

9 Diversity in languages, genes, and the language faculty

JAMES R. HURFORD AND DAN DEDIU

9.1 Introduction

In the literature on language evolution, one frequently finds phrases such as “ancestor language,” “the first human,” and “the language faculty.” The first two of these suggest the existence in the past of single unified entities from which modern languages or humans are descended in their entirety. The third expression, “the language faculty”, suggests a synchronic unity with the implication that it too could have had a single unified source. At the level of expository metaphor, such expressions may have their uses. Here, as a cautionary exercise, we argue that such metaphors widely circulating in both technical and popular scientific discourse are overused and project a too simplified perception of extremely complex phenomena. Our point is quite general, and can be appreciated without recourse to technical detail, although this does not mean that the technical details don’t support our case. In this context, we submit that there are multiple sources (or “cradles”) of:

1. **individual languages**, such as English, Afrikaans and Xhosa—these are varied, there being, for example, as many different varieties of English as there are English speakers; and “genealogical” relations between languages are not consistently divergent;¹
2. the **human genome**, which is not a single, uniform entity across our species, as shown by the HAPMAP project; in this domain, too,

Dan Dediú was funded by an ESRC (UK) postdoctoral fellowship. Both authors wish to thank Karl Diller and Rebecca Cann for their comments on an earlier draft.

¹ A tree diagram is consistently divergent if it is never convergent, that is if there is only ever one path from the root of the tree to any given daughter node.

“genealogical” relationships as revealed by the genetic data are far from simple and tree-like;

3. the **human language capacity**, which is not a single monolithic capacity, but a dynamic, evolving one, resulting from the complex interaction of biology and culture.

Each of these is a mosaic with many sources, and all but the most recent of these was somewhere in Africa, or in the continent from which Africa was formed, but at different times.

A number of theoretical tools and hypothetical concepts circulating in scientific discourse contribute to oversimplified beliefs that the phenomena listed under 1–3 above are unitary and have single sources, which can be pinpointed to a single era in evolution and a single geographical region. Such potentially misleading concepts include: protoworld etymologies implying a single mother language; tree diagrams of language families and of human phylogeny with a single root and no reticulation; “Mitochondrial Eve” and “Y-chromosome Adam” (see below) suggesting a common time when both male and female most recent common ancestors (MRCAs) lived; even speciation, when used to represent a clear-cut evolutionary leap at a particular point in time; the human genome, suggesting uniformity across the species; and finally the human language faculty, as if it were a single monolithic entity uniform across the species. We focus on these because they represent, and generate, the most salient oversimplified ideas about which we wish to encourage due caution.

Mitochondrial Eve existed; there was a woman from whom all extant human mtDNA is inherited. Likewise Y-chromosome Adam existed. An example of the over-simple way in which these labels can be interpreted is found even in a scholarly psychological monograph: “Evolutionists say that there was a first human, and, on the basis of DNA evidence, that this human was a woman” (Paivio 2006: 283). The explicit allusion to the biblical myth, suggesting that this Adam and Eve cohabited, and that we are all the fruit of their union, is of course misleading, as is widely recognized. But even if we are careful to avoid a romantic Garden-of-Eden scenario, the very mention of particular individuals as somehow privileged ancestors of all that is to be found in the modern genome is misleading. For any pair of our roughly 30,000 genes, there is no implication at all that their modern variation can be traced to the same single individual as their most recent common ancestor.

Diversity in languages, genes, and the language faculty 165

When discussing the relative merits of a metaphor, it is always important to specify in which contexts it is useful and where it starts breaking down. Thus, we do not deny the value of the metaphors of the (human and language evolutionary) cradle, genealogical trees or the human language faculty; these certainly represent contextually valid approximations and operationalizations of a complex reality. *In extremis*, without such metaphors, science would be unthinkable. But continuing to use them beyond their limits risks distortion of reality. Tree diagrams, for example, are seductive. They are a handy way of visualizing relationships. Unfortunately, they are often used in diametrically opposed ways, with time correlated with either divergence or convergence of lines in a tree. Figures 9.1 and 9.2 give two common examples.

Both trees are “family trees,” but note that they relate to the dimension of time in directly opposite ways. The phylogenetic tree branches forward in time; the royal family tree branches backward in time. Both trees show ancestry, and both are oversimplifications. A more realistic diagram would combine properties of both figures. Given a chosen period of time, and some chosen set of related entities existing during that time, be it individual people, or species, or languages, the diagram would show all their ancestors and all their descendants within the chosen time frame, resulting usually in a lattice. In such a lattice, there would be examples of both

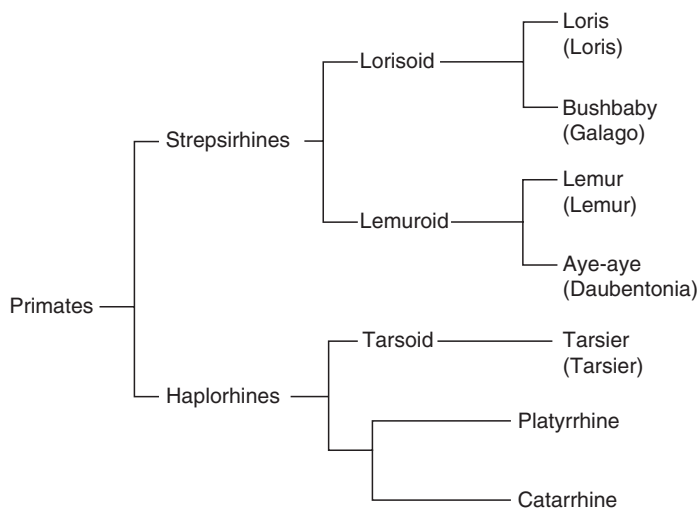


FIG. 9.1 Phylogenetic tree.

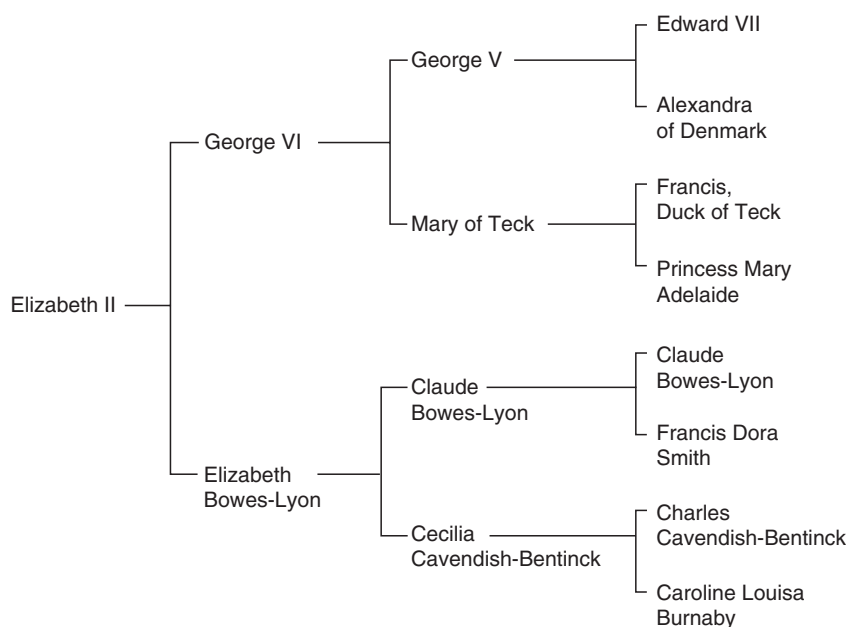


FIG. 9.2 Royal family tree.

divergence (many descendants of one entity) and convergence (many ancestors of an entity) over time—see Figure 9.3 for a simple example. Given the nature of biological species (discussed below), diagrams representing relationships between them would be very well approximated by a tree in most cases and not a lattice, unlike a diagram of family relationships among individuals. Therefore, this metaphor must be used and interpreted in conformity to its actual context, its representational power must be clearly specified in each case, and alternative representational methods must be employed when necessary (Jobling et al. 2004).

There are cases where an element in the lineage of a genome or of a language diverges at some point from other elements and is temporarily (maybe for a long time) passed down along a separate lineage from them, but later rejoins their lineage. In the case of species, this can happen where, for example, a species splits into two populations which have little contact for a long time but then intermix again (hybrid zones due to secondary contact; e.g. Skelton 1993: 382). A possible example relevant for human evolution is represented by a locus on the X chromosome (Xp21.1) for

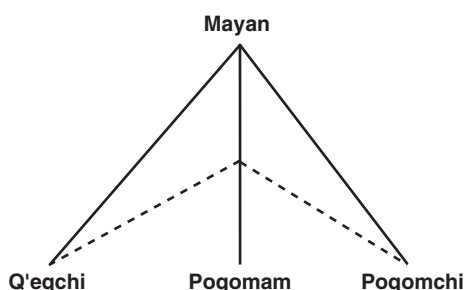


FIG. 9.3 Three separate daughter languages from a common stock. After their separation, a sound change in one is diffused to the others.

which a non-coding sequence of 17.5 kb length has been identified in two African individuals which has not recombined with other lineages for over a million years, suggesting that this X chromosome lineage evolved in isolation from the other lineages (Garrigan et al. 2005a). In the case of languages, an example is given by Campbell (2004: 198): “Q’eqchi’, Poqomam and Poqomchi” [Mayan languages] share change (18) (**ts > s*); however, documents from the sixteenth and seventeenth centuries reveal that this change took place long after these three were independent languages and that the change is borrowed, diffused across language boundaries.” This detail of the history of these Mayan languages is shown in Figure 9.3.

The “cradle” metaphor aptly suggests subsequent growth and change of the entity that starts life as the “baby,” with contributions that were in no way present in any original “blueprint.” Another merit is to bring into focus the special (geographically extensive) place and (evolutionary long stretch of) time represented by Africa between about 2 million years ago and 100 thousand years ago for human and language evolution. And yet another undeniable merit is to highlight the adherence to the evolutionary stance, whereby descent with modification from common ancestors due to random or selective factors represents the fundamental key to modern biology.

It would be wrong to overinterpret the “cradle” metaphor as suggesting a particular moment of conception of a single continuing unified entity (the “baby”) which somehow remains “the same thing” despite all the changes and innovations that it undergoes. For practical purposes, societies are prepared to accept that persons remain in some sense identifiably

“the same thing” throughout their lives, sometimes implicitly qualified by a statute of limitations. But, as we will illustrate, it is misleading to assume a single unified source for all the genes that a person carries. Furthermore, those genes did not all appear on the scene at the same time. Even the term *lineage*, applied to a person, has erroneous connotations, suggesting a single line of descent for the totality of a person’s genes, with no tributaries or distributaries.

We suggest that weight should also be given to another powerful metaphor, the “melting pot,” where new entities are forged from multiple sources. The modern USA is a melting pot, whose population comes from all over the world. It makes no sense to speak of the distinctive “ancestors of (all) modern Americans,” in the sense in which it might just be sensible to conjecture about the distinctive ancestors of the Ainu or the Andaman Islanders. To be sure, all modern American humans are descended from the same stock as all other humans, but that stock has branched out and later recombined, at different time depths.

In successive sections of this paper we will discuss the diversity and multiple sources of human languages, the human genome, and the human language faculty.

9.2 Languages are confluences of features from many sources

Let us first recognize that the notion of “a language” is itself no more than a useful simplification.

Dan Dediu and Jim Hurford both speak English, but is it the same language? It depends on how fine-grained you want the answer to be. At a level of fine detail, we don’t speak the same language; Dan has a Romanian accent and Jim has traces of a British regional accent, and there are differences at lexical, grammatical, and pragmatic levels. But for all practical purposes, it is useful to say that the English that we speak is the same language. OK, so we’ll accept this simplifying idealization of “English” as a unified entity. But does this English have a single unified source, as is implied by traditional family tree diagrams showing it as having a single lineage, back through Proto-West Germanic and Proto-Germanic, to Proto-Indo-European? No.

What is the mother language of English, Proto-Germanic, or Proto-Romance? But why not in part Proto-Afro-Asiatic, as English has borrowed

Diversity in languages, genes, and the language faculty 169

algorithm, alcohol, and other words from an Afro-Asiatic language? Or why not in part Proto-Eskimo-Aleut, as English has borrowed words (e.g. *kayak, igloo*) from languages of this stock? The sensible pragmatic answer is of course that only a few tiny bits of English come from these sources. When using trees, we are neglecting these minor contributions in the interests of highlighting the hypothesis of a single proto-stock that we can think of as representing our *Urmuttersprache*. Therefore, since these few borrowings are such a slim part of English, we will ignore them in the following, and return to the more mainstream question of whether English is Germanic or Romance. The only sensible answer is that it is a bit of both. English, like French and Italian, lacks the case systems and verb-final subordinate clauses of its closest Germanic relatives, German and Dutch. English has vocabulary derived from both Germanic and Romance sources. However, the basic vocabulary of English (including kin terms, numerals, and bodypart terms) is Germanic. But why give those words a privileged status, unless for the purpose of highlighting the Germanic nature of English? That last sentence, all of it impeccable English, had a mixture of Germanic words (*give, word, the*) and Romance words (*status, nature, and even Germanic*). Thus, where is the source of English? There is no single source: “English” has some Germanic, some Romance, some tiny Sino-Tibetan components, etc., etc. The received wisdom about English is that it is Germanic, because that is where its basic vocabulary comes from. But in some basic respects it has French-like syntax (lack of cases, SVO word order).

A completely realistic diagram of the historical sources of English would not be a tree, but a lattice showing how different parts of the language had different sources. This is not to deny that large slices of a language can have common sources. With enough graphic ingenuity it is possible to draw a lattice in such a way that the genuine tree-like relationships stand out, perhaps shown as heavier lines. The discussion here echoes, of course, a debate that raged within historical linguistics in the nineteenth century between proponents of family tree (*Stammbaum*) theory and wave theory. Wave theorists (e.g. Schuchardt 1868; Schmidt 1872) proclaimed that “chaque mot a son histoire” (every word has its own history). This unfortunately ignores generalizations across words that, for example, have undergone the same sound change.

But, equally, “genetic relationship, the only thing represented in family-tree diagrams, is not the only sort of relationship that exists among

languages—for example, languages do also borrow from one another” (Campbell 2004: 212). But note even here, in this quotation from a mainstream historical linguist, the presupposed mutual exclusivity of “genetic relationship” and borrowing. The marginalization of borrowing is endemic in the literature, as a quotation from another linguist, diametrically opposed to Campbell on many issues, shows: “Linguists employ a number of well-known techniques to distinguish borrowed words from inherited items” (Ruhlen 1994a: 279). Why is borrowing less “genetic” than other language changes? The only difference is in the source—a feature inherited from a minority source is labeled “borrowing,” while a feature inherited from a majority source is “genetic.” The Celtic population of Gaul switched during the five centuries of Roman occupation from speaking a Celtic language to speaking a mainly Romance language, leaving behind only a few Celtic relics, such as the partly vigesimal numeral system. This conversion process started by some Celtic speakers borrowing some Romance words, then over time more words were borrowed, until almost the whole vocabulary was of Romance origin. The current allegedly “genetic” Romance status of French is a result of wholesale borrowing! Mufwene (2001: 109–112) makes a similar point, also mentioning contact between Romance and Celtic languages, in a section entitled “How language contact has been downplayed.”

It might be argued that whole Celtic and Vulgar Latin languages existed in parallel, and that speakers were either monolingual or bilingual, with a gradual population shift to monolingual Romance speakers. This idealized scenario would preserve the “genetic” integrity of the two systems, but it ignores the widespread phenomenon of code-switching in such contact situations, giving rise to a mish-mash language, which in this case would have been partly Celtic, partly Romance. It is likely that anyone born in Gaul toward the end of the Roman occupation spoke a variety which was a mixture of originally Latin features and originally Celtic features. Over time, the ex-Latin features came to dominate. It might also be argued that a difference between genetic/inherited traits and borrowed traits is that the former are a product of children learning their language from their parents, while borrowed traits are adopted by adult speakers. We doubt that any such sharp distinction can be sustained.

Ruhlen (1994b: 272) writes, “all the world’s languages share a common origin.” This gives the impression that there was once a language, as complete and complex as any extant human language, which was the

Diversity in languages, genes, and the language faculty 171

Mother Tongue. Ruhlen and his co-author Bengtson have distanced themselves from this view (Bengtson and Ruhlen 1994), but other workers in the same “macro-comparatist” programme have used such suggestive titles as “The mother tongue: how linguists have reconstructed the ancestor of all living languages” (Shevoroskin 1990).

Ruhlen’s dominant theme, which he pursues in common with other macro-comparatists, such as Greenberg and Shevoroskin, is classification of languages as if they were all the same kind of entity in all respects relevant to the classification. Features deemed irrelevant to the classification are ignored, or marginalized as borrowing. Both Merritt Ruhlen and his vociferous opponent in matters of linguistic reconstruction, Lyle Campbell, are staunch family-tree men; they both picture the significant relationships between languages as ever-divergent trees. Where these scholars differ, irreconcilably, is in the time-depth at which it is possible to postulate ancestor forms. Ruhlen and colleagues believe that some form—meaning pairings survive recognizably enough and across such a range of language families that one can postulate “proto-world etymologies.” Such claims have been the subject of fierce controversy, on which we take no stand here.

Suppose that Merritt Ruhlen is right and there were indeed at least 26 single protoforms from which words that can be found in most modern language families are derived; then this only tells us about the mothers of forms for those meanings. It does not reconstruct any single (presumably African) mother language of all human languages. The etymologies of many other words that may have co-existed with proto-world *TIK (= “finger, one”) and proto-world *PAL (= “two”) would have come to evolutionary dead ends long ago. And many completely new words were coined, in different languages, long after the existence of the proposed proto-world. Ruhlen (1994a) himself is careful to say that he is not attempting reconstruction of proto-world, but only postulating global etymologies; however, the subtitle of his other book published in the same year, “tracing the evolution of the mother tongue” (Ruhlen 1994b), definitely suggests an attempt to describe an actual historic entity, the “mother tongue.”

It is sometimes suggested (Bickerton 1990, 1995) that *Homo erectus* may have had “protolanguage,” i.e. a syntaxless vocabulary-only form of communication, and that fully syntactic language came with *Homo sapiens*. If so, and if Ruhlen’s global etymologies have any validity, then these

ancient form—meaning pairings could conceivably be of far greater antiquity than even Ruhlen has dared to suggest. But in this case, the ancestral form of communication that contained such pairings would not have been a language in a fully modern sense, since it had no syntax.

9.3 Multiple sources and heterogeneity of the human genome

Now turning to the human genome, a recent study (Zerjal et al. 2003) suggests Genghis Khan's direct patrilineal descendants today constitute about 8% of men in a large area of Asia (about 0.5% of the world population). Thus the male most recent common ancestor (MRCA) of these men is much more recent than their female MRCA. The MRCA of human mitochondrial genes is probably of greater antiquity than the MRCA of human Y-chromosome genes. The same applies to all characteristic human genes. Some are (much) older than others. mtDNA and the Y-chromosome are a tiny proportion of human DNA. For the general case, Dawkins has made the point that for particular genes, an individual human may be more closely related to some chimpanzees than to some humans. Blood groups are an example; a man may have the same blood group as a chimpanzee but have a different blood group from his wife: "every gene has its own tree, its own chronicle of splits, its own catalogue of close and distant cousins... individuals are temporary meeting points on the criss-crossing routes that take genes through history" (Dawkins 2004).

Of the particular genes affecting human language, they also vary in antiquity. The human variant of FOXP2 is widely claimed to have appeared within the last 200,000 years, although a study presented in the current volume (Diller and Cann this volume) claims much greater antiquity—a claim also supported by the very recent finding that modern humans and Neanderthals share this variant (Krause et al. 2007). According to a recent study (Dediu and Ladd 2007), variants of two more genes, the derived haplogroups of *ASPM* and *Microcephalin*, may also be relevant to human language. These recently evolved variants are rare in Africa, probably originated outside Africa, and are still under positive selection, not being yet fixed in the human population. Dediu and Ladd claim that there is a correlation between the frequencies of these variants in a population and the usage of tone contrasts in the language(s) spoken by it.

Diversity in languages, genes, and the language faculty 173

They argue that this correlation is non-spurious, in the sense that it cannot be explained by other factors. The mechanism linking genes and tone could be tiny acquisition and/or processing biases affecting the cultural transmission of language, and thereby influencing the trajectory of language change.

There is no single story the genes can tell; each bit of DNA potentially has something different to say, if properly asked. Each gene can recount its own version of history, its jumping from body to body across generations, its struggle to outsurvive its competitors by making the bodies it inhabited better than the others in innumerable various ways. We must take the intrinsic diversity of these stories into account while trying to create a faithful reconstruction of the past.

Probably the best-known bits of our genome are represented by mitochondrial DNA and the Y-chromosome, the first (mtDNA) being transmitted down the generations exclusively through the maternal line (Jobling et al. 2004; Seeley et al. 2005; Lewin 2004) while the second contains a segment (NRY: the non-recombining part of the Y-chromosome) which is exclusively transmitted through males (Jobling et al. 2004). This property makes them very well-suited for evolutionary and historical studies, because their history is the history of each sex, separately: mtDNA tells us the adventures of the females while NRY tells those of the males—at least, as a first approximation.

But even in these simple cases things get very complex. There is a much greater difference in fitness (reproductive success) among men than among women. Due to the special way in which both mtDNA and NRY are transmitted, it is a logical necessity that for any group of humans, living, extinct or a combination thereof, there can be found a single individual (female or male, respectively) from which all the group's variants of mtDNA or NRY originated (Dediu 2007; Relethford 2001). This individual represents the MRCA of the genetic variants present in the specific group under study.

In their seminal study, Cann et al. (1987) reconstructed the MRCA of living humans' mtDNA as dating from approximately 200,000 years ago and probably located in Africa, and ignited the popular imagination with an African mitochondrial Eve from which all mtDNA stems. Shortly afterwards, the parallel concept of a Y-chromosome Adam appeared, which, as expected, has a different age than the mtDNA Eve, approximately 60,000 years (Thomson et al. 2000; Underhill et al. 2000).

When leaving the special cases of these sex-linked genetic systems (mtDNA and NRY) and moving into the realm of recombining genes, the story becomes much more complex, as the history told by such a gene has no intuitive counterpart at all. And, again, these histories do differ, sometimes remarkably so. For example, the vast majority of the genes of living humans seem to come from Africa, but the ages of their MRCA are widely different. Some are fairly recent (the derived haplogroups of *ASPM* and *Microcephalin*, estimated at some 5,000 and 37,000 years ago), others are old (predating the chimp—human split, like some alleles of the major histocompatibility system; e.g. Loisel et al. 2006), and yet others are extremely old (predating the vertebrate splits; e.g. Venkatesh et al. 2006).

The complex and varied histories of genes are further illustrated by this example. A segment of the X chromosome (the Xp21.1 locus) presents a very rare lineage confined to certain African populations which seems to have evolved in isolation from the other lineages for more than 1 million years (Garrigan et al. 2005a) suggesting the existence of long-lasting splits inside our species. Other parts of the X chromosome have even stranger stories to tell, including the *HS571B2* locus (Yu et al. 2002), presenting a variant which is suggested to have arisen in Eurasia more than 140,000 years ago, or the segments of the *Dystrophin* gene analysed by Ziętkiewicz et al. (2003), having three lineages, one of them suggesting a non-African origin earlier than 160,000 years ago. But probably the most striking example is represented by the *RRM2P4* pseudogene (Garrigan et al. 2005b), which has an old MRCA (around 2 million years ago) and probably an Asian origin. Of course, all these examples could in fact be due to statistical error, but if not, then not only does their existence highlight the diversity of points of view carried down the ages by different genes, they also throw some doubt on the standard model for human evolution, which posits a recent African origin for modern humans, followed by a rapid expansion across the world with the total replacement of the pre-existing local archaic forms (for a full discussion, implications and class of most probable models, see Dediu 2006, 2007).

One might argue that there may be diversity among human genes, but still there is a single human genome; after all, we are such a uniform species. And in some fundamental way, this is right. However, as shown by the HapMap project (The International HapMap Consortium 2003; www.hapmap.org), taking into account the diversity of our species is important not only for understanding our origins and history, but also for fighting

Diversity in languages, genes, and the language faculty 175

disease and promoting health and quality of life. While it is true that humans are much more uniform than other comparable species (Jobling et al. 2004; Relethford 2001), this does not entail that we are genetic clones. There is a pervasive claim, often cited without any reference, that humans are so uniform and unstructured that the division of *Homo sapiens* into groups is not justified by the genetic data, and people all over the world are much more similar genetically than appearances might suggest. This is formulated by Edwards (2003: 798) as the claim that about 85% of the total genetic variation is due to individual differences within populations and only 15% to differences between populations or ethnic groups, a claim which can be traced to the work of Richard Lewontin (1972). However, this simplifying claim is misleading as it neglects the fact that the structure of the human species is not given by a few independent diagnostic genes, but by the correlations between the frequencies of many different alleles across populations (Jobling et al. 2004; Rosenberg et al. 2002; Bamshad et al. 2003). Thus, there is enough genetic structure to allow reliable prediction of population of origin using a limited number of loci; however, it is not population-specific loci which allow this classification but their correlational structure.

Thus, there is genetic diversity across the human species and each gene has a different history. This inescapable conclusion could potentially have a significant impact on our efforts to understand the evolution of language, suggesting that the evolved language capacity consists of elements with different genetic histories. There has always been a tendency to see language as an all-or-nothing phenomenon, brought into existence by some sort of explosion or sudden revolution. A recent example is Tim Crow's (2002b) effort at identifying a single gene that played a critical role in the transition from a precursor species to modern *Homo sapiens*, hypothesized to be the *protocadherinXY* gene located in the X-Y homologous region. Another theory involving a single gene bringing about language concerns FOXP2, a gene of the forkhead box family which act as transcription regulators (Lai et al. 2001; Scharff and White 2004). Heterozygous carriers of deleterious mutations of this gene develop a complex phenotype including articulatory problems, cognitive impairments and language impairments (Bishop 2003; Fisher et al. 2003; Vargha-Khadem et al. 1998; Lai et al. 2003; Watkins et al. 2002a, b), which suggested to some that this gene might have something specifically to do with language. Moreover, evolutionary considerations suggested that the

human-specific form of the gene appeared during the last 200,000 years of human history, that is, concomitant with or subsequent to the emergence of anatomically modern humans (Enard et al. 2002), boosting the claims that this might be the gene explaining language, modernity, and everything else. However, it turns out that this story is much more complex (Dediu 2007: 111–120), that the estimation of this age is fraught with difficulties, that the human-specific variant is not that specific to humans after all (Webb and Zhang 2005; Zhang et al. 2002), that in birds and vocal-learning mammals FOXP2 does not seem to explain much (Webb and Zhang 2005; Teramitsu et al. 2004; Scharff and Haesler 2005; Haesler et al. 2004; Shu et al. 2005) and, finally, that the human variant is much older (Diller and Cann this volume; Krause et al. 2007). In the end, it seems that the effects of FOXP2 are much more subtle than simply enabling language, probably creating a permissive environment in which vocal learning can evolve if other circumstances/factors come into play (Scharff and White 2004: 342).

Alternative models of language evolution, involving the slow, gradual accretion of various aspects of our linguistic capacity, have been proposed before (e.g. Pinker and Jackendoff 2005; Smith 2006; Corballis 2004; Hurford 2003a). Theories of this type require that small genetic changes impacting (not necessarily directly) on language are selected, and increase in frequency until eventually reaching fixation. However, this standard neo-Darwinian account essentially implies population-level genetic variability concerning language, an idea not seriously considered in linguistics and allied disciplines (e.g. Cavalli-Sforza et al. 1994), which might seem unexpected given the amount of data from behavior genetics suggesting high genetic components of inter-individual abilities and disabilities connected to language (Dediu 2007; Stromswold 2001).

The possible nature of this mechanism was suggested in a recent study (Dediu and Ladd 2007), where the inter-population diversity of two brain growth and development-related genes was related to the distribution of tone languages. *ASPM* and *Microcephalin* are two genes whose deleterious mutations cause primary recessive microcephaly (Gilbert et al. 2005; Cox et al. 2006; Woods 2004) and for which two derived haplogroups have been identified (denoted in the following as *ASPM-D* and *MCPH-D*, respectively), showing signs of ongoing natural selection in humans (Mekel-Bobrov et al. 2005; Evans et al. 2005). These

Diversity in languages, genes, and the language faculty 177

haplogroups have appeared recently (approximately 5,000 and 37,000 years ago, respectively) and *MCPH-D* even seems to have introgressed into the modern human lineage from another archaic form (Evans et al. 2006). In spite of many attempts, the phenotypic effects of these haplogroups which explain the selective pressure have not been found: They seem not to be connected to intelligence (Mekel-Bobrov et al. 2007), brain size (Woods et al. 2006), head circumference, general mental ability, social intelligence (Rushton et al. 2007), or the incidence of schizophrenia (Rivero et al. 2006). The proposal of Dediu and Ladd (2007) is that *ASPM-D* and *MCPH-D* might determine a very small bias at the individual level in the acquisition or processing of linguistic tone, a bias which can be amplified in a population through the cultural transmission of language across generations, and manifested in differences between the languages spoken by such populations. They support this hypothesis by the fact that the population frequencies of *ASPM-D* and *MCPH-D* correlate negatively with the usage of linguistic tone by that population, even after geography and shared linguistic history have been controlled for.

That such biases can work has been suggested previously by both computer models (Smith 2004; Nettle 1999b) and mathematical models (Kirby et al. 2007), but, if confirmed by further experimental studies, this would represent the first case of a genetically influenced linguistic bias manifest at the population level. And this type of bias could represent exactly the mechanism required for gradual, accretionary models of language evolution, whereby small genetic changes appear, influence the capacity for language in various populations, and eventually became part of the universal linguistic capacity. This model suggests that linguistic and genetic diversities are the key for understanding the universal properties of language.

The human language capacity is commonly said to be uniform across the species. Certainly, a baby born of Chinese parents and adopted into a French-speaking family will learn French just as easily as it would have learned Chinese. But the affirmation of uniformity comes with a typical reservation that it excludes pathological cases. The pathological cases are certainly still human, so the language faculty is not in fact uniform, and there is no principled way of separating cases defined as pathological from the tail of a distribution, so it seems likely that even among non-pathological cases there is some variation in the language faculty. It is well

established that there are differences in aptitude for second-language learning (see an extensive literature in applied linguistics with Carroll (1962) as an early example). It would be surprising if some of the differences in second-language learning were not also reflected in differences in first-language learning. If the language faculty evolved by natural selection of advantageous variants (not in reasonable doubt), there must have been variability in the evolutionary precursors of the language faculty. One possible variable is the different dispositions of individuals to innovate linguistically; some language users are more creative with their language than others, pushing it beyond current limits.

Obviously, innovation had to be involved in the evolution of languages to their current complex state. New words, new constructions, and new phonemic distinctions arose. We do not envisage that such innovation was necessarily deliberate or a matter of conscious choice. So a disposition in some individuals to innovate is necessary for a language system to get off the ground. But a disposition to innovate is not necessary to maintain a language in a population, once the system is already up and running. All that is required is a capacity to acquire the language of the community.

This theoretical point is made convincingly by Smith (2002), who computationally modeled various postulated innate strategies for learning arbitrary meaning—form pairings, i.e. vocabulary items. Repeated cultural transmission of the vocabulary is modeled, with one generation producing examples of the form—meaning pairs they have learned, for the next generation to learn from. Initially, at “generation zero,” the population has no common vocabulary, and the whole population is genetically uniform, having the same postulated vocabulary acquisition bias. The learners were modeled with little neural nets mapping between meanings and forms, and the different learning biases investigated were modeled by using different weight update rules. Initially, the members of this artificial population produced random forms for the meanings they were prompted to express, and the observers of these form—meaning pairs responded by internalizing weightings of their preferences of form—meaning mappings, as dictated by their innate learning mechanism (i.e. their weight update rule).

In this way it was possible for Smith to compare the effects of 81 different theoretical innate biases applied to the task of vocabulary learning. And, given that the population always started with no common

Diversity in languages, genes, and the language faculty 179

vocabulary, it was possible to see under what circumstances a common vocabulary emerged, suitable for consistent communication about the meanings involved. In some cases no system emerged at all, with the simulated agents merely continuing to produce random signals at each other, and not building up a common vocabulary. In other cases, with different innate learning biases, a system got off the ground, and could be used for consistent communication. The difference between the two cases is between a population-wide innate bias enabling the group to construct a communication system, and on the other hand a similarly shared bias which does not enable the group to progress beyond producing random signals which cannot be consistently interpreted by other group members. Smith accordingly labeled a particular subclass of biases as system constructors. In some sense, agents with one of these biases could impose order on chaos, very much in the sense in which, in the Chomskyan picture of language acquisition, children induce a coherent linguistic competence from degenerate data. Other innate biases were ineffective at constructing a system in this way, but Smith showed that a further subclass of them, which he labeled maintainers, could acquire a system already established in the population and use it effectively in communication. The behavior of these maintainers was consistent enough for the system to be faithfully transmitted to the next generation of learners. All constructors are maintainers, but not all maintainers are constructors.

While specifically concerning the vocabulary, this result could have more general implications in that it is quite possible for a population that has in the past developed a consistent system to be genetically heterogeneous (polymorphic) with respect to their language acquisition dispositions. The early stages of evolution need a critical mass of system constructors, but once a system is constructed, maintainers who are not themselves richly enough endowed to be constructors can function communicatively in the group and pass on the system to their children. Given the extent of polymorphism generally, in humans as in other species, some degree of polymorphism in the language faculty should not be surprising. If linguistic innovation is occasional and sporadic, it would not be immediately evident that there were different dispositions in the population. Indeed it is theoretically possible, though unlikely, for the constructors to become extinct, with the continuance of the communication system sustained culturally by the remaining maintainers.

9.4 Varying antiquity of the human language faculty

In this section, after some definitional preliminaries, we discuss various aspects of the human language faculty, making a rough division between recent features which have evolved only in humans to any significant degree, and ancient features which are found in other animals, especially primates. For the more recent aspects of the language faculty, such as a specialized vocal tract, and episodic memory, it seems likely that they evolved during the emergence of *Homo sapiens* and therefore in Africa. For the more ancient aspects of the language faculty, such as basic syllabic organization, mental reference to objects, and the rudiments of propositional form (kept private), they certainly evolved or at least began to evolve long before the emergence of humans, and some are probably so ancient as to predate the formation of the continent of Africa, over 100 million years ago.

Hauser et al. (2002) make a useful distinction between the faculty of language in the broad sense (FLB) and the faculty of language in the narrow sense (FLN). FLN includes only that which is special to language and is found in no other human cognitive domain or animal communication system. Hauser et al. (2002) suggest that FLN may consist of nothing more than the human capacity for recursive computation, and perhaps not even that, if examples can be found of recursion in non-linguistic systems, such as animal navigation. This distinction helps to clarify what researchers are interested in as denoted by the vague term *language*. In the recent history of linguistics, generative linguists have focussed on language in the narrow sense, aiming at a theory of FLN. Sometimes they have avoided the overly general term *language* and used *grammar* instead, referring to just the formal organization of the sound—meaning pairing system represented in the brain. Other linguists have cast their net more widely, investigating aspects of language use (e.g. discourse analysts and phoneticians) or the interaction of non-linguistic factors, such as short-term memory, on laboratory examples chosen to highlight grammatical contrasts (psycholinguists). Such researchers are investigating FLB. FLB includes anything involved in the learning, mental storage, and use of language, capacities which may well be also used for non-linguistic purposes. We write here of FLB. It is important to note that even FLB is unique to humans; it is a unique combination of traits that can

Diversity in languages, genes, and the language faculty 181

be found in other activities and also in some animals. The individual components of FLB are not unique to human language (by definition), but their combination, which makes us unique among animals, is unique.

“Used for nonlinguistic purposes” has a paradoxical ring to it in the context of language evolution, where things in fact happened the other way around. The language faculty, in the broad sense, was assembled out of capacities and traits that initially had nothing to do with language (because language didn’t yet exist), but which were exapted (Gould and Vrba 1982) and became used for linguistic purposes.

The vocal apparatus is a prime example. The lungs, trachea, larynx, tongue, and lips were variously used for breathing and eating. These anatomical structures had their earliest “cradle” in the very ancient past, long before the continent of Africa was formed. The vocal tract, like the brain, has undergone radical evolution since the split from chimpanzees, most plausibly in the service of the capacity to make ever finer phonetic distinctions (Lieberman 1984). In the narrow generative view, the cognitive faculty of language is independent of its output modalities, since, as deaf sign languages teach us, the same expressive power can be achieved without the use of the vocal tract. Nevertheless, the vocal/aural medium is the dominant output modality for language, and the human vocal tract is unique among primates in the range of distinctive sounds it can produce.

The physiological details of the human vocal tract are an example of relatively recent evolution, having happened over the past 3 million years, at the very most. It seems likely that there were also very significant cognitive developments over the same period, perhaps including the advent of a developed capacity for recursive computation. One such relatively recent cognitive development is the emergence of episodic memory. Episodic memory is memory for specific events, located at particular points in time. Episodic memory is what is lost in amnesics, who, for instance, cannot recall where they woke up this morning, or any specific events of their former lives. But such amnesics have good “semantic memory” for timeless facts, such as geographical facts and the relationships between words. There is a large and lively literature on whether episodic memory is unique to humans. Naturally, a lot depends on precise definitions. It is clear that animals who hide food for later use have “episodic-like” memory. Scrub jays can recall what kind of food they hid, where, and how long ago (Clayton and Dickinson 1998; Clayton et al. 2001; Clayton et al. 2003; Griffiths et al. 1999). A chimpanzee has been

shown to remember overnight where food was hidden by an experimenter (Menzel 2005), and a gorilla has been shown to remember quite recent specific events, up to fifteen minutes afterwards (Schwartz and Evans 2001; Schwartz et al. 2004; Schwartz 2005; Schwartz et al. 2005). Nevertheless it is clear that there is a very significant difference between humans and non-humans in their capacity for episodic memory (see Hurford 2007: ch. 3). Episodic memory is a component of the language faculty in the broad sense, FLB. Without a permanent way of mentally storing a record of who did what to whom, and when and where, human language would not be what it is today. And this capacity, being of apparently recent origin in its highly developed human form, almost certainly emerged in Africa, since the chimp—human split.

Just as there are examples of recent evolution at both the phonetic and the cognitive-conceptual “ends” of language, there are also examples of very ancient aspects of the human language faculty at both ends. Here, we will give just one phonetic and one conceptual example. The syllable is a basic unit of phonological organization in all languages. Syllables have a characteristic shape, phonetically defined. The basic syllable shape, found in all languages, is CV, a single consonant followed by a single vowel. It has been persuasively argued that, both in ontogeny and in phylogeny, the syllable is more primitive than either of its components, the phonetic segments analyzed as consonant and vowel (Meier et al. 1997; MacNeilage 1998). The basic CV syllable is produced with an articulatory gesture of opening the mouth from a closed position, accompanied by voicing. The close analog of such a gesture in humans can be seen in the cries and calls of many animals. As MacNeilage (1998: 499) writes: “The species-specific organizational property of speech is a continual mouth open—close alternation, the two phases of which are subject to continual articulatory modulation.” He further suggests that “ingestion-related cyclicities of mandibular oscillation (associated with mastication (chewing) sucking and licking) took on communicative significance as lips-macks, tonguesmacks and teeth chatters - displays which are prominent in many non-human primates” (MacNeilage 1998: 499). Meier et al. (1997) refer to the “jaw wags” of infants aged between 8 and 13 months. To acknowledge the ancient origin of the syllable as a basic unit of speech is to recognize a continuous aspect of our evolution from non-human animals. This evolutionary foundation was laid down in its most basic

form hundreds of millions of years before humans emerged, and before Africa was formed.

At the other end of a language system from the phonetic syllable, we can look at the meanings expressed in linguistic utterances. The most common simple clause shape in languages involves a predicating expression, typically a verb, and from one to three nominal expressions. Often these nominal expressions are also directly referring expressions, picking out some particular entity in the world. Examples in English sentences are *Mary frightened John* and *Mary put the book on the table*. Such sentences describe “minimal subscenes” (Itti and Arbib 2006). Many non-human animals are clearly capable of observing an event or situation in the world, involving several participants, and analyzing it into its component entities and the relationship between them.

For example, experiments with baboons in the wild have shown that they exhibit surprise when they hear a recording of a dominant baboon making a submissive noise while a subordinate baboon makes a threatening noise (Cheney and Seyfarth 1999; Bergman et al. 2003). Baboons know the dominance hierarchy of their troop, and they can recognize each other’s voices. The surprise reaction shows that the interaction played back to the baboons is analyzed by them into the components of the two actors, the threatener and the submitter, against the background knowledge of the normal dominance relation between them.

The major difference between humans and non-humans is that we have evolved highly elaborate codes (languages) for telling each other in detail about the events that we observe (and now, of course, about much else). Baboons do not have any shared system for publicly reporting to each other who surprisingly threatened whom, and who surprisingly submitted. They keep their analysis of the event to themselves. And, given their lack of significant episodic memory, as discussed above, they probably don’t keep the perceived and analyzed event in memory for long. But the evidence shows that they do mentally perform such an analysis, into the entities involved and the relationship between them. That is, the basic propositional structure is present in the thought of the baboons, though they don’t express their thoughts in structured sentences. This theme is developed in much greater detail by Hurford (2007); see also Hurford (2003b), where it is argued that neural correlates of basic logical predicate—argument structure exist in many non-human animals, certainly primates, but also other vertebrates. This mental organization of

perceived events and situations is the private substrate upon which human public systems of communication evolved their grammatical subject—predicate structure.

The mental organization of perceived events and situations is a fundamental aspect of the organization of language, and it evolved long before the emergence of humans, and very probably before the emergence of the continent of Africa.

9.5 Conclusion

The three areas that we have surveyed here tend to suffer in the popular imagination from the same type of creation myth, suggesting a single source and a single moment of origin. It is important to stress the multi-stranded nature of languages, genomes, and phenotypic traits. The strands, throughout history, have diverged and recombined in multifarious ways, and new strands are constantly coming onstream through innovation.