Moreover, what about children learning to speak, or animals tested in the laboratory? Are they really attending to the deep grammatical or conceptual structures, or do they use subtle perceptual cues to predict complex outcomes? Throughout this book we will see that the Clever Hans issue pervades much of today's human and animal research, including language processing and other theories of human behavior.

Pidgins and Creoles

Besides discussions about whether animals can or cannot master recursive patterns, and whether recursion is indeed unique to language, in my opinion the main problem with the proposal of Chomsky and his colleagues is that strictly speaking, they do not provide any fundamental insight into language evolution. The main ideas in Hauser, Chomsky, and Fitch's article could have been written by a non-evolutionist, highlighting that we have shared features with other species (that are therefore uninteresting), and we have unique characters provided by recursion. This has been in fact the standard classificatory logic even before the advent of Darwin's evolutionary theory. Species are clustered by their shared characters and differ from each other by non-shared characters. The Australian agnostic anti-evolutionist Michael Denton has in fact argued that the entire classificatory method provided by modern biology and genetics points to a hierarchical or nested categorization of different groups, where larger groups share more characters and smaller groups differ from each other in features that are unique to each group. For him, there is no need for an evolutionary theory (Denton 1985).

Derek Bickerton, a well-known evolutionary linguist mentioned above in relation to Chomsky's concept of Merge, has said that the arguments of Hauser, Chomsky, and Fitch provide no clues as to the process by which we acquired language (Bickerton 2009). Bickerton made his reputation by studying creole languages in Hawaii and the Caribbean, which according to him provide cues about the origin of language. Creoles are people of foreign ancestry (usually African) born in European or American colonies. Creole languages consist of a grammatically simplified mixture of ancestral languages and the dominant language of the colony. Creole languages derive from pidgin languages, which are even simpler forms of communication used by the parents of creoles, who were born in a different land and have the ancestral language as their mother tongue. Pidgin is used for trading with the colony owners, as neither Africans nor Europeans understand the other's language. The mother tongue of creoles is, in a way, the pidgin spoken by their parents. Bickerton has proposed that the cues for a universal grammar can be found in the structure of creole languages, whose rudimentary grammar develops spontaneously from language competent children taught in pidgin. In a way, Bickerton embraces Chomsky's conception of universal grammar and the uniqueness of language, but differs radically in his conception of language evolution. He has proposed some hypotheses (admittedly, some of them quite difficult or impossible to refute or verify) about the evolution of language, which in essence represent a sequence similar to the transition from pidgin →creole → established language with full-fledged syntax.

Bickerton, together with other scholars like Terrence Deacon, have strongly emphasized the capacity for symbolic representation as the initial impetus for language evolution (Deacon 1997). Symbolism is the capacity to evoke a mental representation of an object or an event by providing a sensory stimulus not directly related to that object or event. Symbols are the result of consensual agreements among different people and are in essence arbitrary. For Bickerton, a key event in the acquisition of symbolic capacity was related to the ecology of early *Homo*, who constructed an ecological niche for themselves in high-level scavenging, competing with other hyper-carnivores. Early *Homo* species like *H. erectus* developed primitive tools that were used both for cutting up dead animals and for defending themselves from other animals, and established a vague social organization consisting of large, loosely knit groups that assembled and disassembled over time, which may have been optimal for finding carcasses over a wide area. Under these conditions, it was essential to have ways to call one another in an honest manner (that is, providing a signal that the other considers true), as individuals or small groups needed to recruit as many others as possible when they found a good carcass to fend off dangerous scavengers that competed with them. According to Bickerton, this was the key for the acquisition of displacement, that is, providing an alert signal about something that is not there but somewhere else. Most animals communicate using honest signals (there are also "dishonest" ones that provide benefit only to the sender, but this is not our issue at this point). However, as Marc Hauser had already shown some time before his association with Chomsky, animal communication systems provide information about the here and now, and there are only exceptional examples of animals providing signals for events distant in time and space (Hauser 1996). In the 1980s and 90 s, comparative psychologists Robert Seyfarth, Dorothy Cheney and Peter Marler showed exciting evidence of alarm signals produced by vervet monkeys (a mid-sized African monkey) in response to specific predators that differed whether the predator was a snake or an eagle (see Chapter 10) (Seyfarth et al. 1980). But were they saying "eagle" or "snake"? The current interpretation is that this is not likely, as the two calls trigger different escape responses. Monkeys do not call to talk about an eagle or a snake; rather they may be saying, "Danger from above, lets hide in the bushes". On the other hand, bees and ants, although quite different from us, do show the capacity for signal displacement, as seen in the bee's waggle dance. But again, this kind of signal is very hard wired and genetically determined. Instead, early humans achieved signal displacement by conventionalized learning.

The first displacement calls were probably not truly symbolic, but contained some elements reminiscent of the object they referred to, as for example imitations of the sounds made by the living animal whose carcass had been found, made by vocalizations or pantomime. Eventually, these signals became truly symbolic and allowed concepts to emerge. In the first stage of this sequence, early humans must have acquired the ability to imitate the sounds produced by other animals, or the sounds made by natural elements like wind or water. Other species, notably parrots and songbirds, but also marine mammals and elephants, do show imitative vocal capacity, and have been called vocal learners. Humans are also vocal learners, but our common ancestor with the chimpanzee was probably not a vocal learner. Thus, the acquisition of vocal learning must have been a crucial requisite for speech to evolve.

After the acquisition of symbolic representations in long-term memory, humans developed linguistic concepts, claimed to be different from animal concepts as these are "anchored" by a symbol that permits evoking it in different contexts, and keeping it active online in short-term memory. Short-term, or working memory, also allowed combining these symbols in different ways, and messages began to increase in complexity by arranging strings of a few of these primitive words, generating a pidgin-like protolanguage, with few if any rules governing the order of these utterances. Based on the notable conservation and lack of improvement in tools for some 200 million years in early human evolution, Bickerton claimed that the protolanguage stage remained unchanged for all this time. With the appearance of *Homo sapiens* and its cultural revolution some 100,000 years ago, the hierarchical structure of language developed together with the acquisition of Merge, Chomsky's operation that provides rules for assemblies of words into sentences. This was mainly a cultural, rather than a biological innovation. For Bickerton, the crucial event for the development of language was the acquisition of a lexicon in which the semantic component of words determined the possibilities of joining with other words. For this, many auxiliary words began to appear as links between lexical entities and modern grammar was on its way.

The scenario just summarized is, as I said, quite speculative and largely impossible to verify. Nonetheless, there are important considerations that can be used when tracing the evolutionary history of language. The first is that language conveys displaced learned signals, as opposed to animal communication, which shows no displacement although they may communicate with learned signals, particularly vocal ones. The second is that symbolic representations were probably a critical event in language evolution, and these could be stored in long-term memory and manipulated in short-term or working memory. However, Bickerton said little about the nature of these early words, and how we came to develop a rich vocal system that enables us to imitate sounds and generate long strings conveying a theoretically immense number of possible messages. In this line, Ray Jackendoff suggested a similar sequence of events as Bickerton's in the acquisition of language, more focused on syntactic operations, but perhaps more importantly, emphasizing the early development of a phonological system that combined preexisting sounds to form new ones (Jackendoff 1999). The words of protolanguages may have had a rich phonological structure, a point that I will discuss later. Basically, Bickerton did not put much attention on the development of the vocal system, and seemed to take this for granted. But more likely, the protowords that enabled protolanguage may only have been possible with a previous, rich repertoire of learned vocal signals that were used in social contexts, including mother-child behavior, group cohesion, alliances, and other instances. It is also clear, and Bickerton noted this, that gestural communication was an important component of the social life of early humans.

Toward Biology

One thing that Chomsky and Bickerton have in common is their assumption that the human brain must have been "rewired" for the acquisition of modern language, that is, its connectivity must have diverged from the condition of our close ancestors. However, neither of them says anything about the specific neuronal changes that could have brought about this tremendous achievement. In the 1950s and 60s, the cognitive psychologist Eric Lenneberg teamed up with Noam Chomsky, attempting to provide a biological foundation to Chomsky's innateness hypothesis (Lenneberg 1967). At about the same time, the neurophysiologists Wilfer Penfield and Lamar Roberts 1959 first proposed the existence of a critical period of language acquisition, based on the earlier concept of a critical period for the development of innate behaviors coined by the Nobel laureates Konrad Lorenz 1981 and Niko

Tinbergen (Tinbergen 1951; Penfield and Roberts 1959; Lorenz 1981; Burkhardt 2005). The concept of a critical period was neurobiologically supported by the contemporaneous studies of the also Nobel laureates David Hubel and Torsten Wiesel (awarded in 1981), in the development of the visual system of monkeys (Hubel 1977). Hubel and Wiesel demonstrated that alterations in sensory experience could have dramatic effects on the connectivity of the visual cortex, but only during a brief period after birth. After this period, connectivity stabilizes and becomes more resistant to sensory deprivation. The notion of a critical period in language was based on two lines of evidence. One is the fact that the earlier one learns a second language, the less evident is one´s accent. A well-known example is the Kissinger brothers Walter and Henry (the latter former US Secretary of State), both German Jews that immigrated to the US when Henry was 12 years old, and Walter only 10. While both learned to speak English fluently, Henry never lost his German accent, while Walter is said to sound like a native English speaker. It is commonly said that this is due to the difference in their ages; with the younger Walter still able to reorganize his language network while the older Henry was not. However, it is also possible that Henry was simply not as good at languages as his brother, at any age. The second line of evidence originates from the few known cases of feral children that apparently grew up isolated from human contact, either in a state of confinement by their parents or other people, or simply living in the wild (Candland 1995). Feral children usually have profound difficulties in getting used to living with humans, are incapable of following basic norms like using a toilet, let alone communicating. Their inability to learn to speak is notable, which has been attributed to their lack of early social stimulation. Nonetheless, these children usually show signs of having been abused and mistreated, and it is not clear whether they suffered neural developmental disorders that may have worsened their condition. Today, the study of critical periods for language acquisition (whether a first or second language) in humans represents an intense research agenda, some of which we will address in Chapter 10.

Lenneberg relied on this evidence and on Alvin Lieberman’s motor theory of speech perception mentioned above, to popularize the notion that language is a discrete and separable, species-specific trait, whose biological foundations followed patterns observed in other instinctive behaviors (Steven Pinker's bestselling 1994 book, *The Language Instinct*, is in some aspects a follow-up of these ideas) (Pinker 1994). In the 1970s, when I was doing my undergraduate training in Biological Sciences, I was impressed with all these ideas, embracing animal behavior, neuroscience and human language. I felt particularly motivated with Eric Lenneberg's book *The Biological Foundations of Language*, which at the time represented a brave attempt to conceive language as a biological trait, and it is no exaggeration that reading it was highly significant in my choice to study the evolution of human language.

Contemporary with these developments and a nascent link between linguistics and biology, some scholars began to ask whether language was truly innate in our species. After all, it was known that some bird species like parrots can learn human words even if they do not understand them. Furthermore, at this point there was the widespread notion that apes had a highly developed imitative ability. The first attempt to teach human language to a non-human primate was in the 1940s–1950s by the wife and husband team Catherine and Keith Hayes, who raised a young chimpanzee called Viki (Hayes and Hayes 1952). The Hayes taught Viki to speak by using extensive sessions of speech therapy procedures normally used for children with language disorders. Despite long intense efforts, Vicky was able to voice only four words: "mama", "papa" "up" and "cup", showing that the chimpanzee lacks the vocal motor control necessary to utter human words (however, see Chapter 10). But as noted above, language is not only speech. In the 1960s, the wife and husband team Beatrix and Allen Gardner, both comparative psychologists, followed by their student Francine Patterson in the 1970s and then others, underwent the painstaking task to train apes in hand-sign language commonly used by the deaf (Gardner and Gardner 1969, Patterson and Linden 1985). The Gardners worked with Washoe, a chimpanzee that was 10 months old when the training started, while Patterson worked with Koko, a young gorilla. Their intention was to determine the extent to which apes could acquire linguistic skills, by using a sensorimotor system other than auditory-vocal circuitry. These studies were successful in teaching animals to use signs with humans (after many years, Koko mastered 1,000 hand signs, and understood an even greater number of voiced words). This was, however, the result of very long and intensive training procedures. In comparison, by the age of three, children use around 3,000 words, but these have been learned effortlessly on the part of themselves and their parents. However, the most striking finding was that in none of these long and tedious experiments, were apes able to produce a simple combination of words in something resembling syntax, which is reminiscent of Bickerton's early stages of protolanguage, where there are words but little or no syntactic rules.

In the 1970s, Kenneth Oakley, and later on in the 1980s, researchers like William Calvin and Michael Corballis proposed that language arose as an evolutionary outgrowth of hand dexterity, which was initially used in shaping and handling tools and throwing objects (Oakley 1972; Calvin 1983; Corballis 1993). From that, communication may have developed using body and manual signs, and was finally transferred to the mouth for reasons that are not completely clear (see Chapter 4). This proposal received very intense support from findings by the neurophysiologist Giacomo Rizzolatti and collaborators in Parma, Italy. Rizzolatti's team identified an interesting class of premotor neurons in the macaque cerebral cortex, called "mirror neurons", which were activated both during the execution of a motor act and the observation of another performing the same motor act (or so it seemed) (Rizzolatti 1998). Together with Michael Arbib, Rizzolatti proposed that mirror neurons provided the neuroscientific grounds for the origin of a gestural language that was eventually supplanted by vocal speech (Rizzolatti and Craighero 2004). The role of gestures, and the relevance of the mirror neuron hypothesis in language evolution will be the subject of Chapter 8.

In the last 50 years another research tradition has followed Darwin's original line, which is the study of vocal learning in non-human species, mainly songbirds but also other kinds of birds, as well as mammals like dolphins, bats, and elephants (Chapters 9 and 10). Again, this tradition has experienced an explosion in the last 20 years, in which behavioral, electrophysiological, neuroanatomical, and genetic approaches have converged to analyze the mechanisms underlying the capacity for vocal imitation. Although these vocal learning species are only distantly related to humans, it is expected that they will serve as models to understanding the neurobiological and genetic mechanisms involved in the acquisition of vocal plasticity. For example, the speech-related gene FOXP2 is mutated in specific regions in humans while it is not in non-vocal learning apes and monkeys. It is also known to participate in vocal learning in songbirds and other species.

Our Family

Finally, before we begin with this book, it is necessary to consider human evolution, early human behavior and the fossil evidence for the evolution of the human brain. I said that language does not fossilize, but fossils and human-made artifacts yield important clues about the behavior and the brain capacity of our ancestors, which allow us to make educated guesses about the evolution of human communication. The human lineage started some 6 million years ago, when our common ancestor with chimpanzees split into two lineages. Possibly the closest we have to the last common ancestor is *Ardipithecus*, which lived in trees but was able to walk on two feet. Australopithecines appeared 4 million years ago, and lived until almost 1 million years ago in Africa. *Ardipithecus* and most Australopithecines were small (4 feet tall) with brains not much larger than 500 cubic centimeters, about the same size as that of chimpanzees. The genus *Homo*, our lineage, makes its debut about 2.5 million years ago in early species like *H. habilis, H. rudolfensis* and *H. naledi*, which are gradually replaced by *Homo ergaster* and *Homo erectus* about 1.8 million years ago (Plummer 2004; Anton et al. 2014; Kimbel and Villmoare 2016; Crompton 2016). Brain size was slightly larger in members of the *Homo* lineage, particularly *Homo erectus*, which lived in Africa and Eurasia, with a brain size ranging from about 700 cubic centimeters in the earliest specimens, to somewhat more than 1,000 cubic centimeters in the latest individuals that lived 200 thousand years ago (Cornélio et al. 2016). The more modern *H. antecessor* and *H. heidelberguensis* had significantly larger brains than the last *H. erectus* specimens. Finally, Neanderthal man (*Homo neanderthalensis,* or *H. sapiens neanderthalensis)* and the closely related Denisovan Man

(*H. sapiens denisova*) appeared about 600,000 years ago, while modern humans (*Homo sapiens*) originated 200 thousand years ago. All these late humans had similar brain sizes of about 1,500 cubic centimeters, comparable to the average size of human brains today. Neanderthals and Denisovans disappeared about 30,000–25,000 years ago, possibly at the hands of our direct ancestors. But they did not become totally extinct, as there was an intense interbreeding between these species (or subspecies, to some), and we have inherited many of their genes. The small-sized and small-brained *H. floresiensis* (with a body and braincase the same size as that of Australopithecines) had a very short appearance, between 50,000 and 10,000 years ago. This is a contentious species, as some authors claim that it was a microcephalic or Down syndrome child of normal modern humans. *H. floresiensis* was found associated with Oldowan-like tools that some argue were made by Australopithecines.

Therefore, three stages can be observed in the evolution of the human brain. First, was the Australopithecine stage, in which brains are no larger than that of other apes. Then, the appearance of the genus *Homo* (especially *H. erectus*) set the brain race in motion, slowly increasing brain size over a period of 1.5 million years (Rightmire 2013; Cornélio et al. 2016). However, the increase in brain size was not accompanied by spectacular cultural advances. Early *Homo* species managed primitive stone tools from the beginning, and there is evidence suggesting that at least some Australopithecines used them too, as the oldest stone tools yet discovered date from about 3.3 million years ago. Early *Homo* made sturdy stone tools referred to as the Oldowan industry. The shape of Oldowan stone tools is very stable over time, with very little change in the design of the cuts to sharpen their edges. More sophisticated hand axes appear some 1.7 million years ago, with the Acheulean industry, which overlaps Oldowan technology. Some anthropologists have assigned Oldowan industry to *H. habilis*, and the Acheulean tools to *H. erectus*, with a larger brain size. In any case, both Oldowan and Acheulean tools remained more or less the same over time, evidencing very little evolution in their design. *H. erectus* also used fire and remains of one-million-year old campfires have recently been unearthed, which may support the cooking hypothesis for brain growth that I discuss in Chapter 3. How the use of fire relates to toolmaking and eventually to language evolution is a highly intriguing question to which unfortunately we have as yet no clues. Fire also changes social dynamics, as individuals begin to join around campfires, which increases food sharing and social interactions, and of course the task of making and maintaining fire, which requires strict cooperation among members of the group.

The last step in increased brain size, occurred first in Neanderthals and the closely related Denisovans (with a brain size comparable to ours), and is associated with the origin of more elaborate stone tools that appeared somewhat later, about 300,000 years ago (Sankararaman et al. 2016). This new technology is referred to as Mousterian industry, characterized by sharp flint tools, used for fine cutting and other tasks. Modern humans, who entered the scene slightly later, are also believed to have used Mousterian tools. From then on, there is an evident increase in cultural artifacts in the archaeological record, slowly initiating the Cultural Revolution that is still taking place among us. Why didn't *H. erectus* go further in its cultural development, despite reaching (in later specimens) brain sizes in the lower range of modern humans? Derek Bickerton termed this two-million-year period in *Homo erectus* history "the long stagnation", given that cultural achievements were slow to develop despite the doubling of brain size. The cultural explosion that began in Neanderthals, and then in modern humans could not have been due solely to brain size, as the brain had already been increasing in volume for a long time. What about language and speech? When did they appear? It is likely that modern speech is a very recent acquisition, which evolved together with cultural innovations. Bickerton claimed that *H. erectus* had a simple communication system, consisting only of a small set of word-like elements used to recruit subjects in the search for food. But in what context did these word-like elements appear? What was the biological difference that made our species able to speak about the world? I will touch on these critical questions throughout the book, but this is an advance for which neither I nor anyone else has a definite explanation

Ideally we could obtain additional information about our ancestors' brains by observing the cranial shape and the impressions the brain leaves in the cranial vault. The study of human brain endocasts (molds of the cranial cavity) has been an important discipline in paleoanthropology, yielding information about brain size, shape and growth trajectories in extinct lineages. Nonetheless, this discipline has also been plagued with controversies, the most well-known being that between Dean Falk and Ralph Holloway, which started as an argument about the presumptive presence of an ape-like sulcus in the occipital endocast of the Taung child, an immature specimen of *Australopithecus africanus*, originally described by Falk (Falk and Clerke 2007, 2012; Folk 2009; Holloway and Broadfield 2012; Holloway et al. 2014). Additional findings have been made by Philipp Gunz and colleagues, who observed that modern humans diverge from chimpanzees and from the Neanderthal man in the shape of the braincase, acquiring a more globular shape (Gunz et al. 2010; Neubauer et al. 2010). Furthermore, this unique globular shape is evident in the human neonate, differing significantly from the more elongated skulls of chimps and Neanderthals. These authors attribute the difference in shape to changes in neural organization to support higher cognition and cultural learning. Dean Falk, Emiliano Bruner, and others have focused on reproducing the sulcal patterns in fossil skulls, based on impressions in the inner surface of the cranial cavity (Falk 2014; Bruner et al. 2014). These and other authors claim to have found evidence for reorganization of Broca's region, the parieto-temporo-occipital region and the prefrontal cortex (which is almost the entire brain), a process that started in the Australopithecine brains when compared to modern apes. These arguments are reminiscent of Franz Gall's phrenology doctrine (see the next chapter), or the aforementioned studies on Basque and French skulls by Paul Broca. Can anything about cognition be reliably concluded from these cranial differences? In my opinion, not much. In general, the study of cranial morphology tells us little about the development of the neural networks involved in cognition. The distinct developmental trajectories in the shape of the cranial vault can be attributed to many factors, ranging from obstetric constraints to general craniofacial development, but there is no strong evidence that braincase geometry has anything to do with cognitive capacities. Nonetheless, I have to say that in a very large sample of human adults and children, Michael Gregory and collaborators recently reported that general cognitive capacity correlated with increased gyrification in the inferior parietal lobe, temporoparietal junction, insula and prefrontal cortex, all regions that have been associated with language processing (see next chapter) (Gregory et al. 2016). Whether these differences are reflected in cranial morphology still needs to be confirmed.

Although there is a full Chapter (Chapter 3) discussing the evolution of brain size, I want to mention here that achieving a large brain size has not been easy for humans. The upright posture achieved by Australopithecines implied profound changes in pelvic structure, providing more support to gluteal muscles and constraining pelvic diameter. As the brain size increases, a conflict develops between locomotor adaptation and the development of a larger skull at birth (human newborns are especially large-headed for their body size), which has been termed "the obstetric dilemma". Humans are very special animals that require assisted delivery and possibly have the highest rate of obstetrical complications. This constraint has implied a series of adaptations, like the rotation of the newborn before birth, so that the head comes out first (as opposed to what usually occurs in monkeys and apes), and the development of a circular pelvic canal to facilitate the transition of the newborn (Trevathan 2015, Huseynof et al. 2016; Ponce de León et al. 2016). Instead, others argue that the solution to the obstetric dilemma is rather a consequence of the geometry of growth rather than a specific result of natural selection (Fischer and Mitteroecker 2015; Mitteroecker and Fischer| 2016). This idea is in line with the now classic notion of evolutionary "spandrels", by Stephen Jay Gould and Richard Lewontin, who criticized an overly adaptationist trend in the 1970s that interpreted practically every observable trait as emerging from specific selective pressures (Gould and Lewontin 1979). Perhaps the most important developmental modification associated with the obstetric dilemma consists of delivering the newborn in earlier periods, thus giving birth to smaller and more immature babies. Steven Piantadosi and Celeste Kidd used an estimate of brain size as a proxy for measuring intelligence across species, and reported a strong positive correlation between intelligence estimates and weaning time across many primate species (Piantadosi and Kidd 2016). They propose that selection for increasing brain size leads to progressively immature newborns, which selects back to further increases in brain size as parents need more intelligence to raise their young. Although this is an interesting possibility, comparing intelligence across species is a contentious issue, as different species need to solve different kinds of problems to survive. In this line, Stephen Jay Gould and many others noted that humans retain juvenile characters such as a flat face, hairless skin, a thinner skeleton, and a large brain in relation to body size (Gould 1977). Changes in the developmental timing of different biological characters are called heterochronies, and the process of juvenilization, or keeping juvenile characters until the age of reproduction and beyond, is specifically called neoteny. Humans are neotenic primates according to many standards, including some behavioral characters including playfulness and less aggressiveness. Brain maturation rates are similar in humans and other primates, simply more prolonged in humans than chimps than monkeys, such that the human brain increases at rates expected from its size, only for longer times. In Chapter 11 I will take on this issue again.

This Book

A lasting perspective in the study of speech and language, of which I have talked little as yet, comes from neurology and consists of getting directly inside the speaker's brain. This became possible by the early findings by Paul Broca, and continued through most of the twentieth century. I will provide a short historical account of this tradition in Chapter 2. Notably, the linguistic and the neurologic traditions followed largely parallel histories for most of last century, without much communication between the two disciplines. Only recently, with the revolution caused by new brain imaging methodologies, have these two lines begun to converge. In this book, I will focus on the neuronal and connectional changes that made the emergence of language possible. To do this, in several instances I will delve into the biological aspects of human evolution, to which the acquisition of speech and language is necessarily subordinate. As mentioned, I will refer mostly to speech rather than other forms of language, and will not describe in detail the linguistic aspects of syntax, semantics, phonology or lexical structure, although I will make reference to them as they often impinge on neurobiological discussions. I will take into consideration the different perspectives on the evolution of the neural circuits for language, but will propose the novel approach that a key ingredient was the functional consolidation of a particular circuit connecting auditory and vocal regions in the cerebral cortex, termed the phonological loop by the brilliant psychologist Alan Baddeley (Chapter 6) (Baddeley 2007). This circuit largely overlaps with the classical language areas, but here I emphasize its critical role in verbal working memory in the evolution of speech. The activity of the phonological loop enables us to keep linguistic strings in short-term or working memory while we process them and plays a role in the acquisition of vocabulary and speech. Furthermore, it contributes an internal speech domain that facilitates the recognition and visualization of our own mental states, and may also contribute to engaging in long-term reciprocal conversations with others. Amplification of this circuit astronomically propelled our communication capacity compared to an ancestral multimodal (vocal-gestural) communication system, and may have facilitated the development of a complex grammar. Finally, I will argue that speech was the first instance of elaborate language, and that it allowed us to generate a shared semantics, and consequently a shared mind (Chapters 10 and 11).

In following this approach, it is necessary to discuss several features of the human brain, notably its size, functional asymmetry and the necessary exchange of information in both hemispheres in a lateralized brain. Since Darwin, language has been associated with the size of the human brain (Chapter 3). Is this really causal, and if so, do larger brains give rise to language, or vice versa, does language increase brain size? In addition, language is usually localized in the left hemisphere of the brain, incidentally the same side that controls hand dexterity in right-handers. Why is this so? Do we really have one dominant hemisphere? What is the link between language lateralization and hand preference (Chapter 4)? Finally, interrupting the connections between cerebral hemispheres has significantly informed us about the lateralization of functions and the organization of our brains. But what is the role of this huge tract in non-lateralized animals and how does it contribute to a lateralized brain and the origin of lateralized speech (Chapter 5)?

I will develop my argument starting with the history and current interpretations of the neurology of language (Chapter 2), followed by some general attributes of the human brain like its large size (Chapter 3), lateralization (Chapter 4) and the transfer of information between hemispheres (Chapter 5). I will then introduce Baddeley's concept of working memory and its implications for language development in children (Chapter 6). In the second part of the book, I will review possible evolutionary roots of the phonological loop in the non-human primate brain, presenting the main argument of the book outlined above (Chapter 7). After this, I will discuss the argument that language has its origins in hand gestures, which has been championed by exponents of the mirror neuron hypothesis (Chapter 8). I will then provide an overview of some recent findings on mechanisms of vocal plasticity and learning in other animals, showing how these can be used as models for the early stages of human vocal communication (Chapters 9 and 10). In the last chapter, I will discuss how the phonological loop contributed to amplifying a semantic space that led us to a shared mind and the consequent interpretation of the world, with all its cultural consequences (Chapter 11).

References

Antón SC, Potts R, Aiello LC (2014) Evolution of early Homo: an integrated biological perspective. Science 345:1236828

Baddeley A (2007) Working Memory, Thought and Action. Oxford University Press, Oxford

Berwick RC, Chomsky N (2016) Why Only Us: Language and Evolution. MIT Press, Cambridge

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

Bickerton D (2009) Adam's Tongue. How Humans Made Language, How Language Made Humans. Hill and Wang Press, New York

Bickerton D (2014) More Than Nature Needs. Language, Mind, and Evolution. Harvard University Press, Cambridge

Bruner E, de la Cuétara JM, Masters M, Amano H, Ogihara N (2014) Functional craniology and brain evolution: from paleontology to biomedicine. Front Neuroanat 8:19

Burkhardt RW (2005) Patterns of Behavior: Konrad Lorenz, Niko Tinbergen, and the Founding of Ethology. University of Chicago Press, Chicago

Calvin W (1983) The Throwing Madonna. Mc Graw-Hill Press, New York

Candland DK (1995) Feral Children and Clever Animals. Reflections on Human Nature. Oxford University Press, Oxford

Chomsky N (1957) Syntactic Structures. Mouton de Gruyter, Berlin (2002)

Chomsky N (1965) Aspects of the Theory of Syntax. MIT Press, Cambridge (2014)

Chomsky N (1981) Lectures on Government and Binding. The Pisa Lectures. Mouton de Gruyter, Berlin

Corballis M (1993) The Lopsided Ape. Oxford Univeristy Press, Oxford

Cornélio AM, de Bittencourt-Navarrete RE, de Bittencourt Brum R, Queiroz CM, Costa MR (2016) Human brain expansion during evolution is independent of fire control and cooking. Front Neurosci 10:167

Crompton RH (2016) The hominins: a very conservative tribe? Last common ancestors, plasticity and ecomorphology in Hominidae. Or, What's in a name? J Anat 228:686–699.

Darwin C (1859) On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. Murray Press, London

Darwin C (1871) The Descent of Man, and Selection in Relation to Sex. Murray Press, London

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

Denton M (1985) Evolution: a Theory in Crisis. Burnett Books, Sydney

Desmond A (1984) Archetypes and Ancestors. Palaeontology in Victorian London, 1850–1875. University of Chicago Press, Chicago

Falk D (2009) The natural endocast of Taung (Australopithecus africanus): insights from the unpublished papers of Raymond Arthur Dart. Am J Phys Anthropol 140 (Suppl 49):49–65

Falk D (2014) Interpreting sulci on hominin endocasts: old hypotheses and new findings. Front Hum Neurosci 8:134

Falk D, Clarke R (2007) Brief communication: new reconstruction of the Taung endocast. Am J Phys Anthropol 134:529–534

Falk D, Clarke R (2012) Letter to the editor: response to Holloway and Broadfield's critique of our reconstruction of the Taung virtual endocast. Am J Phys Anthropol 148:483–485

Fischer B, Mitteroecker P (2015) Covariation between human pelvis shape, stature, and head size alleviates the obstetric dilemma. Proc Natl Acad Sci USA 112:5655–5660

Fitch T (2010) The Evolution of Language. Cambridge University Press, Cambridge

Fitch WT, Hauser MD (2004) Computational constraints on syntactic processing in a nonhuman primate. Science 303:377–380

Fitch WT, Hauser MD, Chomsky N (2005) The evolution of the language faculty: clarifications and implications. Cognition 97:179–210; discussion 211–225

Gardner RA, Gardner BT (1969) Teaching sign language to a chimpanzee. Science 165:664–672

Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC. (2006) Recursive syntactic pattern learning by songbirds. Nature 440:1204–1207

Gould SJ (1977) Ontogeny and Phylogeny. Belknap Press, Harvard, Cambridge.

Gould SJ, Lewontin RC (1979) The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proc R Soc Lond B Biol Sci 205:581–598

Gregory MD, Kippenhan JS, Dickinson D, Carrasco J, Mattay VS, Weinberger DR, Berman KF (2016) Regional variations in brain gyrification are associated with general cognitive ability in humans. Curr Biol 26:1301–1305

Gross CG (1993) Huxley versus Owen: the hippocampus minor and evolution. Trends Neurosci 16:493–498

Gunz P, Neubauer S, Maureille B, Hublin JJ (2010) Brain development after birth differs between Neanderthals and modern humans. Curr Biol 20: R921–922

Hauser M (1996) The Evolution of Communication. MIT Press, Cambridge

Hauser MD, Chomsky N, Fitch WT (2002) The faculty of language: what is it, who has it, and how did it evolve? Science 298:1569–1579

Hauser MD, Yang C, Berwick RC, Tattersall I, Ryan MJ, Watumull J, Chomsky N, Lewontin RC (2014) The mystery of language evolution. Front Psychol 5:401

Hayes KJ, Hayes C (1952) Imitation in a home-raised chimpanzee. J Comp Physiol Psychol 45:450–459

Herder G (1800) Outlines of a Philosophy of the History of Man. Bergman Publishers, New York

Holloway RL, Broadfield D (2012) Reply to Falk and Clarke on Taung virtual endocast midline and volume. Am J Phys Anthropol 149:326

Holloway RL, Broadfield DC, Carlson KJ (2014) New high-resolution computed tomography data of the Taung partial cranium and endocast and their bearing on metopism and hominin brain evolution. Proc Natl Acad Sci U S A 111:13022–13027

Hubel DH, Wiesel TN (1977) Ferrier lecture. Functional architecture of macaque monkey visual cortex. Proc R Soc Lond B Biol Sci 198:1–59

Huseynov A, Zollikofer CP, Coudyzer W, Gascho D, Kellenberger C, Hinzpeter R, Ponce de León MS (2016) Developmental evidence for obstetric adaptation of the human female pelvis. Proc Natl Acad Sci U S A 113:5227–5232

Jackendoff R (1999) Possible stages in the evolution of the language capacity. Trends Cogn Sci 3:272–279

Jackendoff R, Pinker S (2005) The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser and Chomsky). Cognition 97:211–225

Kimbel WH, Villmoare B. (2016) From Australopithecus to Homo: the transition that wasn't. Philos Trans R Soc Lond B Biol Sci 371:1698

Lenneberg EH (1967) Biological Foundations of Language. John Wiley Press, New York

Lewontin RC (1975) Genetic aspects of intelligence. Annu Rev Genet 9:387–405

Liberman AM, Mattingly IG (1989) A specialization for speech perception. Science 243:489–494

Liberman AM, Harris KS, Hoffman HS, Griffith BC (1957) The discrimination of speech sounds within and across phoneme boundaries. J Exp Psychol 54:358–368

Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M (1967) Perception of the speech code. Psychol Rev 74:431–461

Lorenz K (1981) The Foundations of Ethology. Springer Verlag, New York

Mitteroecker P, Fischer B (2016) Adult pelvic shape change is an evolutionary side effect. Proc Natl Acad Sci USA 113:E3596

Neubauer S, Gunz P, Hublin JJ (2010) Endocranial shape changes during growth in chimpanzees and humans: a morphometric analysis of unique and shared aspects. J Hum Evol 59:555–566

Oakley K (1949) Man the Tool-Maker. Bulletin of the British Museum of Natural History, London (1972)

Owen R (1837) The Hunterian lectures in comparative anatomy. University of Chicago Press, Chicago (1992).

Patterson F, Linden E (1985) The Education of Koko. Holt, Rinehart and Winston Press, New York

Penfield W, Roberts L (1959) Speech and Brain Mechanisms. Princeton University Press, Princeton

Perruchet P, Rey A (2005) Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates? Psychon Bull Rev 12:307–313

Piantadosi ST, Kidd C (2016) Extraordinary intelligence and the care of infants. Proc Natl Acad Sci U S A 113:6874–6879

Pinker S (1994) The Language Instinct. Penguin Books, London

Pinker S, Jackendoff R (2005) The faculty of language: what's special about it? Cognition 95:201–236

Plummer T (2004) Flaked stones and old bones: biological and cultural evolution at the dawn of technology. Am J Phys Anthropol Suppl 39:118–164

Ponce de León MS, Huseynov A, Zollikofer CP (2016) Reply to Mitteroecker and Fischer: developmental solutions to the obstetrical dilemma are not Gouldian spandrels. Proc Natl Acad Sci U S A 113: E3597–3598

Rey A, Perruchet P, Fagot J (2012) Centre-embedded structures are a by-product of associative learning and working memory constraints: evidence from baboons (Papio Papio). Cognition 123:180–184

Richards RJ (1989) Darwin and the Evolutionary Theories of Mind and Behavior. University of Chicago Press, Chicago

Rightmire GP (2013) Homo erectus and Middle Pleistocene hominins: brain size, skull form, and species recognition. J Hum Evol 65:223–252

Rizzolatti G, Arbib MA (1998) Language within our grasp. Trends Neurosci 21:188–194

Rizzolatti G, Craighero L (2004) The mirror-neuron system. Annu Rev Neurosci 27:169–192

Sankararaman S, Mallick S, Patterson N, Reich D (2016) The combined landscape of Denisovan and Neanderthal ancestry in present-day humans. Curr Biol 26:1241–1247

Schiller F (1992) Paul Broca. Explorer of the Brain. Oxford University Press, Oxford

Seyfarth RM, Cheney DL, Marler P (1980) Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. Science 210:801–803

Tinbergen N (1951) The Study of Instinct. Clarendon Press, New York

Trevathan W (2015) Primate pelvic anatomy and implications for birth. Philos Trans R Soc Lond B Biol Sci 370:20140065