3

Polygenesis of Linguistic Strategies: A Scenario for the Emergence of Languages

Christophe Coupé and Jean-Marie Hombert
CNRS & Université Lyon 2

Abstract

On the one hand, numerous hypotheses have been put forward to account for the emergence of language during the last million years of human evolution. On the other hand, a large majority of linguists considers that nothing can be said about past languages before 8,000 or 10,000 years in the past, given our current knowledge on modern languages. A large gap obviously separates such approaches and conceptions, and has to be crossed to provide a better account of the development of our communicative system.

To partially bridge the gap between the former domains, we aim at proposing a plausible scenario for the emergence of languages, with an emphasis on the development of linguistic diversity. The present study will address the question of the monogenesis or polygenesis of modern languages, which is often implicitly biased toward the first hypothesis. Probabilistic and computational models, as well as palaeo-demographic data and evolutionary considerations, will constitute the key points of our proposals.

1. The Origin of Languages

1.1 Language capacity and languages

Which definition for language?

Even if "language" is often considered as the capacity that separates Man from other species, a precise definition remains controversial. Most of the supposed distinctive features that had for example been proposed by Hockett (1960) in the 1960s have been put into question during the last decades by studies on "talking apes", like the bonobo named Kanzi (Savage Rumbaugh et al., 1998), or other animals.

What, among other characteristics, contributes to the specificity of language lies in the profound unity of its nature in terms of cognitive or informational features, and at the same time the extreme diversity of its superficial forms, namely the languages. Explaining this discontinuity has been one of the major tasks of linguists, leading to the development of elaborate and highly detailed constructions like the generative grammars, and various fields such as typology, the study of language universals, etc. To what extent some features of language are genetically encoded remains at the heart of intense debates (Schoenemann, 1999; Enard et al., 2002; Lai et al., 2001).

While the following study does not deal primarily with this controversy, we rely on a classical, yet sometimes implicit, distinction suggested by the following arguments: first, there is a *faculty of language*, language-specific or derived from more general cognitive abilities, which characterizes the human aptitude for its sophisticated communication. Second, there are instantiations of this faculty, which are the approximately 6,000 *languages* spoken today. As a consequence, we will tend to use the former expressions in italic rather than the term of "language" in isolation, and will focus on the similarities or divergences between the former notions, which are both related to the notion of *linguistic diversity*.

Emergence of the faculty of language and of modern languages

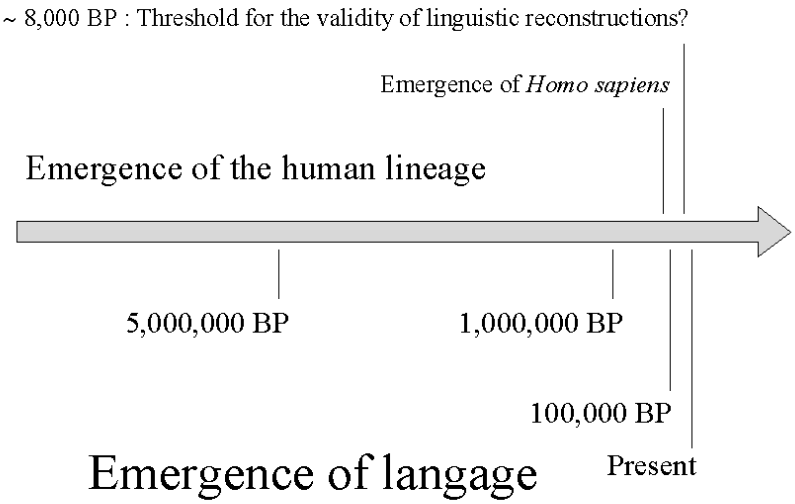
For the last twenty years, the question of the "origins of language" has been revitalized by the cooperative efforts of a number of disciplines. Partly in reaction to the proposal of a genetically-encoded *Universal Grammar* (Chomsky, 1975), their results and paradigms have provided new insights into this topic, among others:

- Progress in the appreciation of our predecessors' cognitive capacities and behaviors (archaeology or palaeo-anthropology)
- Discovery of substantial correlations between modern linguistic and genetic distributions (Cavalli-Sforza, 1994)
- Discovery of plausible neural bases for behaviors that could be related to the faculty of language, for example mirror neurons (Rizzolatti et al., 1996; Rizzolati and Arbib, 1996), etc.

On the basis of these new data, a number of theories have been put forward to explain when, why and how our communication system developed to reach its current state. Following the distinction made earlier, an increasing set of works is dealing on the one side with the origin of the *faculty of language*, while on the other side, another body of research is focusing on the origin of contemporary languages.

Nevertheless, the temporal gap which separates these two fields is a large one, and the methods used to gather and analyze the data in each of them often have few in common. As Figure 1 summarizes, the emergence of a human-specific capacity of language presumably happened along our phylogenetic tree somewhere between some tens of thousands of years and a few million years ago, while the limit for the validity of reconstructions by historical linguists is most often assessed to be around 8,000 BP (Before Present). A consensus now seems to have gained ground which dates the origin of modern language between 50,000 and 100,000 years ago, in line with the modern behaviors of our species *Homo sapiens*. All these dates differ by several orders of magnitude.

Figure 1
The time scale of language evolution



While a number of linguists either try to rebuild the story of recent human languages or the first steps of the development of the faculty of language, e.g. Bickerton (1990)'s protolanguage, few of them have taken an interest in the origin and development of linguistic diversity *per se*. This approach differs from both the work on ancestors of modern linguistic families (e.g. Proto-Indo-European, Nostratic, Austric, etc.) and the study of the origin of the faculty of language. Its intrinsic difficulty lies obviously in the absence of clues from the past (even less than for the faculty of language), but comparisons with contemporary human societies, models or relationships between languages and other cultural features may provide a valuable help.

Considering the evolution of linguistic diversity in itself is not only interesting because it represents a key event to understand the duality between the superficial diversity and the deep uniqueness of the human communication system. It is also useful because since its ties with various variables, for example the size of populations, or with cultural development, can shed light on the history of recent languages. By knowing if the social or demographic conditions of our predecessors were likely to correlate with a large or small number of languages, we may address the plausibility of scenarios put forward by some linguists about the number of languages that would have been spoken 10,000 or 50,000 years ago (Pagel, 2000).

For the sake of simplicity, we will often confound *linguistic diversity* and *diversity of languages*, even if one should keep in mind that several concepts or problems differentiate these two notions, the most significant actually requesting a precise definition of what a language is (Nettle, 1999a: 63).

The origins of linguistic diversity as a matter of *function* or *product* (why?) can be partially differentiated from the origins in *space and time* (where and when?). Since we will mostly consider the second notion in this article, we will just briefly summarize below our position regarding the first issue.

Several theories explain the emergence of the faculty of language (the "why?" question) by social causes. Dunbar (1996) proposed for example that this faculty emerged to replace grooming in its function of preserving social coherency, because of the increase in social group size. This increase was suggested by the correlation between the size of the group and the volume of the neo-cortex in various monkey species (Dunbar, 1993). Another example is Dessalles' political theory of language, which focuses on social aspects in agreement with Darwinian evolution: the development of the faculty of language would have enabled individuals to better express their qualities in order to form coalitions (Dessalles, 2001). Such coalitions are often observed in chimpanzees (De Waal, 1998).

According to a different perspective, but still centered on the social context of linguistic usage, the data of sociolinguistics have largely underlined the weight of social and inter-individual interactions on the evolution of languages and dialects. The social game is in particular partly accountable for linguistic diversity, as highlighted by various examples like Labov's pioneering study of Martha's Vineyard (Labov, 1963).

At the crossing of these various hypotheses, we assume that it is reasonable to consider the origin of linguistic diversity and variability, and therefore of languages, in a concomitant way with the origin of the function of language. If the development of our communication system was since the beginning interlaced with the social life of the communities, and if this link has been preserved until today, then it seems likely that the social game on linguistic forms and the resulting diversity of languages were preserved throughout prehistory. This phenomenon was presumably modulated by a large number of parameters: size of the communities, expressiveness of early forms of languages, development of underlying cognitive capacities, etc. Moreover, the geographic distribution of populations presumably contributed to a very ancient diversification of the communication systems, in a similar way with the evolution of species.

In this conceptual framework, the development of modern languages has to be integrated into the more general evolution of the human communication system, and only represents a "last step". It therefore becomes interesting to wonder what defines the modernity of contemporary languages, in other words what differentiates them from more archaic languages; one may refer here for example to Bickerton's notion of proto-language (Bickerton, 1990) or to Coupé and Hombert (2002)'s proposals regarding "language" in the context of the first sea-crossings to Australia.

Linguistic components

The diversity of the world's languages is naturally expressed by the differences between the structures and elements that compose these languages. Linguistic typology aims at classifying this variety of forms, which appear more or less frequently and in more or less independent ways; linguists for example often use the terms of (implicational) universals or tendencies. An obvious method to study the evolution of linguistic diversity is therefore to rely on a partial individuation of the linguistic forms, rather than studying languages as monolithic entities. Following this line, our work relies on the notion of linguistic item, as defined by Nettle (1999a: 5):

> A linguistic item is any piece of structure that can be independently learned and therefore transmitted from one speaker to another, or from one language to another. Words are the most obvious linguistic items, but sounds and phonological processes are items too, as are grammatical patterns and constructions. . . The distributions of different items in the world's languages need not be statistically independent, and indeed very often are not.

It seems relevant for us to consider linguistic items as communicative tools. We will use the term *linguistic strategies* to reflect the fact that linguistic items before all address functional needs at a cognitive level: the typological elements represent different possible solutions or *strategies* to assemble and bridge external linguistic projections of mental representations:

- word-order or case-markers to express the thematic relationships between the syntagms of the sentence;
- phonemes to encode the acoustic forms of words and overcome the large variability of phonetic forms;
- words as conventions about meanings, etc.

While the term *linguistic strategies* has been defined and used, for example by Croft (1990: 27), in a typological context, the notion of strategy refers for us to the multiplicity of possible functions for the projection from the private cognitive level to the linguistic level, and the competition that may exist among them.

In this line of thought, we may also insist on the fact that the emergence of the cognitive functions themselves has to be considered together with the corresponding linguistic strategies, and that the "grain" of these evolutions may not be as coarse as a single division between a syntactic and a non-syntactic state. Focusing on the precise linguistic correlates of phenomena like an increase in the size of the working memory (Baddeley, 1986), the emergence of a theory of mind or the development of inferential reasoning (Sperber, 1995) seems highly relevant.

From these few comments, it should be clear that we highly favor a progressive and segmented emergence of the function of language and languages, rather than a single step from a simpler stage (iconic, non-syntactic, etc.) to a fully modern one. As we have mentioned already, this article will mostly be concerned with the spatial and temporal characteristics of the various steps that have led to the current linguistic situation, and we will propose plausible hypotheses regarding these aspects in the final discussion.

1.2 Monogenesis or polygenesis of language

Monogenesis versus polygenesis of an innovation

Two scenarios can be put forward to describe the appearance of any innovation in a population. The first one is called *monogenesis* and corresponds to a single emergence of the innovation, possibly followed by its spread in the population. Several independent appearances define the second possible scenario of emergence, namely *polygenesis*. In this case, the different emergences take place at several distinct sites, provided that innovations only appear once at a site.

At least two major cultural innovations of our species seem to have appeared by polygenesis. First, agriculture is believed to have emerged independently in different places around 10,000 years ago, with archaeological proofs offering similar dates but separated by thousands of kilometers. Various regions, such as Mexico, New-Guinea, Europe, the Near-East and China display such evidence. More recently, the development of two or perhaps three different writing systems also points to a polygenesis: the Chinese ideographic system around 3.500 years ago, with inscriptions on bones or turtle shells, and the cuneiform system of the Mesopotamians around 5,000 years ago (Wang, 1973: 50-52) seem too distinct and far apart to have originated from a common origin[1].

---

[1] The possible links between the Egyptian and Mesopotamian systems are hard to trace.

Such questions of course become more difficult to answer for more remote innovations, like for example the domestication of fire around 500,000 years ago.

In the linguistic framework of our topic, it seems natural to wonder about a possible polygenesis of "language", especially if one considers it as a cultural product like agriculture or writing. To be more precise, it is possible to focus on either the emergence of the function of language or the different linguistic components. While we will not further consider the former question and its archaeological bases, we will try to defend the latter approach as the most relevant to study the development of linguistic diversity and languages.

To follow this guideline, we shall begin by introducing some general ideas about the origin of current languages.

The monogenesis of languages

It is often more or less explicitly admitted that all modern languages originate from a single original language; this hypothesis is often described by the term of *monogenesis of languages*. This proposal interestingly gathers researchers who oppose themselves on the possibility to reconstruct languages spoken before 8,000 or 10,000 years ago. Some proponents of strong limitations of the methodology of historical linguistics do not reject the plausibility of a single ancestral language, but think that its content is beyond reach of our current investigations.

This statement is much more a hypothesis than something strongly demonstrated and validated. To this extent, the following points should be recalled:

First of all, the principle of reconstruction itself introduces an implicit bias toward a single ancestor for all contemporary languages. The forms reconstructed from the current states are, above all, "terms of abstract comparisons", as translated from (Auroux, 2001), and therefore do not necessarily represent the linguistic reality of the past. It is for example rather difficult and unreasonable to conclude that Proto-Indo-European was the only language spoken in Europe or Western Asia around 6,000 years ago.

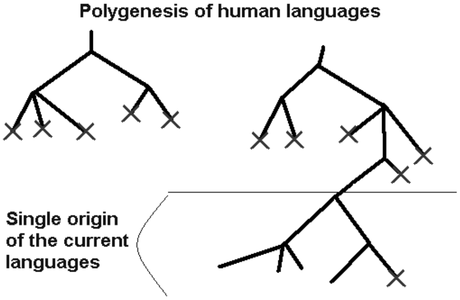
Such a position would rule out the common dialectal variety of languages, as well as the possible existence of now extinct languages. If the linguistic context suggested by the reconstructions is not the one which really took place some thousand years ago, it becomes dangerous to rely on a recurrent application of the process of linguistic reconstruction to conclude a decrease in the number of languages spoken by our ancestors as we go further back in the past. A smaller size of the meta-population in the past might have resulted in many fewer languages, but this argument is not sufficient to conclude the existence of a unique tree for all contemporary languages.

A crucial event that plausibly separates the more recent Neolithic period from the Paleolithic situation is the demographic explosion which took place with the development of agriculture, after a slow initial growth at the end of the Paleolithic. The notion of punctuated equilibria, borrowed from the evolutionary theories in biology and introduced by Dixon (1997) in linguistics, can be applied to such transitions (ibid: 77). Such an approach points to the hidden ties between linguistic diversity and demographic contexts. Some researchers have begun to explore their relationship, either with computational models studying the role of population size (Nettle, 1999b), or considerations about the correlations between densities of speakers and linguistic diversity (Jacquesson, 2001).

A second point is that assimilating a monogenesis of today's languages and the existence of a unique ancestor can be viewed as incorrect. As Figure 2 depicts it, several distinct families could have indeed appeared independently and evolved before all languages but one disappeared. The remaining specimen would have given birth to all modern languages. This scenario is not implausible, since languages often disappear, for example in case of contacts with unbalanced social relationships between populations (Nettle and Romaine, 2000:147-9). Unbalances may have been even stronger in the past due to the small size of populations (Marsico et al., 2000).

However, given the large areas which have been inhabited by humans for tens of thousands of years, it seems unlikely that the descendants of several "initial languages" could have all disappeared

Figure 2
Polygenesis of languages and single origin of contemporary languages



but one. Large areas like Europe have seen the expansion of families like Indo-European, but such developments and replacements did not reach a world scale in recent times[2].

A third argument relies on the link between the origin of contemporary languages and the origin of our species between 100,000 and 200,000 years ago. Two main hypotheses are still discussed to account for the origin of modern Man. The first one, the *Out of Africa* hypothesis, postulates that the speciation event which led to our species took place in East Africa, and that our ancestors subsequently migrated out of this region and colonized the entire Earth (Lahr and Foley, 1994), replacing all the previous species that were living in Africa, Europe or Asia. Based on congruent archaeological and population genetic data (Stringer and Andrews, 1988; Cann et al., 1987), this scenario gathers the favors of most scholars in its opposition to the *Multiregional Continuity* hypothesis. This second proposal, based on Asian fossils and other

---

[2] One may argue that if *Homo sapiens* replaced all other *Homo* species, a single language could as likely have replaced all other languages. However, where in the first case the replacement may be due to a physiological or cognitive advantage, linguists agree that no language is functionally better than others, which may be partially extrapolated back in the past for modern humans.

genetic studies, argues that modern humans evolved from pre-*sapiens* species locally, and that genetic fluxes were dense enough to preserve a single species despite the large geographical areas implied (Thorne and Wolpoff, 1992).

In the framework of the *Out of Africa* theory, an emergence of modern languages, and some may even say of the true function of language, due to neuro-physiological changes as part of the speciation event, would have taken place in the small and geographically restricted population of "new-born" *sapiens*. A monogenesis would then have been more likely, followed by the spread of languages during the migrations leaving East Africa. On the other hand, the link between new cognitive abilities and new forms of language would rather lead to a polygenesis in the case of the *Multiregional Continuity*.

"Rethinking" the origins of modern languages

The aim of our work is not to reject the hypothesis of the monogenesis of modern languages as a whole, but to point at the complexity of this question, and to show how its nuances may partially empty the current consensus from its substance. Our goal is to propose a general sketch of the emergence of linguistic diversity, which could in particular be applied to the development of modern languages.

Our initial statement is that it seems often implicitly assumed that the putative ancestor of modern languages shares with them their degree of complexity and structures. In opposition to proponents of a catastrophic emergence of the function of language, many scholars do not deny the evolution of our system of communication (e.g. the notion of protolanguage), but do not apply this evolutionary way of thinking to the history of *modern* languages. This is of course partly due to the fact that people working on the origin of "language" often do not tackle the prehistory of modern languages to this end, and vice-versa.

We aim at bridging this gap by considering that a large number of differences may have existed between modern languages and their first ancestors in modern Man. We will defend a scenario where part of the current linguistic diversity would have emerged during the several tens of thousands of years that have followed the emergence of our species. The second and third parts of this article will be devoted to this goal, introducing probabilistic and computational models and experiments abstracting the appearance of cultural innovations in group-fragmented populations, as well as considerations about palaeo-demographic and evolutionary data. But before beginning to describe our arguments, these introductory paragraphs shall be concluded with a brief discussion of some of the palaeo-demographic data on which we will rely hereafter.

1.3 Taking palaeodemography into account

Emergence in a population of individuals

The appearance of any new cultural feature in the human species must be considered in a realistic framework, partly built on the specificities of the human population at the period involved. More precisely, if we want to talk about "sites of emergence", as in the definition of polygenesis mentioned above, a clear definition of the meaning of these words in a demographic context is necessary. We shall therefore succinctly describe the data and theories about the societal structures during prehistory.

The structure of the ancestral human population

Several sources of data point at various cues regarding the structure and the size of past populations.

First of all, palaeo-anthropology and archaeology, by the study of characteristics of prehistoric living places such as surface, organization etc., lead us to conclude that the human population was composed of small groups of some tens of individuals, mostly between 20 and 50, during most parts of the Paleolithic (Hassan, 1981: 93-94). The number of 25 individuals is regularly quoted in various studies on prehistoric populations, and appears to be independent of time, density or type of environment (ibid: 53).

Comparison to recent populations of hunter-gatherers inhabiting various ecosystems, e.g. Eskimos with either a caribou and sea mammal hunting economy or a caribou hunting and fishing economy, also gives clues about densities during prehistory. Table 1 is reproduced from (Hassan, 1981: 198) and summarizes Birdsell ‘s proposals in 1972 for three successive periods of the Paleolithic.

Table 1
Estimates of world prehistoric population,
reproduced from (Hassan, 1981:198)

Period	Pop. Density (persons/km²)	Area occupied (1e6 km)	World pop. (1e6 persons)
Lower Palaeolithic	0.015	27.0	ca. 0.4
Middle Palaeolithic	0.032	38.3	ca. 1.0
Upper Palaeolithic	0.039	57.5	ca. 2.2

The total surface that was inhabited by the meta-population of humans can be estimated by the repartition of the prehistoric living places, as well as the carrying capacities of various environments: forests, cold or warm deserts, etc. (Bocquet-Appel and Demars, 2000). Australia or the Americas were, for example, only very lately colonized by our own species (and not by former *Homo* or pre-*Homo* species).

Beyond these first data, population genetics studies propose some evaluations of the global population during the last million years. All the studies conclude a very small population of one or two million people at most, and the existence of a genetic “bottleneck” around 1,800,000 years ago with the speciation leading to *Homo ergaster*. However, their analysis of different genetic markers (mtDNA (Sherry et al., 1994), Alu insertions[3] (Sherry et al., 1997), micro satellites (Zhivotovsky et al., 2000), etc.) feed the debate

---

[3] Alu insertions are primate-specific genetic elements that mobilize via the process of retroposition. They are believed to be non-coding, and account for around 5% of the human genome by mass (which represents 500,000 Alu sequences) (Sherry et al., 1997).

about a possible second bottleneck 100,000 years ago and later expansions, with the appearance of our species *Homo sapiens* (Hawks et al., 2000). These disagreements exemplify the genetic side of the controversy between the *Out of Africa* hypothesis and multiregional continuity.

From these data, it appears clear that the study of the monogenesis or polygenesis of cultural innovations, whether linguistic or not, can be based on the human group as a relevant "base unit", at least until the appearance of larger communities with the development of agriculture around 9,000 years ago.

Table 2
Areas of various lands or continents

	Area in km²
entire Earth	510,072,200
emerged lands	148,939,800
Asia	44,547,800
Africa	30,043,900
Europe	10,404,000

Table 3
Density of human groups for different areas and population sizes

Macro-population size	10,000	25,000	125,000	250,000	1,250,000	2,500,000	5,000,000
Surface (km²) / nb. of groups	400	1,000	5,000	10,000	50,000	100,000	200,000
1,000,000	4e-4	1e-3	5e-3	0.01	0.05	0.1	0.2
5,000,000	8e-5	2e-4	1e-3	2e-3	0.01	0.02	0.04
10,000,000	4e-5	1e-4	5e-4	1e-3	5e-3	0.01	0.02
25,000,000	1.6e-5	4e-5	2e-4	4e-4	2e-3	4e-3	8e-3
50,000,000	8e-6	2e-5	1e-4	2e-4	1e-3	2e-3	4e-3
100,000,000	4e-6	1e-5	5e-5	1e-4	5e-4	1e-3	2e-3

To get a better idea of the quantitative values that are involved in this framework, Tables 2 and 3 report the current sizes of land and the continents, and computations of the densities of human groups for different areas and sizes of the macro-population.

In the next experiments and hypotheses, we will mostly investigate densities of population varying between 0.001 and 0.0001 human groups per km^2. They correspond to average values from 0.025 to 0.0025 individuals per km^2. While they will appear small compared to the values of Table 1 (especially the lower bound), we will explain later why we believe such values to be relevant in some situations.

2. Mathematical Models and Computer Simulations

2.1. A mathematical model to evaluate the probabilities of monogenesis or polygenesis of language

Quantifying the probabilities of monogenesis or polygenesis

Except in a few cases for which deciding between monogenesis or polygenesis of an innovation is possible without ambiguity (as we have seen for the development of agriculture or writing systems), it becomes harder to estimate the “mode” of emergence of an innovation when one goes further in the past and clues become rarer. The main difficulty, as we will see in the next paragraphs, lies in the possibility of an undetected diffusion of the innovation, which might lead to the wrong conclusion of a polygenesis.

If concretely distinguishing monogenesis or polygenesis is a difficult task in some cases, it is however possible to compute the *probabilities* of these events. Indeed, if a mathematical model allows one to conclude that polygenesis of an innovation is much more likely to have taken place than monogenesis, it becomes relevant for the theories on this matter to consider the two possibilities, and not to reject the polygenetic hypothesis without a strong argument. This conception is even more crucial as definitive proofs are lacking when one is interested in the origins of the faculty of language or of languages.

Description of Freedman and Wang's model

Freedman and Wang (1994) have been interested in the possibility of studying the probabilities of the two scenarios of monogenesis or polygenesis. To this end, they have proposed a purely mathematical model that we are now going to describe briefly. One of the targets of the paper was to reformulate correctly the "folk" intuition which assumes that if a rare event is already unlikely to happen once, it will be even less likely to occur twice.

The model focuses on the link which unites the probability **p** of the emergence of language[4] at one site and the probabilities of no emergence, monogenesis or polygenesis at **n** independent sites over a fixed period of time. The mathematical approach enabling the calculation of the probabilities depends on Poisson's probability distribution, which characterizes the occurrences of rare events, that is those with a weak probability of occurrence. The values studied for the probability of emergence at one site **p** are chosen such that the expected number of sites at which language emerges is 1, 2, 3, etc. The probability of emergence is integrated over the entire time period, and the expected number of sites is therefore equal to $\mathbf{p} \times \mathbf{n}$.

It should be clear here that this expected number reflects a statistical approach, i.e. the mean number of sites at which language emerges if one considers a large number of episodes. By episode, we mean a concrete instantiation of the model described above, at the end of which one of the three possible scenarios has occurred: no emergence, monogenesis or polygenesis. The model does not predict the outcome of a single episode, but rather, of a large number of episodes, the percentages of them which end up as no emergence, monogenesis or polygenesis. The issue is then to investigate how the probabilities of the three scenarios evolve with the product $\mathbf{p} \times \mathbf{n}$.

To ease the understanding of the simulations we develop in the

---

[4] The authors were interested in the emergence of language 1 or 2 million years ago. This rather relates to the notion of function of language, but the model may in fact be applied to any cultural innovation.

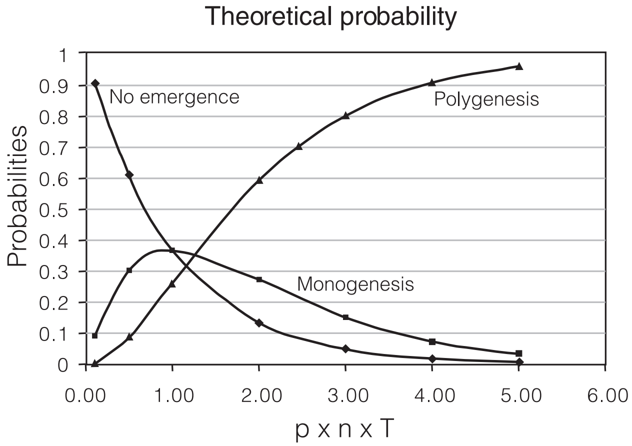
remaining parts of this chapter, we find it useful here to slightly modify Freedman and Wang's model, in a way that does not change any result or interpretation: we modify the meaning of the probability **p**, which was integrated over a whole time period in the original model, to introduce the notion of *time step*. Instead of considering an indivisible time period, we consider a number **T** of time steps, and **p** the probability of emergence at each time step. This new approach can therefore be seen as the discrete counterpart of Freedman and Wang's model, since the global time period they considered has been cut into discrete time units. During an episode, at each time step, a random test is performed against the probability **p** to check whether the innovation emerges or not, *provided that it has not emerged before*. Within this new framework, it turns out that the relevant parameter, i.e. the expected number of sites, can be replaced by the mean number of times the test against the probability **p** is positive in the **n** groups during the **T** time steps. This new parameter is equal to the product $\boldsymbol{\lambda} = \mathbf{p} \times \mathbf{n} \times \mathbf{T}$.

One should notice that since only the value of the product $\boldsymbol{\lambda}$ is relevant, the time variable and the probability of emergence at one site are not independent; this relationship allows us to focus on the relevant values of **p** given **T**: if the values of **p** are chosen such that the product $\boldsymbol{\lambda}$ is very weak (well below 1.0), then the probabilities of monogenesis or polygenesis will be insignificant. For values of $\boldsymbol{\lambda}$ well above 1.0 (for example larger than 10.0), the probability of polygenesis will be close to 1.0, and the two other probabilities insignificant. In these two extreme situations, no qualitative transitions would be observed in the probabilities of the "modes" of emergence.

Results

The graph of Figure 3 provides a good understanding of the situation. For various values of $\boldsymbol{\lambda} = \mathbf{p} \times \mathbf{n} \times \mathbf{T}$, the probabilities that zero, one (monogenesis) or several emergences (polygenesis) take place are displayed.

Figure 3
Evolution of the probabilities of monogenesis and polygenesis
of an innovation at several sites according to the probability of emergence
at one site; adapted from Freedman and Wang (1996)



The evolution of the curves representing the probabilities of no emergence, monogenesis or polygenesis can be summarized in the following way:

- The probability $\mathbf{p_n}$ of no emergence decreases to 0 as λ increases according to the relation $\mathbf{p_n} = \exp(-\lambda)$;
- The probability of monogenesis $\mathbf{p_m}$ first increases for small values of the product, and then decreases to 0 (bell-shaped curve). $\mathbf{p_m} = \lambda.\exp(-\lambda)$;
- The probability of polygenesis $\mathbf{p_p}$ increases as the product increases, with the relation $\mathbf{p_p} = 1 - (\lambda + 1).\exp(-\lambda)$.

The last two behaviors lead to the existence of a threshold for the probability of emergence at one site (depending on the number of sites **n** and **T**), over which polygenesis becomes more likely than monogenesis. In other terms, large values of **p**, **n** or **T** increase the likelihood of a polygenesis of the innovation, and make it more likely than monogenesis over a given threshold.

As a consequence, it is necessary to reformulate more precisely the intuition according to which a weak probability of emergence at one site makes the emergence at several sites even less frequent, since no threshold is set for such statement and the intuition does not take into account the combinatory function of the probabilities at several sites.

It should be pointed out here that it is nearly impossible to estimate the probability of emergence at one site, which widens the gap between the model and real situations. Nevertheless, the knowledge *a posteriori* of the situation may contribute toward a better estimation of this probability with the help of conditional probabilities: as the authors suggest it, for small values of **p**, the model does not account for the fact that the innovation has emerged. This is especially meaningful in the case of linguistic components or the faculty of language, because all human populations possess languages with some recurrent linguistic items, a situation to be contrasted with those of agriculture or writing. This may falsely lead to the rejection of monogenesis; however, it remains impossible to conclude polygenesis or monogenesis for a single episode (see definition above), since for a sample of size 1, an unlikely event like monogenesis cannot be ruled out. The fact that for small values of **p** the model does not account for the emergence of the innovation does not interfere with this issue, but is rather a limitation of the model itself. As a conclusion, polygenesis appears more likely than monogenesis for a large range of probabilities of emergence at one site. It also becomes even more likely as the number of sites increases for a fixed value of **p**.

In this model, sites remain abstract entities and are, in particular, totally independent from each other. But, as we have seen, the reality of prehistory is partly represented by human groups moving in large geographical areas, and as a consequence able to enter into contact and transmit cultural innovations. The frequency of such contacts is hard to estimate intuitively. Freedman and Wang's model therefore calls for further enhancements to take this aspect into account.

2.2 Measuring the frequency of contacts

Relevant parameters

Estimating the frequency of contact between human groups knowing their density in a geographical area is a rather unintuitive task which asks for precise calculations. Were they meeting every month, every six months, every ten years or even less often? Of course, just the density of groups, represented hereafter by the variable **de**, is not enough, and other parameters have to be considered. We shall restrict our attention to three of them:

- The threshold distance **d** between two groups for a contact to occur (when does one group detect another one?);
- The "geometrical" features of the groups' movements;
- The speed of the groups **v**.

The first parameter may be linked to the surface of the *catchment territory* of a group, which may be roughly equated to the region where a second group could be detected, not taking such clues as distant sounds or smoke into account. The size of such territories for hunting or gathering food is estimated to be around several hundred square kilometers, as indicated by ethnographic studies on various groups of hunter-gatherers (Biraben, 1997: 46–7). Regarding the geometrical "features" of the movements, it is reasonable to assume more or less directional migrations, from fully random "brownian-like" trajectories to extremely rectilinear motions, although the latter seem less likely. The speed of displacement may be approximated to several kilometers per year. Dates of archaeological sites related to the migration of the first farmers in Europe 10,000 years ago suggest a speed of 1 kilometer per year for the population wave (Cavalli-Sforza, 1994: 108–9). However, this net distance would only equal the average speed of human groups in the case of unidirectional movement in the same direction as the wave of migration (Hassan, 1981: 200–1). For more random movements, the average speed is necessarily higher. Moreover, tribes of hunter-gatherers were presumably moving faster than farmers cultivating the ground, given their reduced sedentarization.

First theoretical approach

In the case of some specific movements, a useful analogy allows estimating the frequency of contact between groups. Although we will not enter here into the details of the mathematical method (Coupé, 2003), we may just summarize the approach by pointing at the similarities between the collisions of molecules in a perfect gas and the contacts between human groups: the diameter of the circular catchment territory of a human group can indeed be assimilated to the diameter of a molecule. We rely on this similarity to estimate the frequency of contacts in the case of pseudo-rectilinear movements (directional changes only take place after a collision of molecules in a perfect gas). The frequency of contact **f** for the molecules is approximated by the following formula:

$$\mathbf{f} = \mathbf{de} \times \mathbf{d} \times \mathbf{v}$$

The numerical application with a density equal to 0.001 groups per km^2, a speed of displacement of 4 km/year and a radius of 10 km (d = 20 km) for the catchment territory leads to a frequency of around **one contact every 12.5 years**. The frequency of contact for human groups seems therefore at first extremely low. This may be moderated by differences in regional densities, but one decade seems a reasonable magnitude for the period between two contacts. To assert whether such values are reasonable, we will further ground our approach with computer simulations.

Simulations for rectilinear motions

We programmed a simple computational multi-agent model to measure the frequency of contact experimentally. Agents representing human groups were allowed to move in a square bi-dimensional space (with pseudo-rebounds on the frontiers of this space and an initial random distribution) during a large period of time; the number of contacts between them was averaged over this period. The parameters varied as follows:

- Number of agents **n**: 50, 100, 200 and 400. Additional values were also considered in specific cases (see Figure 5);
- Density of groups **de**: 0.001, 0.0005 and 0.0001 groups per km^2. The space area was computed given **de** and **n;**
- Speed **v**: 1, 5, 10, 15 or 20 km per year;
- Threshold distance for contact **d**: 5, 10, 15 and 20 km.

To estimate how close to the theoretical case the experimental results were, we computed the ratio of the experimental frequency of contact **f** over the theoretical frequency $\mathbf{f_{th}}$ for various values of the parameters. Figures 4 displays curves for various values of the speed and size of the catchment territory, while Figures 5 and 6 allow estimating the impact of the density and number of groups.

In addition to these graphs, various correlations were computed between series of values of **f** for different values of the parameters, in order to discover mathematical relationships. This approach was applied to all our simulation results and proved to be useful, as will appear later in this chapter.

Figure 4
Evolution of the frequency of contacts in function of the speed of the agents and the threshold distance for contact (threshold distance for contact in horizontal axis)

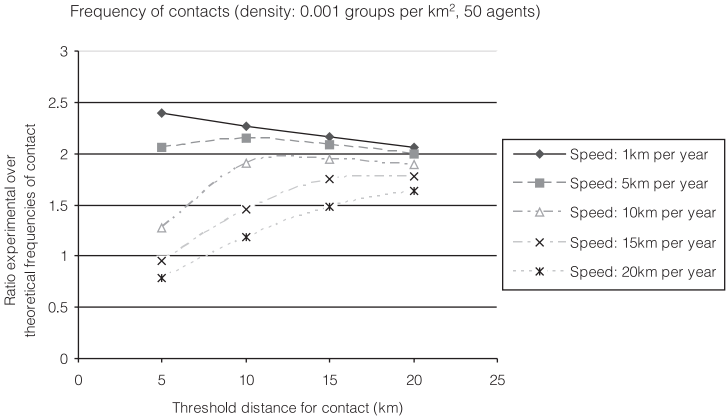


Figure 5
Evolution of the frequency of contacts
in function of the number of human groups

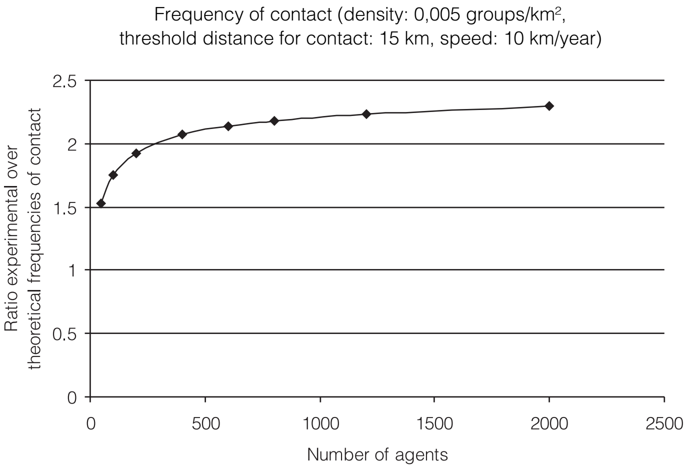
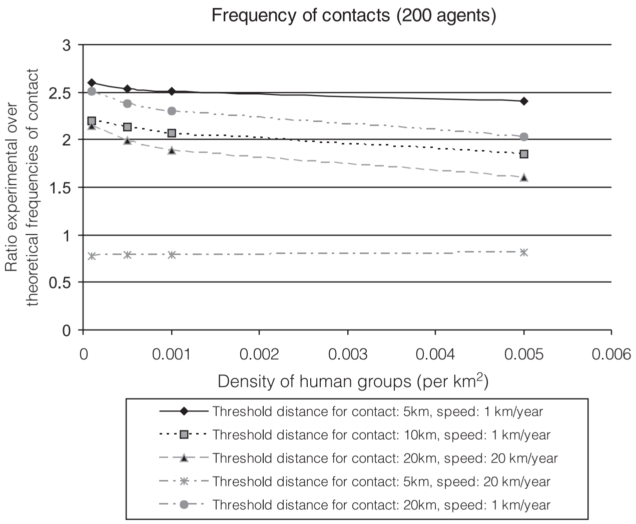


Figure 6
Evolution of the frequency of contacts
in function of the density of human groups



The following results can be drawn from the figures above, other graphs that were not included, and analyses of the series of results:

- The density of groups **d** does not interact with the other parameters. Moreover, the curves displayed in Figure 6 show that the ratio of frequencies $\mathbf{f} / \mathbf{f}_{\mathbf{th}}$ always seems to remain bounded in a narrow interval whatever the value of the density of groups. This means that the frequency of contact **f** is *linearly* related to the density of groups;
- The number of groups **n** does not interact with the other parameters, and the curve in Figure 5 shows that the ratio of frequency seems to reach an asymptote for large values of **n** Since relevant values of **n** are much larger than 1000, **f** can be considered as independent from **n** in a first approximation;
- The frequency of contact **f** can be described by the following formula expressing the independence or interactions between the various parameters: $\mathbf{f} = \mathbf{d} \text{ x } \mathbf{v} \text{ x } \mathbf{de} \text{ x } g(\mathbf{d}, \mathbf{v})$. For the values of interest to us, **g** takes values between 0.5 and 3.

It then appears that the former conclusions should be moderated, in that the experimental frequencies of contact **f** may be different from the theoretical frequency $\mathbf{f}_{\mathbf{th}}$ with ratios from 0.5 to 3. However, even with this multiplicative factor, the frequency of contact can be said to remain very small, especially compared to the life span of an individual during prehistory.

Three reasons may be invoked to explain the discrepancies between the theoretical and experimental cases:

- The theoretical formula approximates the reality by considering a molecule moving in a space containing static elements it can hit; considering a high velocity for these elements (according to a molecular speed distribution) leads to refining the formula by adding a $\sqrt{2}$ multiplicative factor in the right expression: $\mathbf{f} = \sqrt{2} \times \mathbf{de} \times \mathbf{d} \times \mathbf{v}$. This specific numerical factor is adapted to the speed distribution of molecules, but the general principle may explain why the ratios of frequencies we observed were mostly greater than 1.0.

- In the theoretical case, molecules move continuously, where in our case, the groups jump from one location to the next. For a small speed and a large size of the catchment territory, the area covered by a group after a given time span is close to the area covered in the case of a continuous motion, since the areas covered by the group at each time step largely overlap. However, for larger speed and smaller catchment territories, the motion cannot be assimilated to a continuous one, since the regions covered at each time step only slightly overlap or do not overlap at all. Such behavior modifies the frequency of contact, since on the one hand, a larger area is covered if the catchment territories overlap less, and on the other hand, close groups may miss each other in case of distant jumps from one area to another.
- The theoretical case assumes an infinite space, whereas our agents were moving in a closed area. For agents close to the boundaries of this area, the frequency of contact is smaller than for agents in the center of the space. The larger the size of the area, the closer the experiments get to the theoretical case.

Non rectilinear motions

For more random movements, the total area covered by a group during a period of time is intuitively smaller, since this group will more frequently revisit the same places, by often changing its direction. However, it remains unclear whether this affects the frequency of contact: groups may reduce their chance of meeting distant groups through more local movements, but this in turn increases their chance of meeting close groups again and again[5]. Once again, we relied on computer simulations to evaluate the situation.

To model the notion of linearity of movement, we introduced an angle α corresponding to the maximum deviation allowed for the direction of a group at each time step; an angle equal to 0 corresponds to a linear motion, whereas an angle of $2.\pi$ is equivalent to a Brownian motion: at each time step, agents choose a deviation between $-\pi$ and π. In the former experiments with rectilinear motions, the angle was simply set to 0.

We ran experiments for the same values of parameters as above, crossing them with the following values of α: 2π, π, $\pi/3$, $\pi/6$ and 0. As explained earlier, we computed correlations between the series of values obtained for these various angles, and compared the numerical values themselves.

It appears that the frequency of contacts **f** obeys the behavior described for rectilinear motions whatever the value of α, and that this angle in fact plays no role in the frequency of contact. We interpret this result as a balance between the two phenomena that we introduced three paragraphs ago (distant versus local contacts). The frequency of contact can therefore still be computed according to the following formula: $\mathbf{f} = \mathbf{d} \text{ x } \mathbf{v} \text{ x } \mathbf{de} \text{ x } g(\mathbf{d}, \mathbf{v})$.

Time to complete a diffusion

The former results may falsely lead to the conclusion that the randomness of movements plays no role in the impact of contacts on the emergence of an innovation. However, it is not as much the frequency of contact as the speed of diffusion of this innovation which will be relevant in the coming paragraphs. We ran a last series of simulations to measure the time needed for a diffusion to reach all

---

[5] For a Brownian motion in physics, the study of the intersections of two trajectories of particles (called Wiener's sausages) is a hard and still unsolved problem.

the groups, once again according to different sets of values of the former parameters given above. For each set of parameters, the average time **T** for the complete diffusion of the innovation (to all groups, starting from a single group in possession of the innovation) was measured over 50 identical simulations.

An additional parameter was added to the model: $\mathbf{p}_t$ represented the probability that in case of contact, one agent in possession of the innovation transmits it to a second agent not in possession of it.

It appears obviously that **T** is not only dependant on **d**, **de** and **v**, but also on $\mathbf{p}_t$, α and **n**. Indeed, the local aspect of the contacts in the case of more random movements decreases the speed of diffusion, and the more agents for a given density, the bigger the space to conquer for an innovation. The first phenomenon may be interpreted as a weak coefficient of diffusion in the modeling of epidemics (Murray, 1994: 651–5).

According to the analyses of the numerical outputs of the simulations, and by comparing the values of the frequency of contact **f** and the time for complete diffusion **T** for identical sets of parameters, the following relationship links **T** to the other variables of the model:

$$\mathbf{T} = h1(\mathbf{f}, \mathbf{n}, \mathbf{p}_t) \times \exp(\alpha . h2(\mathbf{f}, \mathbf{p}_t))$$

Figures 7 and 8 illustrate the progressive increase of **T** as α increases for two sets of values of **f** and $\mathbf{p}_t$. The different shapes of the exponential tendency curves and their coefficients clearly point to the second term of the product in the former expression of T.

The functions *h1* and *h2* required detailed investigations before their analytical expressions could be found. First, the analysis of series of values of **T** for α=0 led to a further decomposition of the function *h1* as follows:

$$h1(\mathbf{f}, \mathbf{n}, \mathbf{p}_t) = \frac{\eta}{(\mathbf{f}.\mathbf{p}_t)^{\beta+1}} \mathbf{n}^{\delta.(\mathbf{f}.\mathbf{p}_t)^{-\beta}}$$

The coefficients η, β and δ were then estimated: η = 1.89, β = 0.17 and δ= 0.42.

Figure 7

Evolution of the time for complete diffusion in function of the directionality of the movements (first set of parameters)

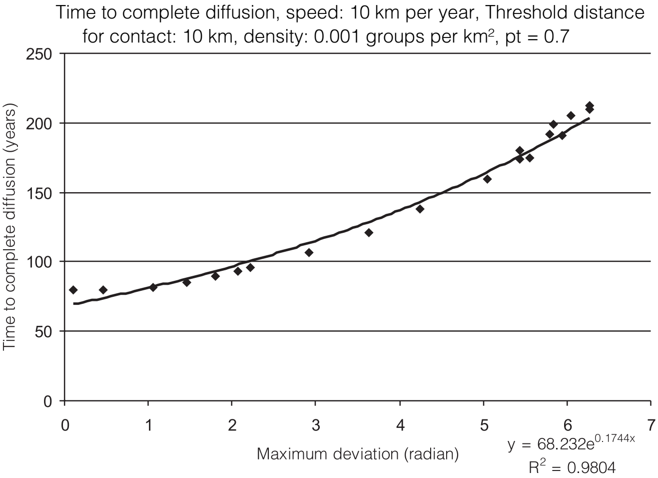
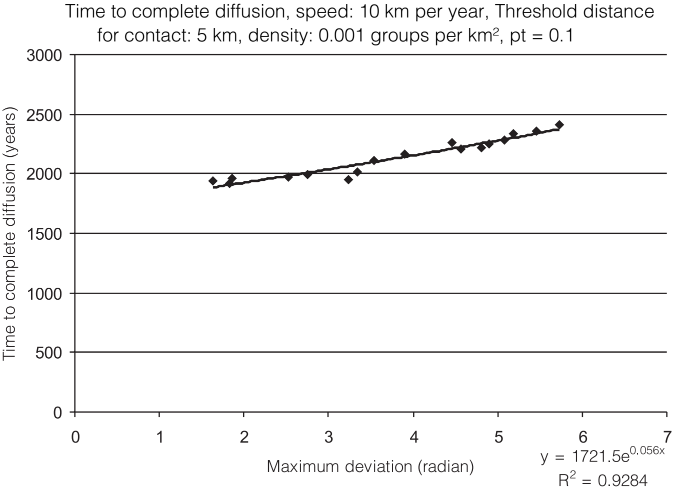


Figure 8

Evolution of the time for complete diffusion in function of the directionality of the movements (second set of parameters)



Paying attention to the coefficients of the exponential tendency curves for various values of **α** further led to the following expression for $h2(\mathbf{f},\mathbf{p}_t)$:

$$h2(\mathbf{f},\mathbf{p}_t)=\varepsilon.\mathbf{f}.\mathbf{p}_t^{\theta}$$

The coefficients ε and θ were estimated as η, β and δ previously. The following values were found: $\varepsilon = 1.21$ and $\theta = 0.69$.

The expression found for **T** is rather complex, and may take other simpler forms. It may also be simplified for large values of **n**, which were however computationally too expensive to be investigated with the necessary high number of simulation runs. Table 4 provides various values of **f** and **T** for different values of **d**, **v**, **de**, **n**, **p** and **α**. Numbers in italic are extrapolations, while other values are experimental results.

Table 4
Values of **f** and **T** for different sets of parameters of the model
(extrapolated values in italic)

v (km/year)	d (km)	de (group/km)	n	α (rad)	p	f	T (years)
10	10	0.0001	258	0	0.7	0.022	810
10	10	0.0001	94	0	0.7	0.022	695
8.81	13.69	0.0034	189	5.40	0.107	0.86	179.88
6.94	10.25	0.001	113	1.57	0.047	0.16	1359.94
10	*10*	*0.001*	*10000*	*1.5*	*0.1*	*0.24*	*1263*
10	*10*	*0.001*	*1000*	*1.5*	*0.1*	*0.24*	*756*
10	*10*	*0.001*	*10000*	*1.5*	*0.01*	*0.24*	*8953*
10	*10*	*0.001*	*1000*	*1.5*	*0.01*	*0.24*	*6329*
5	*10*	*0.0005*	*10000*	*3*	*0.01*	*0.065*	*31055*
5	*10*	*0.0005*	*1000*	*3*	*0.01*	*0.065*	*23524*

On the nature of contacts between human groups

The opportunities for contact between human groups were presumably very slight. One question that may be raised is the possibility for these encounters to have been non-violent and sources of cultural exchanges or transfers of individuals. To partially answer this question, the exchanges of lithic material (Marwick, 2002) are a first clue of non-aggressive contacts, unless we always consider fights for the acquisition of these resources. As it appears in the analyses of databases of parietal art schemes in France and Spain during the Upper Paleolithic (50,000–10,000 BP) (Sauvet and Wlodarczyk, 1995), local stylistic heterogeneities in a globally homogeneous context for the graphic representations strongly suggest the existence of inter-groups relationships. These relationships may be at the origin of both the large-scale homogenization and the preservation of local diversities contributing to the social position of a group among others.

Finally, exchanges of genes through the exchanges of individuals would have usefully preserved the diversity of the gene pool of a group. Such exchanges between groups, namely exogamy, are still common in hunter-gatherers like the Australian aboriginals. Welcoming a man or a woman speaking a different language could presumably play a significant role in the linguistic evolution of a group.

Following the previous hypotheses raises the next question: could the contacts between groups have had an impact on the monogenesis or the polygenesis of innovations? This point is now going to be further investigated

2.3 Combining independent discovery and transmission by contact

Intuitive statements

If one follows the results of sections 2.1 and 2.2, a group may either discover an innovation by itself, or receive it from another group it meets. Intuitively, the resulting probabilities of monogenesis or polygenesis will be the result of the interaction between the diffusion and the emergence of the innovation among the groups: fast diffusion may prevent polygenesis from occurring, since all the groups will be contaminated by the innovation by diffusion before having the possibility to discover it by themselves. Conversely, a slower diffusion will preserve the possibility of polygenesis and the probabilities of Freedman and Wang's model will be more likely to be relevant for the real situation. The next simulations are aimed at checking whether this intuition is relevant or not.

Description of the new model

To test the hypothesis of the last paragraph, both aspects of emergence and diffusion of an innovation among human groups were combined in a single multi-agent model. Contrary to the former models, it did not involve a spatial environment, and relied on a simple discrete probabilistic framework involving two parameters:

- The parameter **r** expressing the probability for an agent to receive the innovation from groups already in possession of it;
- The parameter $\mathbf{p_c}$, giving the probability for an agent to discover by itself the innovation at each time step.

A total number of agents **n** and a time limit for the simulation $\mathbf{T_{max}}$ were set for each run. The relevant value was as in Freedman and Wang's model the product $\lambda = \mathbf{p_c} \times \mathbf{n} \times \mathbf{T_{max}}$.

At each time step, two statistical tests were performed for each agent to decide whether it would discover the innovation by itself, or receive it from another group. In the first case, a random number *rn1* between 0 and 1 was compared to $\mathbf{p_c}$, while in the second case, a second random number *rn2* in the same interval was compared to the product **r** times the percentage *p* of agents already in possession of the innovation. An agent not in possession of the innovation was i) discovering the innovation by itself if *rn1* < $\mathbf{p_c}$, or ii) receiving the innovation from other groups if *rn2* < *p*.**r**. According to this second condition, no diffusion was taking place before a first agent had discovered the innovation. After this initial discovery, the diffusion was taking place according to a logistic growth, *similar to what had been observed in the previous model of diffusion*. This justified the use of a simpler and computationally cheaper non-spatial model. The relation between **r** and **T** was explored with additional experiments, and the following result was established:

$$\mathbf{r} = \frac{2\ln(\mathbf{n})}{\mathbf{T}}$$

This relation found experimentally can in fact be derived from the equations describing logistic growth, provided that the number of human groups **n** is large enough.

Experiments and results

To test our model, large sets of values for the two parameters **r** and **λ** were built, either by crossing random values for both of them or by fixing one parameter and choosing random values for the second. For each set of values of **λ** and **r**, 200 runs were performed to measure the probabilities of no emergence, monogenesis and polygenesis. A run was stopped i) as soon as two independent discoveries had taken place, ii) when a single discovery had diffused to the whole population or iii) if the time limit was reached without any discovery of the innovation.

An initial series of runs was first computed to check the conformity of the computer model to the theoretical results established by Freedman and Wang in the case of no diffusion. A value of **r** equal to 0 reproduced the theoretical values, modulated by a small variability due to the limited number of experiments (an infinite number of simulations should reproduce the mathematical laws exactly).

In order to evaluate the impact of diffusion on Freedman and Wang's model, we investigated the evolution of the ratios of the experimental probabilities over the theoretical probabilities for no emergence, monogenesis and polygenesis, respectively $\mathbf{r}_n = \mathbf{p}_n^{\exp} / \mathbf{p}_n^{\mathrm{theo}}$, $\mathbf{r}_m = \mathbf{p}_m^{\exp} / \mathbf{p}_m^{\mathrm{theo}}$ and $\mathbf{r}_p = \mathbf{p}_p^{\exp} / \mathbf{p}_p^{\mathrm{theo}}$. It appeared that the first ratio was always taking values very close to 1, which means that the probability of no emergence is left unaffected by diffusion. This is rather intuitive since diffusion can only occur when an emergence has already taken place. In order to evaluate the evolution of the two other ratios, we relied on the following intuitions: i) large values of $\mathbf{r}$ should decrease the ratio $\mathbf{r}_p$ until it eventually reaches 0: an extremely fast diffusion prevents polygenesis, ii) very weak values of $\mathbf{r}$ should give a ratio close to 1, since there is then virtually no diffusion, iii) the larger the probability $\mathbf{p}_c$ and $\boldsymbol{\lambda}$, the closer to 1 the ratio $\mathbf{r}_p$: groups discover the innovation by themselves very quickly, which prevents diffusion to occur, iv) before and after a phase of transition for the value of the ratio $\mathbf{r}_p$, variations in $\mathbf{r}$ or $\boldsymbol{\lambda}$ should have minimal impact. A possible relationship between them and $\mathbf{r}_p$ could be:

$$\mathbf{r}_p = \frac{1-\exp(-Cst.f(\lambda,\mathbf{r}))}{1+\exp(-Cst.f(\lambda,\mathbf{r}))} = \tanh\left(\frac{1}{2}Cst.f(\lambda,\mathbf{r})\right)$$

These hypotheses led us to compute for each set of values the expression $-\ln(1-\mathbf{r}_p/1+\mathbf{r}_p)$, and compare it with the corresponding values of $\mathbf{r}$ and $\boldsymbol{\lambda}$. We found a very strong positive correlation between the ratio $\lambda/\mathbf{r}$ and the former expression, and the following relation was established:

$$\mathbf{r}_p = \tanh(9.32e-5.\lambda/\mathbf{r})$$

Finally, the ratio $\mathbf{r}_m = \mathbf{p}_m^{\text{exp}}/\mathbf{p}_m^{\text{theo}}$ could be derived simply from the knowledge of $\mathbf{r}_n$ and $\mathbf{r}_p$.

Analysis

From the previous computational study, we are able to propose a relation between the probabilities of monogenesis $\mathbf{p}_m$ and polygenesis $\mathbf{p}_p$, the number of groups $\mathbf{n}$, the expected number of groups in which the innovation appears when no diffusion occurs, $\boldsymbol{\lambda}$, and the time $\mathbf{T}$ for the complete diffusion of the innovation in the population (this last parameter can itself be decomposed as seen previously):

$$\mathbf{r_p} = \tanh(9.32e-5.\boldsymbol{\lambda}/\mathbf{r})$$

$$\Rightarrow \mathbf{p_p} = \tanh\left(9.32e-5.\frac{\boldsymbol{\lambda}.\mathbf{T}}{2.\ln(\mathbf{n})}\right).(1-\exp(-\boldsymbol{\lambda})-\boldsymbol{\lambda}.\exp(-\boldsymbol{\lambda}))$$

$$\Rightarrow \mathbf{p_m} = 1-\exp(-\boldsymbol{\lambda})-\tanh\left(9.32e-5.\frac{\boldsymbol{\lambda}.\mathbf{T}}{2.\ln(\mathbf{n})}\right).(1-\exp(-\boldsymbol{\lambda})-\boldsymbol{\lambda}.\exp(-\boldsymbol{\lambda}))$$

As we have already made clear, it is often not possible to know the probability of emergence of an innovation at one site; in the former expressions, the parameter **λ** remains an unknown variable. However, we are able to compute the probabilities of monogenesis and polygenesis for various values of this parameter, **n** and **T**. The following array summarizes values of the ratio $\mathbf{p_p}/\mathbf{p_m}$ for various values of these parameters. The values of **T** are identical to the values that were extrapolated in Table 4, and therefore may be related to specific sets of the basic parameters **d**, **v**, **de**, **p** and **α**. $\mathbf{T_{max}}$ was chosen equal to 50,000 years.

Table 5

Values of the ratio $\mathbf{p_p}/\mathbf{p_m}$ for various values of λ, **n** and **T**

T, **n** / λ	0.1	1	2	5	10	50	100
1263 years, 10.000	3.14e-5	0.0027	0.0089	0.032	0.068	0.45	1.29
756 years, 1.000	2.5e-5	0.0021	0.0071	0.025	0.054	0.33	0.89
8953 years, 10.000	2.2e-4	0.019	0.066	0.27	0.74	45.9	4300
6329 years, 1.000	2.1e-4	0.018	0.062	0.25	0.67	35.2	2555
31055 years, 10.000	7.7e-4	0.070	0.26	1.73	11	3.3e6	2.2e13
23524 years, 1.000	4.7e-4	0.041	0.15	0.75	2.85	6822	9.3e7

As we already mentioned it in section 1.3, we have investigated values of the density which appear to be small compared to estimations by palaeo-anthropologists. However, a significant difference between our simple model of diffusion and the reality is that in the former, agents could move without constraints. In reality, natural barriers like warm or cold deserts, mountains, oceans, etc. may have significantly increased the time for diffusion to the whole population. If one admits that groups were living in restricted areas (along rivers or lakes or seas, near abundant resources, etc.), the transmission of innovations between these regions may have taken a long time[6]. We simulated this effect by considering smaller densities of groups. We then propose that the estimated densities for prehistory correspond to quite long periods for complete diffusion, which increases the likelihood of polygenesis.

It appears finally that for values of the parameters that are congruent with our knowledge of prehistory, the probability of polygenesis may still be comparable with or higher than the probability of monogenesis for small values of λ (for example 5 expected sites of emergence with 10,000 human groups). If one deals with a large number of innovations which value of λ is large enough, it becomes statistically very likely that some of these innovations will emerge by polygenesis, and others by monogenesis. To rephrase this conclusion in our linguistic framework, if we assume a λ large enough for modern linguistic strategies, the model predicts that at least some of them appeared by polygenesis, possibly the majority.

In the third and last part of this article, in order to further assess the results of our simulations, we will try to adopt a broader perspective concerning "abstract" models and their implicit statements, as we will try to place the polygenesis or monogenesis of linguistic innovations in a context centered on the cognitive capacities of our predecessors.

---

[6] Extreme cases may be Australia, which required significant sea-crossings to be reached, or the Americas, which were only accessible through the northern path between Siberia and Alaska.

3. Discussion

3.1. Cognitive potential and structural polygenesis

An implicit assumption behind the former results

The model we have described can be applied to any kind of cultural innovation. As we have mentioned in the first part of the article, we think that linguistic items or "linguistic strategies" are the best candidates when it comes to the evolution of languages and the development of linguistic diversity. We propose indeed to view them as cultural innovations, and to apply the results of our model to write the first draft of a scenario of the development of linguistic diversity, and of modern languages. However, doing this requires a better understanding of the probability of emergence at one site for linguistic components.

A hidden assumption hides behind the former results of the model: the probability of emergence of the innovation has to remain constant during all the time period considered. Interestingly, analyzing this assumption leads to various comments that may enrich our point of view.

A first aspect is to be related to the putative link we have already mentioned between the emergence of our species and the emergence of modern languages. Our entire former discussion becomes irrelevant if one cannot assume an approximately constant probability of emergence during the large period of time T considered (typically tens of thousands of years). This is exactly what happens if one assumes that the emergence of our species has to be followed immediately by the emergence of the innovations: all the strategies appear very shortly after the emergence of our species. If this emergence is very localized, as in the *Out of Africa* theory, then monogenesis is very likely. We will try to refute this hypothesis by introducing the notion of *cognitive potential*.

Another objection that may be raised against a constant probability of emergence at one site is that interactions between linguistic items create context-dependant probabilities of emergence: a specific strategy will be more likely to appear in some contexts made by other pre-existing strategies.

The notion of cognitive potential

Evolutionary biologists often quote the following saying: "The function makes the organ". This means that an organ will not first appear randomly before being attributed a function, but will appear or change to satisfy a specific functional requirement. This assumption is of course related to the Darwinian law of natural selection. However, if a functional requirement may spawn an organ, this one may also be used for another function than the one it was first developed for. This phenomenon, which has received the name of *exaptation*, is another possibility to create a link between a physiological device and a given function. An example is the wing of the bat, which was formerly an upper leg like for other mammals, but gradually changed to fill another function. More and more biologists estimate that most of functions appear as exaptations, rather than with the primary appearance of an organ. This may especially be the case for the brain, where neural circuits and structures may have played different roles during evolution. One may refer here for example to MacNeilage (1998)'s Frame-Content theory centered on the evolution of the Broca's area.

We may slightly extend this last point of view by saying that the mechanism of exaptation could be translated to the domain of higher cognitive functions, and not only restricted to low-level neural networks. A function could then take advantage of already existing cognitive mechanisms to become active in the behavior of an individual. Various authors have already defended this position. Wang (1991) promoted, for example, the idea of a "mosaic" of cognitive functions which would have led to the emergence of language. Writing is a demonstrative example of a cognitive activity which is not the result of an evolutionary process, but relies on a collection of cognitive abilities and neural areas which primarily evolved for other functions.

We propose to use the term "cognitive potential" to describe the fact that some cognitive functions could *potentially* emerge given an organic or cognitive background. At a given time, what would make them concretely exist or not is a matter of external conditions or events rather than internal requirements.

Modern linguistic strategies may especially be considered as cognitive potentials, relying on various cognitive abilities as memory, integration of various spatial and temporal frames, attribution of thematic roles etc. A strong argument in favor of this analysis is that many linguistic strategies do not necessarily appear in aa languages, but are readily learnt by young infants receiving them as part of their linguistic input. This notion of potentiality is highly significant for our discussion, since it implies that the emergence of modern linguistic strategies did not necessarily take place at the same time as the emergence of our species. This opens the door to a progressive emergence of the features of today's languages, according to cultural innovations taking place in different human groups. It becomes then more plausible to assume a relatively constant probability **Pc** over a larger period of time, which in turn validates the results and hypotheses derived from our computational experiments.

What "events" may trigger the emergence of a new linguistic strategy? This question remains hard to answer, because many factors have to be taken into account. On the one side, common linguistic phenomena, such as grammaticalization processes, may lead to new linguistic forms after a while: primary forms which were more likely to appear first gradually evolve into more complex states. Other elements can be related to cultural facts: behaviors requiring for example to share cognitive representations about spatially and temporally distant situations may have led to the emergence of new linguistic forms to express time and space, just as many new words have been created throughout history to name new concepts or tools.

It is interesting to recall here the gap between the emergence of *Homo sapiens* and the emergence of behaviors like the first sea-crossings to Australia around 60,000 BP, the religious burial of the dead[7]or rock painting. If one accepts that several tens of

---

[7] Archaeologists still debate on the first burials: whether Neanderthals or early *H. sapiens* were burying some of their dead remains very controversial. However, *H. sapiens'* burials with offerings become obvious during the Upper Palaeolithic (50,000 – 10,000 BP) (Klein, 1999: 468–70, 550–3).

thousands of years separate the emergence of our species and the former behaviors, and that such behaviors require specific linguistic abilities, it is reasonable to assume that linguistic evolution accompanying these deep changes occurred a long time after the emergence of our species. It is especially relevant to notice here that these behaviours appeared when *Homo sapiens* had already colonized a significant part of Eurasia and Africa, which implies large areas and multiple natural barriers that prevent fast diffusions from occurring.

Structural genesis

Linguistic strategies interact with each other. Some of them are exclusive or rarely appear together (Greenberg, 1978). One reason could be that once a cognitive demand is satisfied at a linguistic level, there is no need to have another strategy for the same purpose. Case-markers or word-order are such quite redundant strategies. Others may occur together with a high frequency in today's languages. Once again, some explanations may be found in cognitive constraints or economy: two strategies may induce the same cognitive operations and hence save computational time for real speech processing. The X-bar rule of the *Government and Binding Theory*, which defines the position of all heads and specifiers, see (Black, 1999) for definition, may be explained by such computational savings or costs.

Because of these interactions, all strategies are not as likely to occur in a context formed by already existing strategies. As a consequence, instead of considering steady probabilities of emergence for all strategies during the whole time period, probabilities at one site should in theory be recomputed at each new emergence to take the various interactions into account.

Starting from one or several initial forms, families of languages follow their own pathway, and each new emergence modifies the possible directions of evolution the system may take. The term *bifurcation*, as introduced for example in linguistics by Ehala (1996: 2-3) to describe the evolution of linguistic systems, appropriately characterizes the fact that each new emergence, according to external contingent events, forces the system to choose a restricted path toward some of the initial potential configurations it could adopt. This is very similar to the evolution of species, where changes can only occur in the frame defined by the biological characteristics of the organism (Maturana & Mpodozis 2000).

This specific scheme of interactions between the strategies blurs the situation by adding a lot of conditional relationships between the various paths that might be taken by a linguistic system from an initial state. However, one feels intuitively that such structural constraints do not contradict the main result of the model, which is that linguistic strategies are likely to emerge according to polygenesis at various locations. For some of them, the constraints will increase the probability of emergence at one site, while others will appear less likely because of pre-existing strategies. We propose to use the term *structural genesis*, to summarize the fact that linguistic innovations appear under structural constraints.

3.2 A scenario for the development of linguistic diversity and the emergence of modern languages

Given all the previous proposals and hypotheses, we propose the following sketch for the development of languages:

- Due partially to its social origin, linguistic diversity is as ancient as what we may call the human function of language, likely long before the emergence of *Homo sapiens*, as suggested by the behavioral achievements of *pre-sapiens* species.
- Since we postulate that linguistic strategies represent devices to transfer data from a cognitive internal level to an external and shared one (or vice-versa) relying on general cognitive abilities, it is reasonable to assume that these strategies have become increasingly complex with the evolution of our cognitive capacities in the past, especially along the several speciation events which constitute the nodes of our phylogenetic tree.

- Once new cognitive capacities become available, the linguistic “appropriation” or usage of these capacities is not immediate. Cognitive potentials appear that may be instantiated only after a period of time. The evolution of the probability of emergence at one site and in a given linguistic state is determined by the nature of the new cognitive abilities, the benefit for communication of instantiating the strategy, the structural constraints that weigh on the emergence of the innovation and the cultural evolution and requirements at a more general level.
- The former assumptions lead to the possibility of monogenesis or polygenesis of linguistic strategies. Until recently, around 10,000 years ago, the densities of population were very low, therefore leading to rather slow diffusions of linguistic innovations by contact between human groups. The progressively increasing number of sites and larger areas colonized seems to favor the polygenesis of various innovations, as was the case for both agriculture and writing.
- According to Freedman and Wang’s arguments, polygenesis is not necessarily less likely than monogenesis on probabilistic grounds. This proposal has to be revised because of the possible diffusion of an innovation leading to the bypass of polygenesis. However, the results of our experiments and the large number of linguistic strategies show that the polygenesis of at least part of these strategies cannot be ruled out. In case of a large probability of emergence at one site, the model predicts that most of the strategies would have appeared by polygenesis (what the model cannot predict is exactly what the probability of emergence at one site is).

If we now turn to the origin of contemporary languages, the former hypotheses can be reformulated in the following way (if one assumes the *Out of Africa* hypothesis):

- Prior to the emergence of our own species, our predecessors already possessed a function of language. This function was taking diverse surface forms, according to their cognitive abilities (and physiological structures). Neither human language nor linguistic diversity appeared with our species.

- With the emergence of our species, new cognitive potentials opened the way to new linguistic strategies. However, these strategies did not all appear right after the speciation event, but during a long time span of several tens of thousands of years. The various linguistic strategies took place according to external events, especially related to the more general cultural development of our ancestors.
- Whether there were one or several ancestors to modern languages soon after the first *Homo sapiens*, these systems of communication were less complex than today's languages. Assuming a single ancestor for all modern languages makes little sense if this ancestor only shared a few of the characteristics of contemporary languages, most likely the simplest ones, and if most of the modern linguistic features appeared later in various populations (partially constrained by the structural interactions between them, and therefore following different paths starting from the initial emergence). Considering the polygenesis of numerous linguistic strategies, or their monogenesis *because of diffusion*, leads to partially emptying this hypothesis of its substance and interest. Moreover, if the transition from pre-modern to modern humans took place slowly and in a large number of groups, various sources could have been at the origin of the linguistic families that later spread all over the world.
- Trying to reconstruct a unique ancestor to today's languages on the basis of comparisons of words or typological structures can only be validated if alternative scenarios built on the polygenesis of the considered features are demonstrated as significantly less likely than the monogenetic hypothesis.

4. Conclusion

The question of the monogenesis or polygenesis of languages is a difficult one, because it involves numerous factors as diffuse as the palaeo-demographic conditions of our ancestors, the mechanisms of a speciation event, the relationships between general cognitive abilities and linguistic "tools," etc. Most of our proposals are not firm demonstrations, but aim nevertheless to shed light on elements that are often forgotten in current debates. We have aimed to present this question in the more general framework of the evolution of linguistic diversity. By putting forward a scenario based on the possible polygenesis of linguistic strategies over long periods of time, we conclude that the fact of having a single ancestor to all modern languages makes little sense, if this ancestor was much simpler than today's languages, and if many features evolved later independently in many human groups scattered all over Earth. Moreover, we strongly believe that language did not emerge with our species, and that many linguistic strategies, even already sophisticated, were used by our former ancestors.

This position seems to run counter to the hypotheses of linguists such as Merritt Ruhlen (1994), who proposed words of the *proto-sapiens, Homo sapiens'* "first" language, or Murray Gell-Mann, who defends the fact that there are various visible arrows of time in today's languages, one of them being word-order (see contribution in this volume). Despite the fact that alternative models based on polygenesis have not been taken into account to contrast their proposals, their positions do not necessarily contradict our hypotheses, since some (central) features, such as word-order or core lexicon, might have already existed in the language of pre-modern men[8]. It is actually reasonable for us to assume that strategies such as word-order already existed long before the last 100,000 years: as experiments on monkeys have shown, sequential

---

[8] We do not enter here into the debate whether word replacement totally erases the traces of such ancient lexicons or not (Ringe, 1992).

ordering is an ancient ability (Terrace, 2000), and does not seem to require highly sophisticated cognitive capacities. This has to be compared with linguistic strategies requiring more complex internal representation, such as expressions of time, aspect, mode, causation (Shibatani and Pardeshi, 2001), etc. This view disagrees with some scholars' proposal of a catastrophic transition from a proto-language to a fully syntactic one, and asks for the cognitive loads of various linguistic strategies.

Contact between small groups of humans and possible linguistic innovations play a central role in our proposals. Such emphasis will appear close to the emphasis on contact in historical linguistics. These elements are often put forward to criticize the simplified, but useful, model of the *Stammbaum*. We therefore propose to project the controversies of the recent evolution of languages to their origins. As well as for recent contacts and consequent language changes, these considerations may be useful to refine our knowledge and reconsider well established theories of the prehistory of languages that may reveal themselves to be implausible in regard of such arguments.

Acknowledgments

We thank James Minett for his precious help and useful comments to prepare this paper and improve its quality. This work was partially supported by grant #9040781 from the Research Grants Council of the Hong Kong SAR, China.

References

Arbib, Michael A. and Rizzolatti, Giacomo. (1996) Neural expectations: a possible evolutionary path from manual skills to language. *Communication and Cognition* 29 (3/4), 393–4242.

Auroux, Sylvain, and Mayet, Laurent. (2001) Entretien avec Sylvain Auroux, Le mystère des racines. *Sciences et Avenir* 125, 12–5.

Baddeley, Alan D. (1986) *Working memory*. Oxford: Clarendon Press.

Bickerton, Derek. (1990) *Language and Species*. Chicago: The University of Chicago Press.

Biraben, Jean-Noël, Masset, Claude, and Thillaud, Pierre L. (1997) Le peuplement préhistorique de l'Europe. In *Histoire des populations de l'Europe, tome 1 — Des origines aux prémices de la révolution démographique* (pp. 39–92). Paris: Fayard.

Black, Cheryl A. (1999) *A step-by-step introduction to the Government and Binding theory of syntax*. Summer Institute of Linguistics, http://www.sil.org/mexico/ling/E002-IntroGB.pdf.

Bocquet-Appel, Jean-Pierre, and Demars, Pierre-Yves. (2000) Population kinetics in the upper Palaeolithic in Western Europe. *Journal of Archaeological Science* 27, 551–70.

Cann, Rebecca L., Stoneking, Mark, and Wilson, Allan C. (1987) Mitochondrial DNA and Human evolution. *Nature* 325, 31–6.

Cavalli-Sforza, Luigi Luca, Menozzi, Paolo, and Plazza, Alberto. (1994) *The history and geography of human genes*. Princeton: Princeton University Press.

Chomsky, Noam. (1975) *Reflections on language*. New York: Pantheon Books.

Coupé, Christophe and Hombert, Jean-Marie. (2002) Language at 70,000 BP: Evidence from sea-crossings. In *Proceedings of the Fourth International Conference on the Evolution of Language* (p. 27). Harvard.

Coupé, Christophe. (2003) *De l'origine du langage à l'origine des langues: Modélisations de l'émergence et de l'évolution des systèmes linguistiques*. PhD dissertation in Cognitive Sciences, Univ. Lyon 2.

Croft, William. (1990) *Typology and Universals. Cambridge Textbooks in Linguistics*. Cambridge: Cambridge University Press.

De Waal, Franz. B. (1998) *Chimpanzee Politics: Power and Sex among Apes*. Baltimore: Johns Hopkins University Press.

D'Errico, Francesco, Henshilwood, Christopher, and Nilssen, Peter. (2001) An engraved bone fragment from c. 70,000-year-old Middle Stone Age levels at Blombos Cave, South Africa: implications for the origin of symbolism and language. *Antiquity* 75 (288), 309–18.

Dessalles, Jean-Louis. (2000) *Aux origines du langage — Une histoire naturelle de la parole*. Paris: Hermes Science Publications.

Dixon, Robert M.W. (1997) *The rise and fall of languages*. Cambridge: Cambridge University Press.

Dunbar, Robin I. M. (1993) Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences* 16, 681–94.

Dunbar, Robin I. M. (1996) *Grooming, Gossip and the Evolution of Language*. London: Farber and Farber.

Ehala, Martin. (1996) Self-organisation and language change. *Diachronica* 13 (1), 1–28.

Enard, Wolfgang, Przeworski, Molly, Fisher, E. Simon, Lai, Cecilia S. L., Wiebe, Victor, Kitano, Takashi, Monaco, Tony and Pääbo, Svante. (2002) Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* 418, 869.

Freedman, David. A. and Wang, William S.-Y. (1996) Language polygenesis: A probabilistic model. *Anthropological Sciences* 104 (2), 131–8.

Greenberg, Joseph H. (1978) Typology and cross-linguistic generalizations. In Joseph H. Greenberg (Ed.) *Universals of human language. Method and Theory, volume 1* (pp. 33–61). Stanford: Stanford University Press.

Harpending, Henry C., Batzer, Mark A., Gurven, Michael, Jorde, Lynn B., Rogers, Alan R., and Sherry, Stephen T. (1998) Genetic traces of ancient demography. *Proceedings of the National Academy of Sciences of the United States of America* 95, 1961–7.

Harpending, Henry C., Sherry, Stephen T., Rogers, Alan R. and Stoneking, Mark. (1993) The genetic structure of ancient human populations. *Current Anthropology* 34 (4), 483–96.

Hassan, Fekri A. (1981) *Demographic archaeology. Studies in Archaeology.* Academic Press.

Hawks, John, Hunley, Keith, Lee, Sang-Hee, and Wolpoff, Milford. (2000) Population bottlenecks and Pleistocene human evolution. *Molecular Biology and Evolution* 17 (1), 2–22.

Hockett, Charles F. (1960) The origin of speech. *Scientific American* 203, 88–96.

Jacquesson, François. (2001) Pour une linguistique des quasi-déserts. In Anne-Marie Loffler-Laurian (Ed.) *Etudes de linguistique générale et contrastive. Hommage à Jean Perrot* (pp. 199–216). Paris: Centre de Recherche sur les Langues et les Sociétés.

Klein, Richard. G. (1999) *The human career, human biological and cultural origins*. Chicago; London: The University of Chicago Press.

Labov, William. (1963) The social motivation of a sound change. *Word* 19, 273–303.

Lahr, Marta Mirazon and Foley, Robert. (1994) Multiple dispersals and modern human origins. *Evolutionary Anthropology* 3, 48–60.

Lai, Cecilia S., Fisher, Simon E., Hurst, Jane A., Vargha-Khadem, Faraneh, and Monaco, Anthony P. (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413, 519–23.

MacNeilage, Peter F. (1998) The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences* 21 (4), 499–511.

Marsico, Egidio. (1999) What can a database of proto-languages tell us about the last 10,000 years of sound changes? In *Proceedings of the XIVth International Congress of Phonetic Sciences*. San Francisco.

Marsico, Egidio, Coupé, Christophe, and Pellegrino, François. (2000) Evaluating the influence of language contact on lexical changes. In *Proceedings of the third Conference on the Evolution of Language* (pp. 154–5). Paris: Ecole Normale Supérieure des Télécommunications.

Marwick, Ben. (2002) Raw material transportation as an indicator of hominid symbolic linguistic capacity during the Pleistocene. In *Proceedings of the Fourth International Conference on the Evolution of Language* (p. 74). Harvard.

Maturana, Humberto and Mpodozis, Jorge. (2000) The origin of species by means of natural drift. *Revista Chilena de Historia Natural* 73, 261–310.

Murray, James D. (1994) *Mathematical biology. Second, Corrected Edition*. Berlin; Heidelberg; New York: Springer Verlag.

Naccache, Albert F. H. (2002) Sociolinguistic approaches to the prehistory of the Mashriqian (Semitic) languages. In *Proceedings of the Fourth International Conference on the Evolution of Language* (p. 80). Harvard.

Nettle, Daniel. (1999a) *Linguistic Diversity*. Oxford Linguistic. Oxford: Oxford University Press.

Nettle, Daniel. (1999b) Using social impact theory to simulate language change. *Lingua* 108, 95–117.

Nettle, Daniel. (1999c) Is the rate of linguistic change constant? *Lingua* 108, 119–136.

Nettle, Daniel and Romaine, Suzanne. (2000) *Vanishing voices*. Oxford: Oxford University Press.

Nichols, Johanna. (1992) *Linguistic diversity in space and time*. Chicago; London: University of Chicago Press.

Pagel, Mark. (2000) The history, rate and pattern of world linguistic evolution. In Chris Knight, Michael Studdert-Kennedy, and Jim Hurford (Eds.) *The evolutionary emergence of language* (pp. 391–416). Cambridge: Cambridge University Press.

Ringe, Donald. A. J. (1992) On calculating the factor of chance in language comparison. *Transactions of the American Philosophical Society* 82 (1), 1–109.

Rizzolatti, Giacomo, Fadiga, Luciano, Gallese, Vittorio, and Fogassi, Leonardo. (1996) Premotor cortex and the recognition of motor action. *Cognitive Brain Research* 3, 131–41.

Ruhlen, Merritt. (1994) *The origin of language. Tracing the evolution of the mother tongue*. New York: John Wiley and Sons.

Sauvet, Georges and Wlodarczyk, André. (1995) Eléments d'une grammaire formelle de l'art pariétal paléolithique. *L'Anthropologie* 99 (2–3), 193–211.

Savage-Rumbaugh, Sue, Shanker, Stuart G., and Taylor, Talbot J. (1998) *Apes, language, and the human min*d. Oxford: Oxford University Press.

Schoenemann, P. Thomas. (1999) Syntax as an emergent characteristic of the evolution of semantic complexity. *Minds and Machines* 9, 309–46.

Sherry, Stephen T., Rogers, Alan R., Harpending, Henry C., Soodyall, Himla, Jenkins, Trefor, and Stoneking, Mark. (1994) Mismatch distributions of mtDNA reveal recent human population expansions. *Human Biology* 66 (5), 761–75.

Sherry, Stephen T., Harpending, Henry C., Batzer, Mark A., and Stoneking, Mark. (1997) Alu evolution in human populations: Using the coalescent to estimate effective population size. *Genetics* 147, 1977–80.

Shibatani, Masayoshi and Pardeshi, Prashant. (2001) The causative continuum. The grammar of causation and interpersonal manipulation. In Masayoshi Shibatani (Ed.) *Typological Studies in Language 48* (pp. 85–126). Amsterdam; Philadelphia: John Benjamins Publishing Company.

Sperber, Dan. (1995) How do we communicate? In John Brockman and Katinka Matson (Eds.) *How things are: A science toolkit for the mind* (pp. 191–9). New York: Morrow.

Stringer, Chris B. and Andrews, Peter. (1988) Genetic and fossil evidence for the origin of modern humans. *Science* 239, 1263–7.

Terrace, Herbert S. (2000) Serial expertise and the evolution of language. In *Proceedings of the Third International Conference on the Evolution of Language* (pp. 154–155). Paris: Ecole Nationale Supérieure des Télécommunications, Paris.

Thorne, Alan G. and Wolpoff, Milford H. (1992) The multiregional evolution of humans. *Scientific American* 266, 76–83.

Wang, William S.-Y. (1973) Chinese language. *Scientific American* 228, 50–60.

Wang, William S.-Y. (1991) Explorations in language evolution. In William S.-Y. Wang (Ed.) *Explorations in Language* (pp. 105–130). Taiwan: Pyramid Press.

Zhivotovsky, Lev A., Bennett, Lynda, Bowcock, Anne M., and Feldman, Marcus W. (2000) Human population expansion and microsatellite variation. *Molecular Biology and Evolution* 17 (5), 757–67.

Zubrow, Ezra B. (1989) The Demographic Modelling of Neanderthal Extinction. In Paul Mellars and Chris Stringer (Eds.) *The Human Revolution: behavioural and biological perspectives on the origins of modern humans* (pp. 212–31). Princeton, New Jersey: Princeton University Press.