Homo sapiens Populates the Earth: A Provisional Synthesis, Privileging Linguistic Evidence

PATRICK MANNING
Northeastern University

Recent discoveries have provided much new information on the emergence and spread of modern humans.[1] Scholars in the field of genetics have established that *Homo sapiens* originated in Africa in about 200,000 B.P., and that our species subsequently displaced all previous hominid species. Recent results in paleontology have gone far toward confirming these views.[2] Further, while only a few scholars with degrees in history have undertaken analysis of the earliest human migrations, the comprehensive methodological approach associated with world history has been important in developing new insights into early human history.[3] That is, geneticists, paleontologists, archaeologists, and earth scientists have tended increasingly to overcome the

---

[1] The author expresses thanks to Luigi Luca Cavalli-Sforza, Christopher Ehret, Merritt Ruhlen, and an anonymous reader of this journal for comments on an earlier version of this essay.

[2] For conciseness I identify our species as *"Homo sapiens"* rather than use the more precise *"Homo sapiens sapiens."* By "B.P." I mean "before present" or "years ago." For an authoritative but argumentative survey of genetic and archaeological interpretation of human evolution and migration, see Christopher Stringer and Robin McKie, *African Exodus: The Origins of Modern Humanity* (New York: Henry Holt, 1996); see also Sally McBrearty and Alison S. Brooks, "The Revolution That Wasn't: A New Interpretation of the Origin of Modern Human Behavior," *Journal of Human Evolution* 39 (2000): 453–563. For an accessible summary of recent archaeological debates on early *Homo sapiens*, see Kate Wong, "The Morning of the Modern Mind," *Scientific American*, June 2005, pp. 86–95.

[3] David Christian and Christopher Ehret are two historians who have analyzed early human migrations in print. Christian, *Maps of Time: An Introduction to Big History* (Berkeley: University of California Press, 2003), pp. 176–202; Ehret, *The Civilizations of Africa: A History to 1800* (Charlottesville: University Press of Virginia, 2002), pp. 20–25. For a

---

parochialism of their disciplines, linking and comparing various sorts of evidence. Taken together, scholars from these disciplines have begun to meet on the terrain of world history to revolutionize our understanding of the early life of *Homo sapiens*.

Yet there remain major gaps in our understanding of human expansion. While it is accepted that all humanity came "out of Africa," there remain disputes on the path and timing of migration from Africa to other regions. The maps and descriptions of early human migration tend to neglect migrations within Africa and include arrows suggesting a general dispersion of migrants from Africa in several directions.[4] Disciplinary parochialism reasserts itself from time to time: for instance, geneticists have not yet worked sufficiently to link their results to results from other fields of study or to develop alternative models within genetics that may yield different interpretations.[5]

Information from another field of study—linguistics—has the potential to clarify the paths of early human migration. This article argues that evidence on language classification can and should be used systematically in interpreting early human migrations.[6] In it I apply techniques for analyzing language-group distributions that have led successfully to reconstructing Indo-European, Bantu, and Austronesian expansions of the past four thousand to eight thousand years. I combine these techniques with the argument that they may appropriately be applied to earlier times. This is not the first application of linguistic data to the interpretation of human dispersal, though I argue that this interpretation is distinct in its conclusions and more systematic in its approach than previous interpretations.[7]

My narrative of early human migration begins with the movement

---

thoughtful journalistic synthesis of human origins and early migrations, see Steve Olson, *Mapping Human History: Genes, Race, and Our Common Origins* (Boston: Houghton Mifflin, 2003).

[4] Luigi Luca Cavalli-Sforza, Paolo Menozzi, and Alberto Piazza, *The History and Geography of Human Genes,* abridged ed. (Princeton, N.J.: Princeton University Press, 1994), p. 156. For a map closer to the present interpretation, see Olson, *Mapping Human History,* p. 135. See also Christian, *Maps of Time,* p. 193.

[5] For a genetic argument on migration unmediated by cross-disciplinary analysis, see Bo Wen et al., "Genetic Evidence Supports Demic Diffusion of Han Culture," *Nature* 431 (2004): 302–305.

[6] Luigi Luca Cavalli-Sforza has been exemplary among geneticists in using evidence of language to confirm his analysis of genetics. Yet his approach, as I will argue, has been to appropriate the most general results of language classifications rather than inquire more deeply into language dynamics and linguistic methods, so that his linguistic insights are muted and, in some cases, incorrect. Cavalli-Sforza, Menozzi, and Piazza, *Human Genes,* pp. 164–167, 220–222, 263–266, 317–320, 349–351.

[7] Merritt Ruhlen, *The Origin of Language: Tracing the Evolution of the Mother Tongue* (New York: Wiley, 1994); Cavalli-Sforza, Menozzi, and Piazza, *Human Genes.*

of the densest human populations from equatorial East Africa to the northern savannas of Africa. It proceeds then to trace waterborne migration across the mouth of the Red Sea to South Arabia, then eastward along the shores of the Indian Ocean to the South China Sea, and later across the oceanic straits to Australia and New Guinea, all by about 50,000 B.P. Thereafter, the analysis considers four possible routes by which humans might have moved from the tropics into temperate zones of Eurasia, and concludes that the easternmost route, along the eastern coast of Asia, is attested most clearly by linguistic evidence. As I argue, this movement into temperate regions took place from about 45,000 to 30,000 years ago; it included the human occupation of Europe and the displacement of its preexisting Neanderthal population. Further, I argue that this same wave of migration continued north of the Pacific and to the Americas, also in the period before the great Ice Age beginning 30,000 B.P. Thereafter, the initial populations in each major world region continued to differentiate into subgroups. Thus, well before the beginnings of agriculture about 15,000 B.P., the populations of the various world regions had settled into place, and the languages of their descendants give us strong evidence of their ancestral migrations.

As will be shown, linguistic data are central to the details of this interpretation. Why have language data not been used more in interpretations of early human history? Language may provide substantial information on early migrations, but linguistics is a field riven with controversy. Conflicting priorities in language classification leave us with contradictory classifications of the world's languages: do languages reveal a global pattern or are the patterns restricted to localities? In part, the current contradictions in linguistic interpretations echo those of recent years in genetics and paleontology. But while both geneticists and paleontologists carried on vigorous debates until each field had confirmed a widely accepted interpretation of the data—one that confirmed the "out of Africa" vision of human origins and dispersal—historical linguists have chosen not to give priority either to resolving their classificatory differences or to developing broad interpretations of human migration. In a second area of dispute, while some linguists think that language data provide important indications on human origins and dispersal, others argue that linguistic data give no information at all for times more than 10,000 years ago.[8]

---

[8] For contending viewpoints, see Colin Renfrew, April McMahon, and Larry Trask, eds., *Time Depth in Historical Linguistics,* 2 vols. (Cambridge: McDonald Institute for Archaeological Research, 2003).

The next section of this article demonstrates the differences among linguists on the classification of languages. It shows why I have accepted the view that virtually all the languages of the world may be classified into twelve phyla, each with a time depth of more than 20,000 years, in contrast to views arguing, for instance, that there are over one hundred separate language families, no one of which may be traced back further than 10,000 years. The third section of the article summarizes the methodology I use to propose interpretations of early human migration: analyzing data on language classification and using a world-historical approach of combining language data with other data from other fields. The two final sections apply this global combination of methods to address, chronologically, the tropical migration of humans from Africa to the Pacific in the era from about 80,000 to 50,000 B.P. and then the human occupation of the temperate Old World and the Americas from about 40,000 to 30,000 B.P.

CLASSIFICATION OF LANGUAGES: DEBATES ON LINKAGE AND TIME FRAME

Evidence from historical linguistics has been central to resolving puzzles about the origins and migrations of several populations. The most fundamental example is that of speakers of Indo-European languages. While disputes continue about the precise location and especially the timing of Indo-European origins, the linguistic data affirm that the homeland must be near to the Black Sea, and other data support this conclusion. For the Austronesian languages—spoken throughout Southeast Asia and the Pacific and in Madagascar—analysis has shown that the languages originated in coastal south China (where they are no longer spoken) and that speakers migrated to Taiwan and then migrated by stages to wider regions. In the most controversial and most definitively resolved instance, the Bantu languages—spoken throughout central, eastern, and southern Africa—are shown conclusively to have originated in southeastern Nigeria, where their nearest neighbor languages are spoken.[9] Despite the success of these analyses, world

---

[9] On Indo-European languages, see J. P. Mallory, *In Search of the Indo-Europeans: Language, Archaeology, and Myth* (London: Thames and Hudson, 1989), p. 262; on Austronesian languages, see Peter Bellwood, *Prehistory of the Indo-Malaysian Archipelago,* 2nd ed. (Honolulu: University of Hawai'i Press, 1997), pp. 96–127; on Bantu languages, see Christopher Ehret, "Bantu Expansions: Re-Envisioning a Central Problem of Early African History," *International Journal of African Historical Studies* 34 (2001): 5–41.

historians have not found it easy to address linguistic data globally. The obstacle is that the inconsistency of language classification has impeded historians from using language data on a world-historical level. While language classification has led to successful historical analysis at the regional levels identified above, it has been difficult to utilize language data for global comparisons because the language units currently in favor in various parts of the world are inconsistently defined.

What is the best summary of current knowledge on classification of languages? The nineteenth-century work of Franz Bopp in classifying the Indo-European family of languages set the standard for more than a century of language classification worldwide.[10] The basic principle is that of "genetic" linguistic evolution: any given language may give birth to several "daughter" languages through gradual change in both lexicon and grammar. Detailed empirical analyses of lexicon and grammar in various languages are conducted to identify the patterns of such change and should enable partial reconstruction of ancestral languages. While linguists accept this principle, they disagree on the priorities in its implementation. Some analyze two or three languages at a time; others analyze larger numbers. Some linguists set the very exacting standard of creating a completely reconstructed system of sound changes between any two languages before confirming a genetic relationship between the languages.[11]

Linguists accept in general the existence of large-scale linguistic phyla. Linguistic phyla or super-families are classifications including all languages that can be demonstrated to have genetic relationships with each other. While the genetic logic of language evolution makes inevitable the postulation of phyla, many claim that that it is practically impossible to identify phyla, again because of the difficulty of identifying complete systems of sound changes.

Thus, despite the apparent clarity of principles that ought to yield consistent classification of the world's languages and interpretation of their migratory history, it is easy to demonstrate the inconsistency of the currently prevailing language classifications. Table 1 and the appendix on which it is based summarize roughly one hundred lan-

---

[10] Franz Bopp, *Vergleichende Grammatik des sanskrit, zend, armenischen, griechischen, lateinischen, litauischen, altslavischen, gothischen und deutschen*, 3 vols. (Berlin: F. Dümmler, 1833–1837).

[11] The divergences in practices of language classification seem to have grown since 1950. In this study, rather than trace linguists' debates in detail, I have chosen—especially through Table 1—to focus on demonstrating the contradictory nature of their conclusions.

guage families of the world as they are identified on the *Ethnologue* Web site, an authoritative summary of the current classifications of linguists. I have organized the families to show that they reflect three competing but coexisting categories of breadth in language classification. The numbers of languages in each family and the indentation of terms in the table help to identify the divergences among linguists on classification of languages. These categories distinguish the approaches to classification, favoring the identification of small groupings, larger groupings, and phyla; contested language groupings are identified in parentheses. Category 1 contains eight major language groups (all but two of them with seventy-five or more languages), whose existence is accepted by virtually all linguists. (Some call these groups phyla and others call them families.) In Category 2, there are twenty-two major language groups (all but four of them with ten or more languages) whose existence is accepted by virtually all linguists; the dispute is that some linguists see these families as subphyla of the phyla listed below each group of families, while others treat these families as independent of each other, and contest the existence of the encompassing phyla. Category 3 contains seventy-three groups (nearly fifty of them with fewer than ten languages each) and a total of roughly 950 languages. Those who accept phyla in general recognize an encompassing Amerind phylum with 950 languages, and identify six subphyla within it.[12] Most linguists who specialize in these languages claim that few linkages can be established among the seventy-three groups.

There exists no "consensus" view of human language classification. Rather, there is what might be called an "armed truce" of localized camps, each armed with a different approach. Overall, those who accept the practicability of identifying phyla see human languages as consisting of about twelve phyla of roughly parallel extent.[13] Those who deny the practical knowability of phyla, especially specialists in Amerindian languages, see a patchwork of languages with little overall pattern.[14] Others fall between these limits. The encyclopedias of

---

[12] Na-Dene and Eskimo-Aleut, the language families that are outside of Amerind, are accepted as families even by those who deny the grouping of American languages into large families.

[13] Even within the theorists of Dene-Caucasion there are differences and evolution in viewpoint. For instance, if Dene-Caucasian is accepted as a phylum, then Sino-Tibetan within it loses its status as a phylum.

[14] Scholars in this group, however, tend not to deny the existence of such large groupings as the four African phyla, though they would not use the term "phyla" in describing them.

TABLE 1.   Three categories of language groups*

Category 1: Large language groupings (existence uncontested)

Afroasiatic (372)
Australian (258)
Dravidian (75)
Khoisan (29)
Niger-Congo (1489)
Nilo-Saharan (199)
Sino-Tibetan (365)
South Caucasian (5)

Category 2: Large groupings (uncontested) vs. larger groupings (contested)

Altaic (65)
Chukotko-Kamchatkan (5)
Eskimo-Aleut (11)
Indo-European (443)
Uralic (38)
Japanese (12)
Eurasiatic (similar to "Nostratic")**—570 **(contested)**

Austroasiatic (168)
Austronesian (1262)
Hmong-Mien (32)
Tai-Kadai (70)
Austric—1530 **(contested)**

Andamanese (13)
East Bird's Head (3)
East Papuan (36)
Sepik-Ramu (104)
Torricelli (48)
Trans-New Guinea (552)
West Papuan (26)
Indo-Pacific—770 **(contested)**

Basque (3)
Na-Dene (47)
North Caucasian (34)
Sino-Tibetan (365)
Yenisei Ostyak (2)
Dene-Caucasian—450 **(contested)**

Category 3: Small groupings (uncontested) vs. large groupings (contested)

Alacalufan (2) . . .
Zaparoan (7)
(73 groups total)
Amerind (with 6 subphyla)—approx. 950 languages **(contested)**

Map 1. World language groups, ca. 1500.

linguistics, rather than sharpening these differences, speak vaguely of language "families" and include a mix of both points of view.[15] In the remainder of this article I assume that the best summary of existing knowledge on language classification is that there exist twelve phyla, as indicated in Map 1.

How far back in time can major language groups be traced? I argue, along with some linguists, that present linguistic phyla have existed for at least twenty thousand years and in some cases as much as eighty thousand years. More commonly, linguists argue that present linguistic families or phyla can be traced back no more than 10,000 years and thus are of relevance to the study of human migrations only in the past ten thousand years. Many historical linguists, knowing the relatively rapid rate at which much vocabulary changes, accept the view that the ancestors of today's languages would be different beyond recognition if one tried to trace them back beyond 10,000 years ago. Even those who accept the existence of language phyla have been daunted by the limitations of "glottochronology." This early attempt to estimate the absolute dates for separation of languages sought to apply a linear model at too large a scale.[16] For a standard list of some two hundred words, one assumed a constant rate of change in words over time, so that in comparing any two languages, the percentage of cognates shared by the two gave an indication of the time of their separation. This procedure, which in any case was considered to be applicable to changes only for the last several thousand years, rapidly became controversial, and its use declined, both because of the difficulties in agreeing on cognates and because it became clear that the rate of change in words was not constant over time.[17]

---

[15] Major resources on languages include R. E. Asher and J. M. Y. Simpson, eds., *The Encyclopedia of Language and Linguistics,* 10 vols. (Oxford: Pergamon Press, 1994); Merritt Ruhlen, *A Guide to the World's Languages,* vol. 1, *Classification* (Stanford, Calif.: Stanford University Press, 1987); and Kenneth Katzner, *The Languages of the World,* 3rd ed. (London: Routledge, 2002). See also the extensive collection of data on languages on the *Ethnologue* Web site, www.ethnologue.org.

[16] In the 1950s Morris Swadesh coined the terms "lexicostatistics" and "glottochronology," based on the notion of a fairly regular rate of change in the core vocabulary of languages, at the rate of some 14 percent over a thousand years. Swadesh, *The Origin and Diversification of Languages,* ed. Joel Shertzer (Chicago: Aldine, Atherton, 1971). For a recent discussion, see Christopher Ehret, "Testing the Expectations of Glottochronology against the Correlations of Language and Archaeology in Africa," in Renfrew, McMahon, and Trask, *Time Depth in Historical Linguistics,* chap. 15.

[17] In particular, the more basic vocabulary terms seem less likely to change than terms that are less commonly used and less central to existence. In a genetic parallel to this varying rate of linguistic change, some parts of the genome mutate at different rates than others.

TABLE 2.   Language groups ancestral to Polynesian and Bantu
            (showing numbers of languages at each level)

Austronesian (1262)
  Malayo-Polynesian (1239)
    Central-Eastern (706)
      Eastern Malayo-Polynesian (541)
        Oceanic (502)
          Central-Eastern Oceanic (234)

Niger-Congo (1489)
  Atlantic-Congo (1390)
    Volta-Congo (1316)
      Benue-Congo (938)
        Bantoid (668)
          Southern (643)
            Bantu (501)

A different approach to language history, based on tree diagrams of the genetic relationships within a language family, is clearer in presenting the case that language phyla represent communities of great age. Table 2 shows portions of the family tree for two thoroughly studied groups of languages: the Bantu languages within the Niger-Congo phylum and the Polynesian languages within the Austronesian family. The Bantu languages are about five hundred languages distributed across central, eastern, and southern Africa, and their origin has been traced back to about 4,000 years ago; the Central-Eastern Oceanic languages are more than two hundred languages of the Pacific, including the Polynesian languages, and their origin is traced by archaeological remains to at least 2,500 years ago. As indicated in the table (based on the *Ethnologue* Web site), the work of classification has identified some six previous branches in Niger-Congo languages before the development of Bantu; similar work has identified some five previous branches in Austronesian before the development of Central-Eastern Oceanic.[18] If the previous branches took anywhere near the same amount of time to develop as the last grouping listed has existed (that

---

[18] Table 2 is based on data from the *Ethnologue* Web site, www.ethnologue.org. On the time frame of the emergence of Central-Eastern Oceanic and Bantu language groups, see Bellwood, *Indo-Malaysian Archipelago,* pp. 113–116; and Ehret, "Bantu Expansions."

is, two thousand to four thousand years for each branching), then it is clearly implied that the ancestors of all the Austronesian speakers or all the Niger-Congo speakers have been traced to a time well before 10,000 B.P.

A larger-scale case for the deep historical depth of language groups lies in the languages of Australia and New Guinea. The languages of Australia and the Indo-Pacific phylum centered in New Guinea appear to have come into existence with the settlement of these regions some 50,000 years ago—they were the only language groups spoken in those regions until the recent arrival of Austronesian speakers.[19] If these two phyla remain identifiable after so many years of language change, then other phyla may represent a similar time depth. Of course the tasks of determining the chronological depth of the various language phyla or groupings will be difficult, and our methods are very crude so far. Thousands of individual languages have been lost in recent times, and more were lost in earlier times. Sometimes the disappearance of a language resulted from the populations dying out, but more commonly it resulted from the populations adopting other languages.[20] Nevertheless, I believe that linguistic analysis, linked to studies of archaeology and genetics, will confirm the longevity of language phyla and the consistency of language data with other evidence on early humans.[21]

The conflicting summaries of language data leave historians with a major dilemma. First, if one recognizes phyla as having great time depth, then language data appear to confirm and strengthen interpretations of early human migration based on genetic and archaeological data, as I argue below. Second, if we interpret human migration through a hundred independent language families that can be traced back no more than five thousand to ten thousand years, we would con-

---

[19] Australian languages include sharply different subgroups, but most specialists assume they are related to each other. The Trans-New Guinea family (over 550 languages) is widely accepted, but the broader classification of Indo-Pacific is not accepted by all.

[20] By a similar logic, one can imagine that not only individual languages but whole phyla of languages have ceased to exist, as their populations became absorbed into others for which the populations managed to reproduce themselves more successfully. Frances Karttunen and Alfred W. Crosby, "Language Death, Language Genesis, and World History," *Journal of World History* 6 (1995): 157–174.

[21] A fuller demonstration of the case for this longevity of language phyla will require modeling of how languages within the twelve phyla of today, changing structure and lexicon at known rates, could be shown to have descended from ancestral languages of 50,000 or more years ago. This presentation does not take up that task but instead focuses on portraying the interpretation of migration that should result if such longevity of language phyla can be demonstrated.

clude that there had been many tiny populations in the Americas, moving only small distances, while Eurasia and especially Africa had large-scale population expansions. Third, if we rely on the same hundred language families but assume they are relevant for earlier times, we might conclude that the Americas were the ancestral human homeland, and that Eurasia had been settled from the Americas, since there was greater differentiation of language and population in the Americas than elsewhere. By the same logic, New Guinea and Southeast Asia would be seen as a center from which population expanded.[22] Yet a fourth approach would be to conclude that language data are not relevant to long-term studies of migration, and this in practice is the approach that has prevailed until now.

How did this interpretive confusion arise? Linguists are divided very unequally among the languages they study, and the process of classification has been slow. There are many issues to address in the study of language, and linguists are interested more in current than historical language. Classification studies have been relatively marginal, as linguists have concentrated more fully on grammatical and lexical characteristics of individual languages. Glottochronology, the statistical analysis of language change, ran into early obstacles and has remained limited by them. These are not trivial problems, but there may be ways to solve them other than giving up and concluding that the history of languages cannot be reconstructed beyond that of localized groups in recent times. At a time when such rapid strides are being made in early human history, historians have an interest in learning everything possible from the analysis of language. While it will take the work of linguists themselves to sort out the contradictions in their analysis, the encouragement of historians and the perspective of global interpretation may be helpful in clarifying the historical interpretation of language. It may be useful to remember the experience of Alfred Wegener, whose early insights on continental drift were long ignored, but helped nonetheless to elucidate the very specific mechanisms of plate tectonics that are now known to sustain global geographic patterns.[23]

---

[22] To phrase these views with reference to Table 1, the first approach accepts the twelve phyla listed and assumes they apply to the past 50,000 years; the second approach rejects the notion of phyla and assumes that the families listed apply to the past 10,000 years; the third approach rejects the notion of phyla but assumes that the families listed apply to the past 50,000 years.

[23] Alfred Wegener, *Die Enstehung der Kontinente und Ozeane* (Braunschweig: F. Vieweg, 1915); Martin Schwarzbach, *Alfred Wegener, the Father of Continental Drift* (Madison, Wisc.: Science Tech, 1986).

Data and Assumptions in Analyzing Early Human Migration

Language Phyla and "Tree Models"

My analysis of language classifications relies most fundamentally on the research of the late Joseph E. Greenberg. Greenberg did more than anyone else to assemble a coherent and balanced picture of the main groupings of human languages. Over a long career, he classified the languages of Africa, the Americas, much of Eurasia, and parts of the Pacific.[24] Greenberg also wrote extensively on the methodology of language classification; such classification began with the work of Sir William Jones, who in a 1786 book on Sanskrit suggested that it might be related to Greek, Latin, and Persian. In 1816, the German philologist Franz Bopp published the first comparative grammar on what became known as the Indo-European languages and expanded it in later editions. Indeed, Greenberg explicitly invoked the heritage of Bopp's comparative methodology in defense of his approach to language classification.[25]

The basic data are presented in Map 1, which shows the approximate geographic distribution, in the year 1500, of twelve language phyla into which virtually all of the world's many thousands of languages surviving at that time can be classified.[26] These twelve groups represent (for those linguists who accept that large groupings of languages can feasibly be reconstructed) a rough summary of current knowledge. Of the twelve phyla, the Dene-Caucasian (including Sino-

---

[24] In effect, Joseph Greenberg classified seven of the twelve known phyla of the world's languages. Greenberg's pioneering classifications of major language groups of the Old World tropics are summarized in *The Languages of Africa* (Bloomington: Indiana University, 1966); and "The Indo-Pacific Hypothesis," in *Current Trends in Linguistics,* vol. 8, *Linguistics in Oceania,* ed. Thomas A. Sebeok (The Hague: Mouton, 1971), pp. 807–871. The basic analyses of linguistic classification for northern Eurasia and the Americas are Joseph Greenberg, *Language in the Americas* (Stanford, Calif.: Stanford University Press, 1987), and Greenberg, *Indo-European and Its Closest Relatives: The Eurasiatic Language Family,* 2 vols. (Stanford, Calif.: Stanford University Press, 2000–2002). A more accessible summary, including the argument for an early migration associated with Dene-Caucasian languages, may be found in Ruhlen, *Origin of Language.*

[25] Joseph H. Greenberg, Christy G. Turner II, and Stephen L. Zegura, "The Settlement of the Americas: A Comparison of the Linguistic, Dental, and Genetic Evidence," *Current Anthropology* 27 (1986): 477–497 (see especially p. 493); Bopp, *Vergleichende Grammatik.* See also Joseph H. Greenberg, *Essays in Linguistics* (Chicago: University of Chicago Press, 1957), p. 43.

[26] A linguistic phylum is a maximal group of languages demonstrated to be related to each other through descent from a common ancestral language. It is roughly parallel in the logic of its construction to a biological phylum.

Tibetan) and Eurasiatic language groups had the largest number of speakers; the Niger-Congo and Austric groups had the largest number of languages.[27]

Greenberg's classifications—of four African language phyla, plus Amerind, Indo-Pacific, and Eurasiatic—each encountered substantial debate, though a firm consensus has developed on modified versions of his four African phyla.[28] Overall, the full range of Greenberg's classificatory work reveals the consistency in the pattern of ancestry and differentiation in human languages.[29] Details of classification within phyla are likely to change with further research, and links among phyla are likely to be discovered, but the overall classification of human languages will almost certainly remain within the boundaries summarized here. Following the tradition of Indo-Europeanists, Greenberg used a tree-model approach in structuring his proposed language groups. Working with existing languages to identify their relationship through the closeness of their grammatical patterns and the proportion of their cognate words, he assembled languages with a common ancestor, and then assembled the ancestral languages to postulate a more distant ancestor, and so forth. Greenberg modeled his proposed trees on the implicit assumption of a simultaneous separation of daughters from parent languages in each generation; subsequent scholars in African languages have modified this model with closer analysis and have proposed the sequence of separations within each "generation."[30]

Geographic Homeland: The "Least Moves" Principle

Identifying the homeland for a dispersed population is a key task in analysis of early migrations. The full determination and verification of

---

[27] The map has been drawn based on language distribution in 1500, because migration since then has changed the pattern of language distribution greatly.

[28] A significant group of linguists, often known as "structuralists," decline to recognize phyla or subphyla unless the ancestral language has been reconstructed, and unless a full map of regular sound changes among languages has been established.

[29] Merritt Ruhlen, a former student of Greenberg at Stanford, continues the work the two of them began on hypothesizing that there was an original human language and trying to identify elements of it. Ruhlen, *Origin of Language*.

[30] Examples of the moderate changes in classification of African languages since the work of Greenberg are the recognition of Omotic as a major group within Afroasiatic and the recognition of Ijo and Dogon as major groups within Niger-Congo. For examples of recently drawn language trees showing sequential separation of groups, see Bernd Heine and Derek Nurse, eds., *African Languages: An Introduction* (Cambridge: Cambridge University Press, 2000) pp. 18, 274, 289–293; for comparison, see Greenberg, *Languages of Africa*, pp. 8–9, 46, 49, 85–86, 130, 177.

the points of origin and the paths of movement of populations and their languages are complex and require the assembly of expertise drawn from many fields.[31] The most important single element in identifying the homelands from which languages spread, however, is the mapping of language subgroups. For this reason, through simple application of the "least moves" principle, a layperson can make quick and remarkably valuable estimates of the points of origin and direction of migration of past populations. Only two sorts of information are required, and both of these are provided by linguists in many cases: (1) a genetic classification of related languages, distinguishing the broader groupings of languages for earlier times from the narrower groupings of more closely related languages for more recent times; and (2) a map showing the locations of populations speaking these same languages and groups of languages.[32]

Let us take the example of speakers of the Portuguese language. Where was the homeland from which their ancestors came? Linguists have classified Portuguese as a Romance language, and have identified the major other Romance languages as Spanish, French, Italian, and Romanian. To estimate the homeland for the ancestor to Romance languages: (1) on the map, locate and mark the point that is the geographical center for each Romance language; and (2) locate the point that minimizes the total distance from it to each of these points. Thus, if we placed points at the geographic center of Portugal, Spain, France, Italy, and Romania on Map 2, then our estimate of the point of origin for the whole language group would be somewhere in northwestern Italy. This is the point from which the total length of the lines drawn to each of the language centers would be minimized. In fact, it gives a pretty good representation of the fact that Latin-speaking Romans, especially from the northern half of Italy, colonized all of these areas

---

[31] On Indo-European expansion see Colin Renfrew, *Archaeology and Language: The Puzzle of Indo-European Origins* (Cambridge: Cambridge University Press, 1987); on Bantu expansion see Christopher Ehret, "Bantu Expansions"; on Austronesian expansion see Bellwood, *Indo-Malaysian Archipelago,* pp. 96–127. Bellwood, an archaeologist, relied significantly on the work of Isidore Dyen and other linguists in developing his interpretation.

[32] For an early and detailed formulation of this identification of linguistic homelands through a "least moves" approach, see Isidore Dyen, "Language Distribution and Migration Theory," *Language* 32 (1956): 611–626; reprinted in Dyen, *Linguistic Subgrouping and Lexicostatistics* (The Hague: Mouton, 1975), pp. 50–74. Dyen developed ideas earlier suggested in 1916 by Edward Sapir in analysis of North American languages and applied them especially to Austronesian languages.
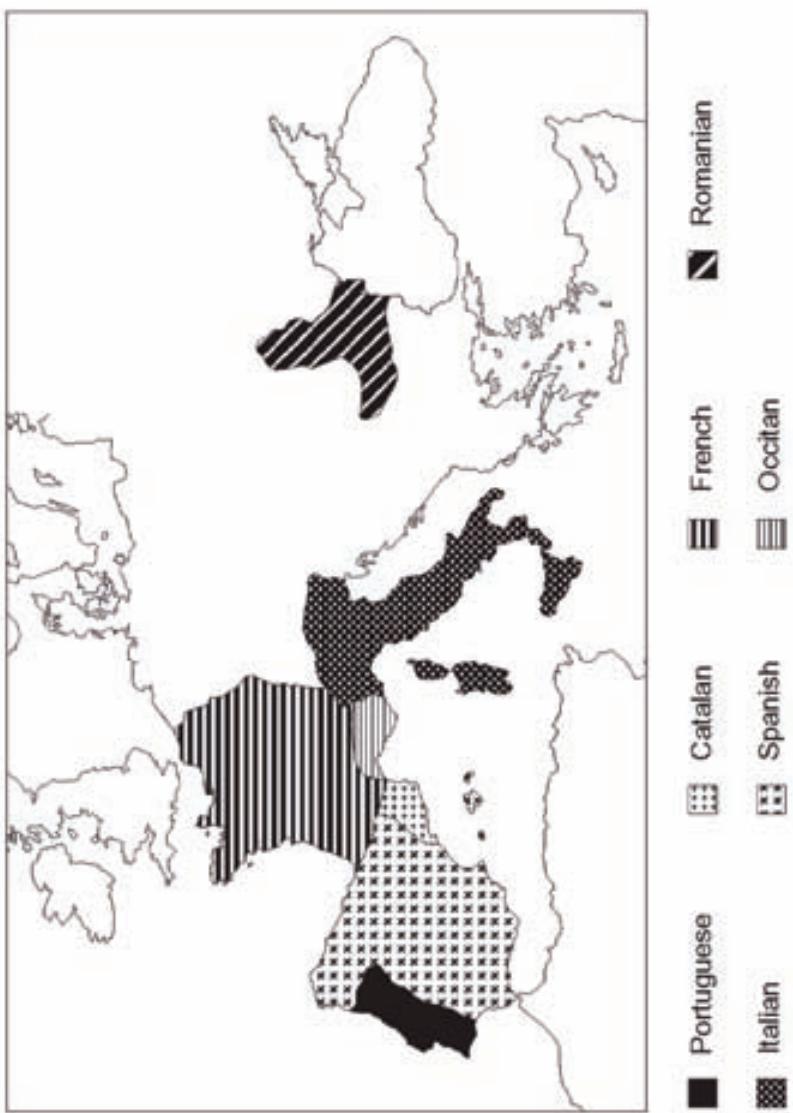
MAP 2. Romance language homeland.

more than 2,000 years ago and launched the process leading to the languages of today.

This statement of the principle of least moves is highly simplified and in this presentation has left out a great deal of available information. For instance, there were many more Romance languages than the five I listed, and the others were clustered in the area around the homeland.[33] Further, the center of origin for Portuguese (or any of the other languages) can be located more precisely by accounting for the various dialects within the language; there are huge populations speaking Portuguese, Spanish, and French outside of Europe (though these are known to have grown up in recent centuries), and so forth. Nonetheless, this simple least-moves approach enables the lay reader to participate actively in the interpretation of past human migrations through study of evidence on language classification.[34]

We can trace the ancestry of Portuguese language to an earlier stage, since Romance languages are one of the categories in the Indo-European language family. Map 3 shows the distribution of Romance and the other ten known subgroups of Indo-European languages.[35] As shown on Map 3, the least-moves estimate for the Indo-European homeland is near the shores of the Black Sea.[36] The language evidence does not lead to a straightforward estimate of the time of Indo-European origins. In fact, linguists and archaeologists have debated fiercely the question of the location of the Indo-European homeland and also the timing of Indo-European origins.[37] But our simple least-moves esti-

---

[33] Other Romance languages include Provençal of southern France, Calatan of northeastern Spain, Corsican, Sardinian, and other small groups in northern Italy.

[34] In classroom exercises with Nilo-Saharan, Afroasiatic, and Niger-Congo phyla, I have created these simple estimates of the homeland assuming that all the major subgroups diverged at once and compared them with more complex estimates of the homeland accounting for the differing times at which subgroups emerged. The two estimates of each homeland were very close to each other, thus confirming that the simple least-moves estimate is a valuable technique.

[35] Two of the groups, Tocharian and Anatolian, are no longer spoken but are known from written records.

[36] As an assist in locating the least-moves center, find the latitude at which half of the groups are centered to north and south, and the longitude at which half of the groups are centered to east and west. The intersection of these two lines is very close to the least-moves center.

[37] Mallory proposes a homeland at the northeast edge of the Black Sea, Renfrew proposes Anatolia (south of the Black Sea), and Marija Gimbutas argues for the northwest coast of the Black Sea. Mallory, *In Search of the Indo-Europeans*, p. 262; Renfrew, *Archaeology and Language*, p. 266; Gimbutas, *The Civilization of the Goddess* (San Francisco: Harper, 1991), pp. 352–353. I argue that the origins of this group must go back before the development of agriculture, to at least 15,000 years ago.
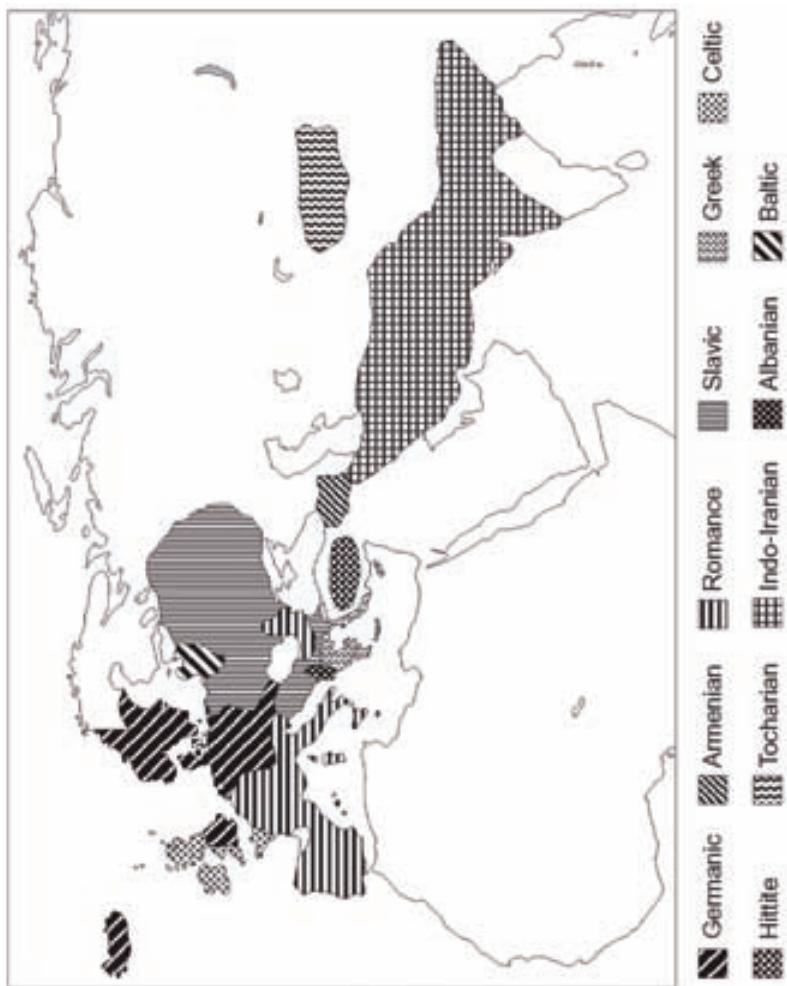
MAP 3. Indo-European languages.

mate is sufficient to get us into the thick of the argument—it is precisely one of the main areas proposed by scholars as the Indo-European homeland and is definitely within a thousand kilometers of any of the candidates for the homeland. In short, through this method ancient homelands can be picked out of contemporary language distributions with some confidence.

Continuing back into the deeper past, we may ask whether Indo-European was part of a broader and earlier grouping of languages. Indeed, the answer is yes, and the most authoritative description is that of Joseph Greenberg, who identified the super-family of languages he labeled as Eurasiatic. The Eurasiatic super-family comprises seven major families of languages of Eurasia and the Arctic, of which the Indo-European languages are but one. As I will show, the least-moves estimate of the Eurasiatic homeland is near the Pacific coast of north Asia.

World-Historical Linkage of Data

For a world-historical approach to the issue of early human migration, the analyst should pose the issue at broad scope (preferably planetary), consider both long-term and short-term relationships, incorporate data from a wide range of disciplines, and utilize a range of methods. The geneticist L. L. Cavalli-Sforza pioneered the linkage of different sorts of data—genetic, paleontological, and linguistic—in projecting the spread and differentiation of human populations. He has published "tree" diagrams showing estimates of the genetic distance of human populations of today, has compared them to tree diagrams of language groups of human populations of today, and has included measures of bodily characteristics of human populations.[38]

While the combination of many sorts of data enables a more comprehensive analysis, it also has its difficulties. Each type of data has its own logic. For language, genetic composition, and physical type, we assume that present data indicate the remnants of earlier communities.[39] But the definition of "earlier community" is different for each

---

[38] Cavalli-Sforza, *Human Genes,* p. 99. The genetic data included recent analysis of DNA but especially earlier analysis of blood types and other protein data; measures of bodily characteristics included skin and eye color, height and skull measurements; language data were drawn from Greenberg. Links among these data were proposed by Cavalli-Sforza and his associates.

[39] As Cavalli-Sforza has noted, there do not now exist ancestral populations from which others have descended, either for language or genetics. Since mutations occur in all

type of data, so the tree diagrams of genetic, linguistic, and skeletal change in humans have slightly different meanings. Genetic descent is sexual, so that each offspring has two ancestors at the level of each generation; further, one's genetic composition is set at conception. Linguistic descent is asexual, so that each offspring has only one ancestor in each generation; on the other hand, an individual can change language by an act of the will. Body type is inherited biologically, but is also subject to environmental pressures after birth. The tree models of these three types of descent convey certain common characteristics. When they can be mapped, it is generally the case that the areas of greatest diversity (among groups that have some relationship) correspond to regions where populations have differentiated through long residence in a single place; these are typically a homeland from which dispersal took place.[40]

But each sort of tree has its own patterns, and a tree model is not sufficient to capture all elements of variation in the evidence it summarizes.[41] Because of the single-ancestor characteristic of the linguistic "tree model," language gives more evidence on the path of migration than does genetics, because it allows for fewer possibilities among ancestors. Quantitative measurement of linguistic differences is difficult, however, because of the substantial qualitative differences between one aspect of language and another. Genetic variation is more readily susceptible to quantitative estimates, to the degree that it is comparison of base pairs on the genome from one population to another. For these reasons, percentages of genetic variation cannot be compared directly with percentages of linguistic variation.

DNA, and since changes in vocabulary and syntax occur in all languages, all the populations and languages we encounter now are modern. In genetics, it is now possible to determine the degree of relationship between the composition of any two populations. In language, within phyla (but not yet between phyla) it is possible to determine the degree of relationship of any two populations.

[40] For Romance languages, the diversity of languages is greatest along the Mediterranean coast from Italy to Spain. For Indo-European languages, the diversity is greatest in the area including Greek, Albanian, Hittite, and the southern range of Slavic.

[41] Cladistics is a type of analysis, developed especially among biologists, for constructing analytical trees to reflect patterns of descent and evolution. In particular, cladistics has shown that multiple trees may fit a single set of data in genetic or linguistic descent. (The "wave model" for languages reflects an attempt to account for the types of influence striking all languages at the same time—especially borrowing of terms resulting from innovations.) Cladistic models for languages, meanwhile, may differ from those for genetic descent because languages have no equivalent to bisexualism. Ian J. Kitching, *Cladistics: The Theory and Practice of Parsimony Analysis* (Oxford: Oxford University Press, 1998).

Two more types of data play a central role in this analysis. First is the study of climate—the rise and fall of temperature and precipitation, habitability of various world regions, and sea level. Recently developed data, presented especially as changing sea levels, play a key role in the interpretation of migration paths. Second is archaeological studies, which provide evidence on lifestyle and environment for human populations. The combination of these two types of evidence, I will argue, emphasizes the importance of life at water's edge and the use of watercraft at all stages of human history. As human communities grew and spread, they were confronted repeatedly by a choice: concentrate at water's edge or range across open grassland. Earlier hominids had faced this choice and tended to stay close to waterways.[42] Early communities of *Homo sapiens,* at each stage of developing technologies and exploring new ecologies, found new ways to benefit from life in the grasslands and also from life at water's edge. Studies of human evolution have long tended to emphasize hunting and the grasslands. To achieve some balance, I want to emphasize the continuing importance of rivers, lakes, and the ocean among early *Homo sapiens.* Gatherers found a rich variety of plant and animal life along the seashore, along rivers, and at lakeside. Humans are likely to have been swimmers from the first and to have developed rafts and boats. Though the evidence is indirect, maritime archaeologists have shown the logic of the construction of the first watercraft. Logs might serve as rafts, but, more practically, the gathering and bundling of reeds—available at water's edge throughout the tropics—provided materials for lightweight and maneuverable craft.[43] The balance of human reliance on the produce of the soil and the produce of the waters has been adjusted in each new region and with each new technology. Here I argue that this pattern of reliance on the waters and watercraft can be projected back to the earliest days of human migration and that it fits with patterns revealed in archaeology, genetics, and historical linguistics.

These principles are now applied to the data on language distribution and other data to yield a provisional synthesis, an interpretation of four stages in the migration and differentiation of human populations.

---

[42] On overcoming the oversimplified model of man the hunter, focusing on foraging, and noting consistent linkage of humans to lakes and streams and littorals, see Stringer and McKie, *African Exodus,* pp. 29–33.

[43] Paul Johnstone, *The Sea-Craft of Prehistory* (London: Routledge and Kegan Paul, 1980), pp. 7–16.

PEOPLING THE OLD WORLD TROPICS: 100,000–40,000 B.P.

In their first migration out of Africa, modern humans moved into the region east of the Mediterranean as early as 100,000 B.P. The archaeological record shows that there were alternations of modern humans and Neanderthals in the region, even in the occupation of individual caves, and that Neanderthals continued to live in the area until about 40,000 B.P.[44] For modern *Homo sapiens,* this was an early but limited movement out of Africa, which left no linguistic remains and for which the population did not become sizeable. Desiccation of the Sahara in the period from 90,000 B.P. suggests reasons why this northern region might not have remained hospitable to humans.

Within Africa, meanwhile, substantial migrations took place, as indicated in the patterns of language groups. African populations moved from being centered in the savannas of eastern and southern Africa, where their hominid ancestors had always been most numerous, to being centered in the east-west belt of the northern savanna between Ethiopia in the east and Senegal in the west. Four great language groups are based in the African continent and reflect the placement and movement of people for tens of thousands of years. I believe that recent language distributions can be projected back with sufficient confidence to show that as of about 80,000 B.P., the Khoisan languages were based in the savanna areas of eastern and southern Africa, where humans had first evolved. The Nilo-Saharan languages were based in the middle Nile Valley, and the Afroasiatic languages were based in a nearby region of the middle Nile Valley. The Niger-Congo languages were centered to the west of the last two, and included groupings both east and west of Lake Chad. All of these were areas where hominids had lived before, but the regional emphasis had now moved from eastern and southern Africa to the grasslands and waterways of the northern savanna. In addition, and in continuity with earlier hominid patterns, we must assume that humans populated the shores of the Indian Ocean and the Red Sea.

The next move out of Africa, along the Indian Ocean littoral, was to be of a far larger scale. In this colonization of new lands, *Homo sapiens* migrated east along the tropical lands bordering the Indian Ocean. This tropical migration appears to have stemmed from the

---

[44] Brian M. Fagan, *Journey from Eden: The Peopling of Our World* (New York: Thames and Hudson, 1990), pp. 90–100; Stringer and McKie, *African Exodus,* pp. 76–80. Results of newer archaeological work are expected.

development of new technologies and social systems, allowing humans to occupy a steadily wider range of ecologies. Then one stream of migrants, relying on water's-edge technology, including the use of boats, crossed the narrow waterway between Ethiopia and Yemen (less than 20 kilometers at the time) and expanded eastward. These migrants colonized the Indian Ocean coast with relative ease, and from that vantage point gradually spread to the interior of islands and mainland areas. The preexisting populations of *Homo erectus* provided little resistance to the migrants and may not have been numerous in the coastal zones along which the settlers moved. There was one significant change in ecology in the course of this eastward transit: east of the Ganges River, thick forest—populated especially with bamboo—covered the lands right up to the coast.

Perhaps the most remarkable step of this migration was the movement across what is now the Indonesian archipelago to the lands that are now New Guinea and Australia. Indonesia was then a subcontinent, but the only way to get to New Guinea and Australia was to cross open stretches of ocean of at least 100 kilometers. Archaeologists have shown, through dating of human remains and artifacts in Australia, that humans had achieved that task by about 50,000 B.P.[45]

An essential part of the information for creating this interpretation comes from the work of geologists. Their work has demonstrated that the earth went through a long cooling phase between about 130,000 and 20,000 B.P., after which it warmed rapidly. During this long era of cooling the polar ice pack grew, ocean levels declined, and the climate became steadily drier because so much water was in frozen form. Figure 1 shows the summary results of recent research, using measurements from the island of Barbados to estimate the rise and fall in sea level over that time. It suggests that in the time from 80,000 to 50,000 B.P., sea level was from 60 to 80 meters lower than it is today. Thus the migrants who first worked their way eastward along the tropical coast were on a coastline that has since been inundated by the rise in waters at the end of the Ice Age. Those lower sea levels revealed an expanded Southeast Asian subcontinent that geologists have called Sunda. The lower waters also linked Australia and New Guinea into a continent that geologists call Sahul.

---

[45] The date for the human remains at Lake Mungo, New South Wales, has now been reduced to 40,000 B.P., but it is assumed that the first human arrivals reached western Australia (at the other end of the continent) about 10,000 years earlier. James M. Bowler et al., "New Ages for Human Occupation and Climatic Change at Lake Mungo, Australia," *Nature* 421 (2003): 837–840.
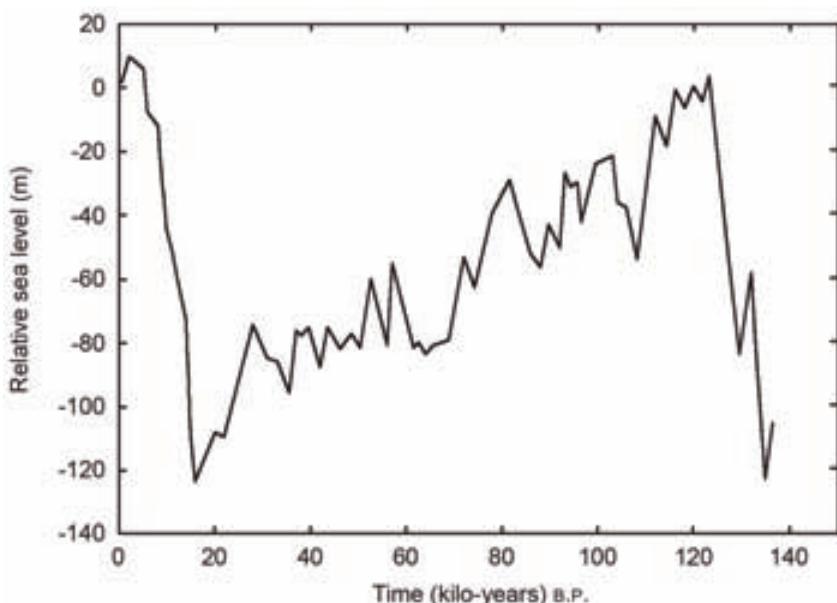
FIGURE 1. Global changes in sea levels from 2000 to 140,000 B.P.
Source: Edouard Bard, Bruno Hamelin, and Richard G. Fairbanks, "U-Th Ages Obtained by Mass Spectrometry in Corals from Barbados: Sea Level During the Past 130,000 Years," *Nature* 346 (1990): 456–458.

Even with the maximal amount of land revealed by low levels of the ocean, the human migration eastward entailed the task of island-hopping across distances of up to 100 kilometers by boat. The boats may have been reed craft or bamboo rafts. The crossing was made not once but several times, according to genetic evidence showing differences within the populations of Australia and New Guinea.[46] After making this crossing, the settlers were able to spread throughout Sahul.

I think that this idea of a water's-edge migration from Africa to Australia, within the period from 80,000 to 50,000 B.P., is more than plausible. If a technology were developed that enabled humans to prosper at the boundary of tropical ocean and land of somewhat varying rainfall, there were thousands of kilometers of coastline of similar ecology from the Horn of Africa to Sahul. Vegetable and crustacean

---

[46] Fagan, *Journey from Eden,* pp. 129–138.

nourishment from this neighborhood provided the basis of subsistence, perhaps along with fish. Boats were a necessary part of life.[47] The result was that Indo-Pacific and Australian language groups, and probably the ancestors of Sino-Tibetan, Austric, and Dravidian groups, were set in place by 50,000 B.P. Map 4, summarizing existing knowledge of existing tropical language groups and their homelands, gives a clear overview of human occupation of the tropics.

What were the languages of those who left Africa and headed east along the coast? They could have been in any of the four language groups of Africa today, or of yet another language group that has since disappeared. Of the current African language groups, I argue that the Nilo-Saharan languages are the most likely source of the eastward migrants. I base this estimate on the geographical distribution of Nilo-Saharan languages, for which the homeland would appear to have been within reach of the Red Sea coast, and on the significant emphasis of Nilo-Saharan speakers in more recent times on what Christopher Ehret has called an "aquatic tradition."[48] As a second candidate for the origin of the eastward migrants, I suggest the Afroasiatic languages: these too appear to have a homeland along the frontier of modern Ethiopia and Sudan and were geographically well placed to send migrants eastward.

Two other groups are less likely candidates as the source of colonists in Asia, but cannot be excluded. For the Niger-Congo languages, their homeland appears to be rather far to the west (at least as far as Kordofan in western Sudan), but many of the Niger-Congo speakers in recent times have emphasized life at water's edge. For the Khoisan languages, the Khoisan speakers of today live rather far from the East African coast and have very little involvement in boating. (On the other hand, genetic comparisons suggest Khoisan-speakers are closer to Asians than other African groups, though this might reflect recent rather than early connections.)[49]

If the Nilo-Saharan languages were the source of the eastward migrants, then one would expect ultimately to find all the tropical Asian and Oceanic language groups to be related to Nilo-Saharan, presumably as daughter language groups. These include Dravidian, Sino-Tibetan (or Dene-Caucasian), Austric, Indo-Pacific, and Australian.

---

[47] Brian Fagan has assumed that humans developed boats in Southeast Asia, as a result of their encounter with bamboo. He assumes a journey by land from Africa to Sahul—see Fagan, *Journey from Eden*, pp. 121–138.

[48] Ehret, *Civilizations of Africa*, pp. 68–75.

[49] Cavalli-Sforza, *Human Genes*, pp. 175–176.
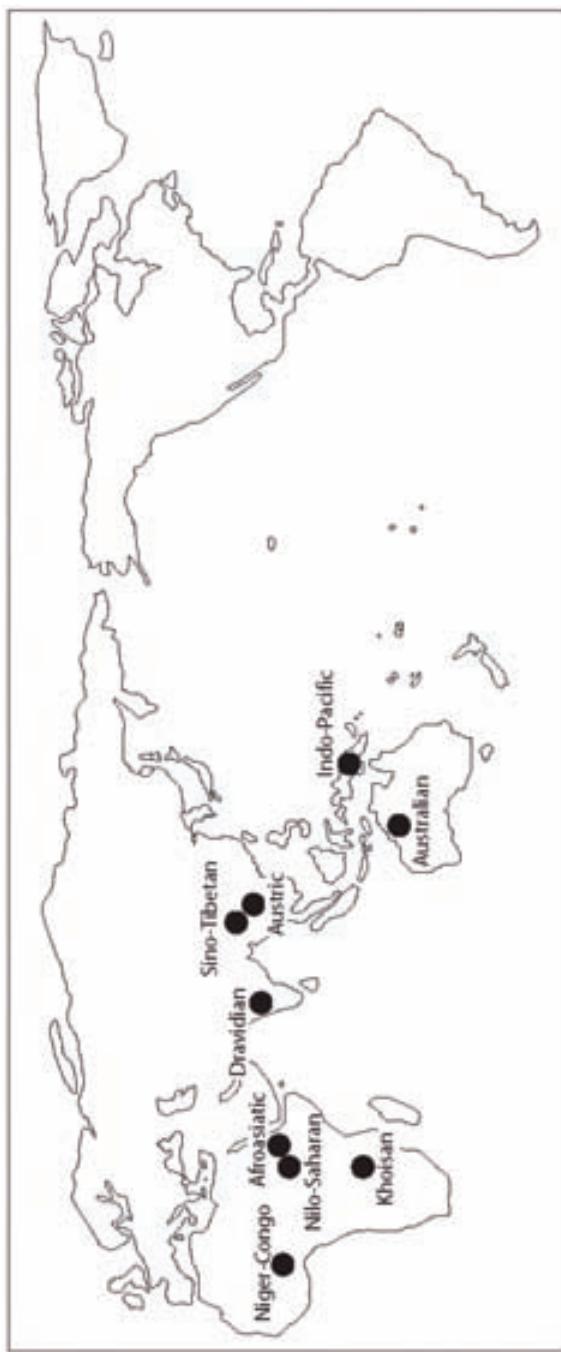
MAP 4. Tropical language families.

The continuing work of language classification is almost sure to clarify these linkages.[50]

Peopling Northern and American Regions: 40,000−15,000 b.p.

By 50,000 b.p. humans had become a set of communities expanding their activities along coastal and inland areas of the tropics from West Africa to the South Pacific. The lifestyle of these humans likely depended on the gathering of animal and vegetable materials from water's edge, from oceans, rivers, and lakes. It appears, however, that this technology was not adequate for life in the cooler or drier climates of regions north of the tropics. Humans remained restricted to the tropics until they developed techniques for living under different ecological conditions.[51]

Occupation of temperate regions required development of a technology based on gathering of different sorts of vegetable materials and associated with more effective hunting of large animals. The new technology included better spears and (later) throwing sticks, techniques for isolating large animals, and sewing to make clothing for cold weather as well as to sew hides around wooden frameworks for boats. These techniques, once developed, allowed for rapid occupation of the northern two-thirds of Eurasia. Once gaining the ability to live comfortably in temperate zones, whatever their point of entry from the tropics, humans spread easily to occupy the lands and water's edge from the Atlantic to the Pacific.

The explanation of the human movement eastward from Africa along the fringe of the Indian Ocean to the Sahul continent, as presented above, is a rather straightforward analysis, once its basic presumptions are accepted. The evidence of archaeology and genetics, confirmed by that of language, gives a consistent picture of the tropical expansion of *Homo sapiens*.

---

[50] As a skeptical note on this vision of human occupation of the tropics, I should note that the islands of Madagascar and the Comoros, off the southeast coast of Africa, were not occupied by humans as part of the initial human expansion, and may not have been settled by humans until some 3,000 years ago. Madagascar and the Comoros, however, each lie some 400 kilometers from the African coast, a far greater distance than those crossed by mariners crossing from Africa to Arabia or from Sunda to Sahul.

[51] Of particular importance is the question of whether, in this time from 90,000 to 40,000 years ago, the ecology of Egypt, Sinai, and Palestine was sufficiently close to that of the African tropics to make a landward migration out of Africa as feasible as the movement

Reconstructing the human occupation of northern Eurasia and the Americas, in contrast, is a complex problem. It involves the sorting out of several possible routes of migration and requires resolving conflicting evidence on genetics, archaeology, and language. The overall scenario I propose is as follows. As late as 40,000 B.P., *Homo sapiens* remained restricted to the tropical areas of Africa, Asia, and Oceania.[52] By 30,000 B.P., *Homo sapiens* had expanded to occupy all of Eurasia, displacing previous hominids (*Homo erectus* in the eastern zones and Neanderthals in the western zones), and had established communities in the Americas. The archaeological record for widely dispersed regions of temperate Eurasia shows dates for remains of modern humans as far back as 30,000 and 40,000 B.P. The dates for Europe and the Middle East are more numerous and somewhat earlier than for central and eastern Asia, but the central and eastern regions have not been as thoroughly studied.[53]

In the analysis to come, I contrast regions of linguistic commonality with regions of linguistic diversity. The most impressive region of linguistic unity was that of the Amerind languages, which expanded without interruption to occupy all of South America and most of North America (though they have since lost out significantly to Indo-European languages). A close second in linguistic unity is Eurasia, where the single, large Eurasiatic language family is spoken today from the Atlantic to the Pacific and even to the Indian Ocean and parts of North America.[54] A third pattern of linguistic unity, characterized by a wide scattering of related groups, is the Dene-Caucasian languages.

In contrast, I want to point out four major centers of linguistic diversity: regions where the existence of distinct but related languages in a small area gives the impression that these were regions from which migrants departed. (The reader may consult Map 1 to locate these regions.) One such region of diversity is the Caucasus. There in the low mountains between the Black and Caspian Seas, we find the North Caucasian languages (including modern Chechen) and Kartvelian languages (including modern Georgian)—each related only distantly to

---

across to South Arabia. My assumption here is that this northern route was too different to be attractive to humans at the time.

[52] The exception to this pattern is the presence of modern *Homo sapiens* in the Eastern Mediterranean for a period about 100,000 years ago. Fagan, *Journey from Eden,* pp. 90–100; Stringer and McKie, *African Exodus,* pp. 77–80. Further archaeological results are expected from this region.

[53] Fagan, *Journey from Eden,* pp. 141–198.

[54] In land area, the Amerind languages dominated some 40 million square kilometers in the Americas, and the Eurasiatic languages dominated roughly 20 million square kilometers.
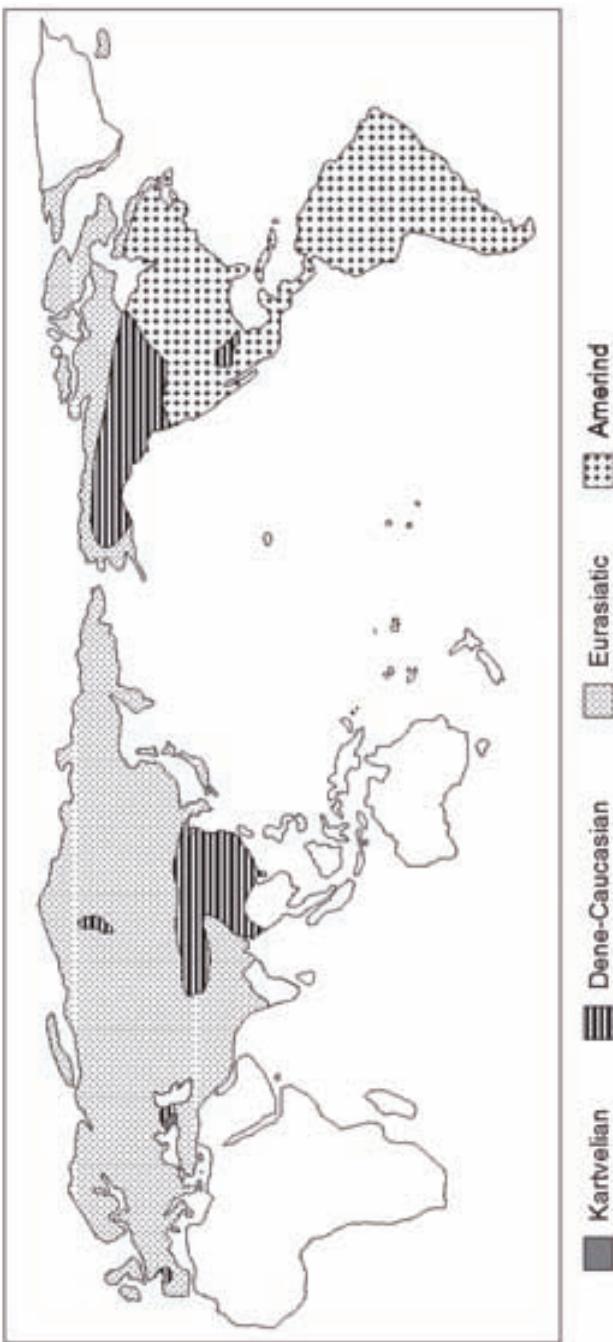
Map 5. Language families of temperate and American zones.

Kartvelian   Dene-Caucasian   Eurasiatic   Amerind

other languages—and representatives of the Indo-European and Altaic families of Eurasiatic languages. The Caucasus has long received attention as a possible center of human dispersion, and its significance as a center of linguistic diversity is striking.[55]

A second region of linguistic diversity has received far less attention. Within the great linguistic commonality of the Eurasiatic languages, the greatest diversity of languages is to be found on the northeast Asian coast, where four of the seven subgroups of Eurasiatic languages appear to have their homelands.[56] (See Map 6.) The Gilyak and Chukotian language groups have not been studied in great detail, and Greenberg's classification of Korean, Japanese, and Ainu as a single group is recent; deeper linguistic research on this region is surely a priority. The Altaic languages exhibit the greatest diversity in the eastern part of their range, suggesting that the group emerged in the east, near the Pacific. A least-moves estimate of the homeland for Eurasiatic as a whole places it near the Pacific coast and suggests that the Eurasian grasslands may have been settled from the east rather than from the west. The Indo-European languages, while now the largest and most populous group within the Eurasiatic family, are also the most far-flung from the apparent homeland. They may have begun, therefore, as western outliers among Eurasiatic speakers.

A third region of linguistic diversity goes further back in time. All four of the major subgroups of the Sino-Tibetan languages are represented in Yunnan, in today's southwest China, along the major rivers of southeast Asia.[57] In much the same area, and only slightly downriver, is the homeland of the Austric languages (a phylum that is commonly discussed in terms of its four constituent subgroups: Austroasiatic, Miao-Yao, Dai, and Austronesian).[58] This double-barreled center

---

[55] The use of "Caucasian" as a racial term stems from an eighteenth-century argument that the Caucasus was the home of a pure, "Caucasian" race, and from nineteenth-century assertions that the same region was the homeland for the Indo-European languages. Since geneticists now argue that the characteristics of "race" are genetically superficial rather than of any depth, the relevance of the Caucasus for racial analysis has become dubious; however, the relevance of the Caucasus for its linguistic diversity remains significant. On Blumenbach's 1776 coining of the term "Caucasian," see Emmanuel Chukwudi Eze, ed., *Race and the Enlightenment: A Reader* (Oxford: Blackwell, 1997), p. 86.

[56] Greenberg, *Indo-European and Its Closest Relatives,* vol. 1, *Grammar,* pp. 1–23.

[57] R. L. Rankin, "Sino-Tibetan Languages," in Asher and Simpson, *Encyclopedia of Language,* 7: 3,951–3,953; and Ruhlen, *Guide to the World's Languages,* 1: 141–148.

[58] Paul Benedict has led in denying that Austric is a single phylum, but I follow Ruhlen in treating it as one. Indeed, given the proximity of the homelands of the subgroups of Austric and the subgroups of Sino-Tibetan, I think it should be suggested that a linguistic relationship and a shared migratory history may ultimately be unraveled for all the groups speaking Austric and Dene-Caucasion (including Sino-Tibetan) languages. Paul K.

Eurasiatic

Indo-European   Uralic   Chukotian   Gilyak

Altaic   Japanese-Korean-Ainu   Eskimo-Aleut
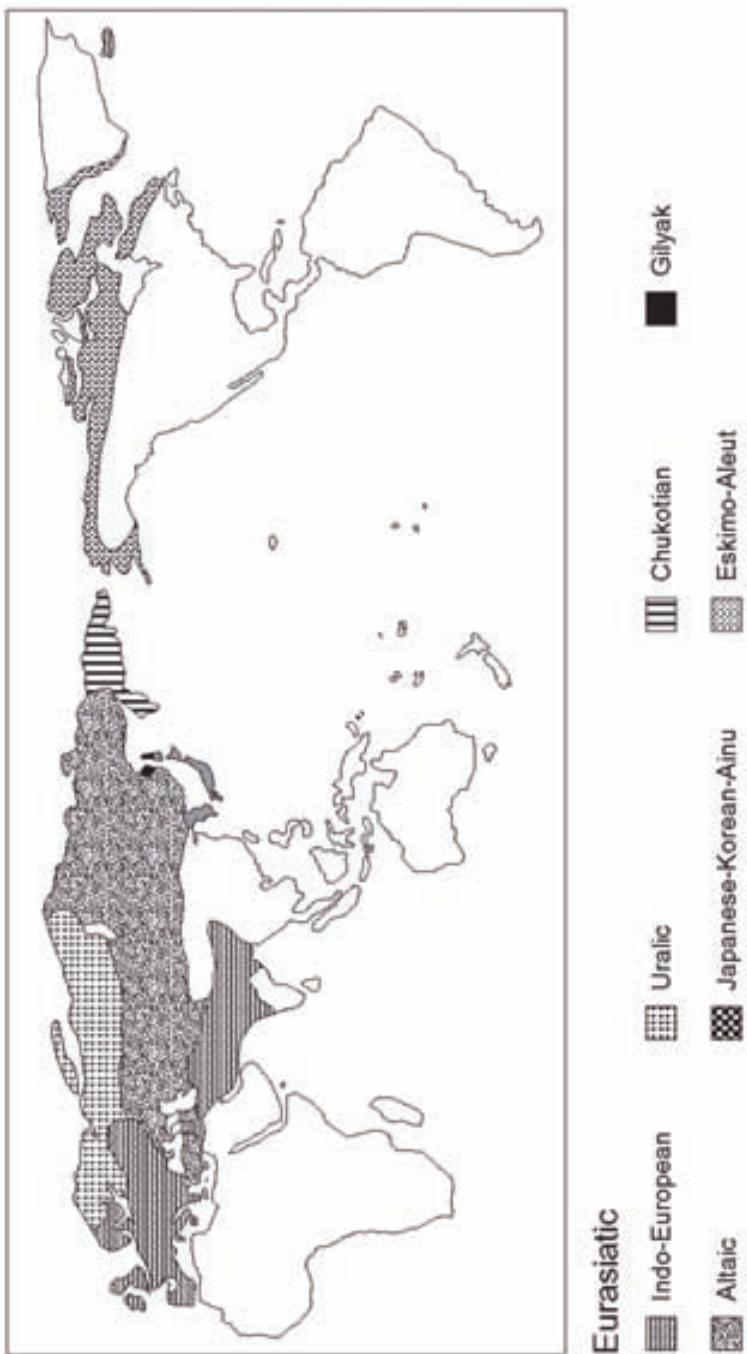
MAP 6. Eurasiatic language family.

of tropical linguistic diversity may have been a source of migrations to the north and in other directions.

The fourth region of linguistic diversity goes even further back in time: the middle Nile Valley, where Afroasiatic and Nilo-Saharan language groups have their homeland and where a small but important group of Niger-Congo languages is located just to the west.[59] The middle Nile was arguably the region that started the whole process of expansion to the east about 80,000 B.P.; in addition, it may also have been a source of expansion to the north in later times.

The archaeological record shows *Homo sapiens* as inhabitants of temperate Eurasian regions from Atlantic to Pacific beginning about 40,000 B.P.—somewhat later for the arctic fringe of Eurasia. There had been a pause, it appears, between the occupation of the tropics in the years up to 50,000 B.P. and the movement into temperate Eurasia. Some sort of breakthrough in technology and perhaps social organization was needed to enable significant numbers of humans to move north.

With this introduction, let us turn to an investigation of the Eurasiatic languages, the language phylum now occupying the great majority of the territory of Eurasia. The map of Eurasiatic languages, as proposed by Joseph Greenberg, covers such an immense area that one is readily tempted to view it as reflecting a rapid move to occupy all of northern Eurasia, stemming from a single region in the tropics. This is a first approximation to the argument that I will make, though I will also add a number of complications to the story. The identification of this phylum (sometimes called a super-family) of languages is a substantial accomplishment: it is a major advance over the previous century's emphasis on Indo-European languages, now shown to be one of seven constituent groups of Eurasiatic. The history of the Eurasiatic language group goes back much further and includes a far wider range of populations than does its Indo-European subgroup. Linguists have suspected this possibility for some time; Greenberg's analysis of Eurasiatic paralleled the work of a series of European-based scholars (working particularly in Russia) who developed the term "Nostratic" to refer to the combination of Indo-European, Altaic, Uralic, and other lan-

---

Benedict, "Austric: An 'Extinct' Proto-Language," in *Austroasiatic Languages: Essays in Honor of H. L. Shorto,* ed. J. H. C. Davidson (London: School of Oriental and African Studies, University of London, 1991); Ruhlen, *Guide to the World's Languages,* 1: 148–158.

[59] Some linguists have raised the possibility that Niger-Congo might be a branch of Nilo-Saharan. Further, based on proximity of homelands, one may ask whether Nilo-Saharan and Afroasiatic might be descended from some earlier common language.

guage groups. While there remain differences on the proposed linkage of Afroasiatic, Dravidian, and Kartvelian to Nostratic, there is great similarity between Aharon Dolgopolsky's vision of Nostratic and Greenberg's vision of Eurasiatic.[60] Thus we have significant agreement on the composition of a language family covering most of Eurasia.

The next stage in unraveling the puzzle of occupying the temperate regions is analyzing the languages of the Americas. Prior to his classification of Eurasiatic languages, Greenberg published in 1987 a classification of the languages of the Americas.[61] His identification of Amerind as a single family encompassing the great majority of American languages brought a stormy response from Americanist linguists who declined to accept the existence of this larger grouping of languages.[62] Important statements from each camp appeared as a result, and one must wait for the debate to run its course, but here I have unhesitatingly accepted Greenberg's classification because its patterns fit so well with those accepted for languages elsewhere in the world.

Greenberg argued that Amerind is a sister group to Eurasiatic. (See Map 7.) If he had seen Amerind as a daughter group, he would have classified it along with Eskimo-Aleut as a subgroup of Eurasiatic. This classification implies that Eurasiatic and Amerind are both descendants of some ancestral stock, one that linguists can presumably seek out. Thus, if Eurasiatic came into existence in about 40,000 B.P., perhaps among fishers and hunters of the northeast coast of Asia, then one is prompted to argue that Amerind arose at much the same time, among hunters and fishers of the same region who continued to move north and east. Amerind speakers moved across the Bering straits to the Americas, either on a land bridge during the Ice Age or by sea before it. Greenberg's own clear opinion was that Eurasiatic and Amerind both emerged between 15,000 and 11,000 B.P. among populations that occupied lands given up by receding glaciers.[63] On the other hand, genetic evidence, as summarized by Cavalli-Sforza, tends to support

---

[60] Aharon Dolgopolsky, *Nostratic Macrofamily and Linguistic Paleontology* (Cambridge: McDonald Institute for Archaeological Research, 1998); Greenberg, *Indo-European and Its Closest Relatives*, vol. 1, *Grammar*, p. 9.

[61] Greenberg, *Language in the Americas*. Greenberg first proposed the outlines of this classification some thirty years earlier, in a paper presented in 1956 and published as "The General Classification of Central and South American Languages," in *Men and Cultures: Selected Papers of the 5th International Congress of Anthropological and Ethnological Sciences, 1956*, ed. Anthony Wallace (Philadelphia: University of Pennsylvania Press, 1960).

[62] See the responses of Americanist linguists in Greenberg, Turner, and Zegura, "Settlement of the Americas," pp. 488–492.

[63] Greenberg, *Language in the Americas*, pp. 333, 335; Greenberg, *Indo-European and Its Closest Relatives*, vol. 2, *Lexicon*, pp. 2–3.

Amerind

▦ Chibchan-Paezan	▨ Andean	▥ Central Amerind
⊞ Ge-Pano-Carib	■ Equatorial	▦ Northern Amerind

MAP 7. Amerind language family.

the earlier date of about 35,000 B.P. for the settlement of the Americas and also for the occupation of temperate Eurasia.[64] I too accept the earlier period as the time for expansion of these languages, as it is consistent with the hypothesized expansion of Eurasiatic and with the evidence of genetic difference.

To these two large groupings of languages beyond the tropics we may add a third. Linguist John Bengtson has confirmed and expanded the case for a grouping he calls Dene-Caucasian.[65] (See Map 5.) He finds a family relationship among six sets of languages that are widely separated geographically: Sino-Tibetan, North Caucasian, Basque (in the Pyrenees of Spain and France), Burushaski (in Pakistan), Yeniseian (isolated languages in northeast Siberia), and the Na-Dene languages of North America. Three of these groups—Basque, North Caucasian, and Burushaski—can easily be seen as remnants of earlier populations that lost ground to expanding Eurasiatic-speaking groups. The Na-Dene group of North America, in contrast, clearly arrived in North America after the Amerind speakers and found its advance into the continent limited by the previously established populations.[66] Sino-Tibetan, meanwhile, is as much a tropical as a temperate language group, in that most of its subgroups are located in the subtropical highlands of the Southeast Asian river valleys.

The evidence for the Dene-Caucasian language family suggests that there have been at least two waves of advance by humans into the Eurasian temperate zone: first by Dene-Caucasian speakers and then by Eurasiatic speakers. To clarify this possibility, it is important to establish the place of the Sino-Tibetan languages in the larger Dene-Caucasian family. I have argued that Sino-Tibetan was one of the founding families left by the eastward-moving colonization of the tropics. Under this assumption, the other groups listed in Dene-Caucasian are in practice part of Sino-Tibetan. But if Sino-Tibetan is only part of a larger family, one may have to look beyond Southeast Asia for the

---

[64] This conclusion is based on comparison of genetic distance between speakers of Amerindian languages and populations of northeast Asia. Cavalli-Sforza, *Human Genes*, pp. 325–326; L. L. Cavalli-Sforza, A. Piazza, P. Menozzi, and J. Mountain, "Reconstruction of Human Evolution: Bringing Together Genetic, Archaeological, and Linguistic Data," *Proceedings of the National Academy of Sciences of the USA* 85 (1988): 6002–6006.

[65] John D. Bengtson, "Notes on Sino-Caucasian," in *Dene-Sino-Caucasian Languages*, ed. Vitaly Shevoroshkin (Bochum, Germany: Brockmeyer, 1991), pp. 67–129.

[66] Bengtson argues that Basque, Caucasian, and Burushaski form a subgroup within Dene-Caucasian, but treats Yeniseian and Na-Dene as later movements from East Asia. Ruhlen, *Origin of Language*, pp. 74, 143, 164–166.

location of its homeland. A different homeland would lead to a different interpretation of paths of migration.[67]

Let us turn explicitly to exploring the four main possible routes for the occupation of temperate Eurasia in about 40,000 B.P. First, as implied above, there is the argument for migration up the Pacific coast. Maritime peoples of the Southeast Asian tropics, in advancing northward, could have gradually accommodated themselves to the changing seaside species. (The importance of seafood in the cuisine of Korea and Japan today may thus be the reflection of an ancient tradition.) At a region of the coast opposite Hokkaido and Sakhalin, these coastal populations may have developed the techniques of hunting, boating, and gathering that made possible life beyond the coast. They then could have moved west, spreading out and diverging to become the various Eurasiatic-speaking populations. The Amur River valley presents the interesting possibility of a waterway by which coastal peoples could gain acquaintance with inland regions.[68] This approach focuses on the concentration of Eurasiatic subgroups on the northwest Pacific shore: Korean-Japanese-Ainu, Gilyak, Chukotian, and, nearby, Altaic. In this case, Eurasiatic would most likely have descended from Austric, though other possible linguistic ancestors include Sino-Tibetan and Indo-Pacific.

An additional dimension to the story of this first route is the development of a new type of boat: skin boats. These are boats in which animal skins are sewn and stretched over a wooden framework. Maritime archaeologist Paul Johnstone has noted the distribution of such boats all over northern Eurasia and into arctic North America.[69] This is rather precisely the distribution of the Eurasiatic languages. Skin-boat technology was invented at some place and time, and it may have been along the northeast Asian coast some 40,000 years ago. While reed boats were probably the main watercraft of tropical populations as they began to move north along the Pacific coast, they had disadvantages that would have become increasingly problematic as people moved northward into colder climates. First, the reeds necessary to

---

[67] Ruhlen argues that Dene-Caucasian originated somewhere in the Near East, with groups moving east and west from that point; he also argues that Eurasiatic originated somewhere in the Near East. But if Basque, Caucasian, and Burushaski (in Pakistan) turn out to form a group that is parallel to others in Sino-Tibetan, then it makes sense to argue that the highlands of Yunnan were the homeland not only of Sino-Tibetan but of the larger Dene-Caucasian group. Ruhlen, *Origin of Language*, p. 74.

[68] One complication is that the Amur Valley is mostly forested; to its west and south begin the grasslands that stretch across Eurasia.

[69] Johnstone, *Sea-Craft*, pp. 36–43.

make the reed boats became sparser in temperate climates; second, and more importantly, reed boats sit low in the water and expose the mariners to the water. The invention of skin boats required the ability to hunt large animals efficiently and also required development of effective awls to puncture the skins and sew them together with either animal or vegetable ties, plus the ability to construct a sturdy wooden framework. Skin boats, once created, had the advantage of riding high in the water and keeping their passengers relatively dry. They were also light and portable. They could first have been tried out in rivers, and then extended to use in the seas. In one way or another, the development of skin boats seems to have been important to the occupation of temperate and arctic Eurasia.[70]

A second path to the north was from the Sino-Tibetan homeland to the Eurasian steppes. This trail would have led from what is now South China, with the migrants moving up and down various river valleys and learning how to live in progressively drier zones that brought changing systems of rains. Movement eastward toward the Pacific should have been easy at any point, but movement westward was easy only north of the Himalayas, from the latitude of the Huang He River. In effect, then, such migrants would have followed what later became the Silk Road to reach and settle in Central Asia, the Caucasus, and Europe. This might have been the path of Dene-Caucasian speakers as they moved north from a tropical homeland, then branched out east and west as they reached the grasslands. But the present wide dispersion of the communities speaking Dene-Caucasian languages makes it difficult to reconstruct the timing and the steps of their migration.

A third path to the temperate zone might be labeled the Nile–Fertile Crescent–Caucasus path. This path is often assumed to be the path by which humans left Africa and settled the Eurasian heartland. For instance, geneticist L. L. Cavalli-Sforza, in his authoritative survey of the genetics of human migration, has assumed that this was the path for the human migration out of Africa.[71] It is a superficially plausible route, but when examined in detail it reveals three types of difficulty, in the linguistic, ecological, and genetic arguments in favor of such a

---

[70] From east to west, the five great basins of the Amur, Lena, Yenisei, Ob, and Volga, linked by portages, make it possible to cross northern Eurasia by boat. For description of travels across this region in recent times, see James Forsyth, *A History of the Peoples of Siberia: Russia's North Asian Colony 1581–1900* (Cambridge: Cambridge University Press, 1992), pp. 5–10.

[71] Cavalli-Sforza, *Human Genes*, p. 64.

route. I can state the ecological point concisely, while the other two points must be explained at greater length. The ecological differences between the middle Nile and the Fertile Crescent or Caucasia—the differing vegetation, temperatures, and patterns of rains—while easily surmounted by human technology in more recent times, were not necessarily easy for humans to overcome 60,000 years ago. We need clearer archaeological documentation of *Homo sapiens* in the Fertile Crescent before 40,000 B.P. than is now available to argue that this was the main route out of Africa.[72]

Recent linguistic analyses give no clear support to the Nile–Fertile Crescent–Caucasus route, in contrast to the linguistic logic behind the first two routes. This point is worthy of some emphasis because it contradicts earlier linguistic analysis that claimed that such links could be shown. The Semitic languages are spoken in southwest Asia and northeast Africa. Because the Semitic languages were so influential in the development of writing and such important texts as Hammurabi's legal code and the Hebrew Bible, scholars of the nineteenth century sought to link Indo-European to Semitic.[73] And since social-scientific analysis in the nineteenth century focused especially on racial identity, there was reason to try to link Semitic speakers to Indo-European speakers on the grounds that both were part of a Caucasian race, based especially on assessment of skin color. Scholars seeking to identify a "Nostratic" group of languages related to Indo-European in their early work revealed a continuation of this thinking. They correctly included Altaic, Uralic, Korean, and Japanese in this larger grouping, but also sought to include Semitic and Dravidian in what have been shown to be incorrect classification within Nostratic.[74]

In particular, the Semitic languages have shown to be one of seven subgroups within the Afroasiatic language family, and the homeland of the Afroasiatic family has been shown with increasing clarity by recent evidence to have been in the middle Nile Valley—so any route from the Afroasiatic homeland to the Eurasiatic homeland was a long one and not a short one.[75] In sum, this third route remains possible as a path for occupation of temperate Eurasia, but the evidence for it is

---

[72] Stringer and McKie, *African Exodus,* pp. 54–114.

[73] Greenberg, *Indo-European and Its Closest Relatives,* vol. 1, *Grammar,* p. 9.

[74] The relationship of Kartvelian to Eurasiatic and to Afroasiatic languages remains unresolved. Dolgopolsky, *Nostratic Macrofamily;* Greenberg, *Indo-European and Its Closest Relatives,* vol. 1, *Grammar,* p. 9.

[75] Ehret's classification divides Afroasiatic into Omotic and Erythrean, Erythrean into Cushitic and North Erythrean, North Erythrean into Chadic and Boreafrasian, and Boreafrasian into Egyptian, Berber, and Semitic. According to this classification, any Afroasi-

not strong. If there were a link between Afroasiatic and Eurasiatic, such that Eurasiatic emerged from Afroasiatic (or from an ancestor to Afroasiatic or a descendant of Afroasiatic such as Semitic), then the migration route of early Eurasiatic speakers could indeed have encompassed the Nile–Fertile Crescent–Caucasus route. From a center in this region, humans could have occupied the forested and steppe areas before the Ice Age. A linguistic relationship between Afroasiatic and Eurasiatic remains conceivable, but no clear statement of it has been offered. In addition, if Greenberg is correct that Eurasiatic and Amerind are sister stocks, then Afroasiatic should have the same relationship to Amerind as it has to Eurasiatic.

A further difficulty with the Nile–Fertile Crescent–Caucasus route is in the genetic evidence. Although genetic evidence is commonly argued to support the case for a path of migrants from the Nile Valley through the Fertile Crescent and to Eurasia generally, I think the historical projections of genetic evidence need to be recalculated. In particular, the current projections contain a consistent bias that underestimates the genetic distance among populations geographically close to each other and exaggerates the genetic distance among geographically distant populations.[76] Cavalli-Sforza's extensive research and careful summaries reflect the seriousness of his attempt to correlate work from all fields of study contributing to the study of early humanity. Yet there remain curious results that do not fit in. Systematically, the most isolated populations are those calculated as having the greatest genetic distance from others, and hence as being the oldest. As a result, he estimates the divisions among populations in the central parts of Eurasia as being relatively recent.[77] In another curious decision, Cavalli-Sforza uses inherited racial terms to classify phenotypes, though genetic work has made clear that physical appearances repre-

---

atic-speakers who were early colonists of the Caucasus would not have been Semitic speakers, but would have been from the earlier Erythrean or North Erythrean language groups. Christopher Ehret, *Reconstructing Proto-Afroasiatic (Proto-Arasian): Vowels, Tone, Consonants, and Vocabulary* (Berkeley: University of California Press, 1995), pp. 489–490; Ehret, "Language and History," and Richard J. Hayward, "Afroasiatic," in Heine and Nurse, *African Languages*, p. 292 and pp. 83–86, respectively.

[76] If humans migrated from Africa to Southeast Asia (and Australia and New Guinea), and then to temperate Eurasia, then the genetic distance between Africans and temperate Eurasians should be greater than that between Africans and Australians. But subsequent and repeated mixture of populations within temperate Eurasia, and mixture of these populations with those of the northern half of Africa, has reduced the genetic distance between Africans and temperate Eurasians. So far, genetic analysis tends to report on the similarities and differences of populations, but not on *when* the similarities and differences emerged.

[77] Cavalli-Sforza, *Human Genes*, pp. 79–80, 135; see also pp. 248–254.

sent a small part of genetic difference.[78] A look at the global map of
skin colors in the same volume, showing the differences of skin color
within the Americas, suggests strongly that environment and not just
heredity affects human phenotype.[79]

There is, finally, a fourth path from tropical to temperate Eurasia
that can be hypothesized: a path leading from the Dravidian-speaking
zone of the Indian Ocean littoral across the mountains and north-
ward. In more recent times other populations have migrated in the
opposite direction, from Central Asia into India, so it is possible that
a previous migration might have gone northward. I know of no serious
attempts to make this case in either archaeological or linguistic terms,
though one could imagine the possibility that Eurasiatic languages are
descended from Dravidian. The route from tropical waters to temper-
ate grasslands, though mountainous, was rather short in this case.

Here is my assembly and summary of the complex possibilities out
of which we must reconstruct the human occupation of temperate
Eurasia. Overall, I would argue that there were three substantial migra-
tions from the tropics to temperate Eurasia, and one cannot yet be cer-
tain about their relative timing. A movement overland (or in part
along rivers in the valleys east of the Himalayas) from South China to
the Eurasian steppes may have given birth to a temperate population.
This group, speaking Dene-Caucasian languages, made initial adjust-
ments to life in temperate zones. The second substantial migration
moved north along the western Pacific shore. This movement led to
formation of the Eurasiatic language group, which then spread to dis-
place or assimilate earlier groups except for some Dene-Caucasian rem-
nants. At the very least, the linguistic diversity of the north Pacific
coast suggests that it was a place of early settlement and a homeland
for groups of migrants. Third, a northward movement of African-
based Afroasiatic speakers may have contributed to settlement of tem-
perate Eurasia. Because of the clear demonstration that the Semitic
languages (along with Egyptian and Berber) are relatively recent sub-
groups within the Afroasiatic languages, I think it is most likely that
the Semitic speakers moved from Africa to Arabia and the Fertile
Crescent after the most recent glacial maximum.[80]

The ability to occupy northern Eurasia prepared humans for entry
to North America, either on foot or by boat. As they entered the

---

[78] For instance, he uses the term "Caucasoid" when referring to North Africans. Ibid.,
p. 167.
[79] Ibid., p. 145.
[80] See n. 77.

Americas, humans found no hominid competitors. But as had been the case in Australia and northern Eurasia, they did encounter mega-fauna—in this case large mammalian species—and the expansion of humans correlated provocatively with the disappearance of the mega-fauna.[81] The archaeological remains of early humans in the Americas have been sparse so far, indicating that populations were either late to arrive or slow to grow. I believe, however, that linguistic and genetic evidence argues for an early occupation of the Americas—before the last great Ice Age.[82]

Between 30,000 and 15,000 B.P., the earth experienced one more wave of cooling: massive sheets of ice formed at both poles and extended to cover most of Europe and North America. Sea level fell by 40 meters to a level more than 100 meters below today's sea level (see Figure 1). The small human population in northern Eurasia and the smaller population in the Americas had to withdraw to more southerly regions, and every human population had to adjust to a climate that was cooler and also drier (since so much water was congealed in icy form).

I think that Eurasiatic and Amerind language groups both had their origins on the western shores of the north Pacific. Amerind then spread into the Americas, before the last Ice Age took hold in 35,000 B.P., while Eurasiatic spread westward across the Eurasian steppes. I think that both groups relied on boats as well as on the soil: they stayed close to rivers as they moved inland, and they hunted large animals as well as small on land and at water's edge.[83]

Regardless of the outcome of my hypothesis, it is clear that Eurasiatic and Amerind must be compared with other major language groups, to see if it can be determined with which tropical groups they are affiliated. The full list of candidate groups from which Eurasiatic and Amerind might have sprung includes Nilo-Saharan, Afroasiatic, Dravidian, Sino-Tibetan (or Dene-Caucasian), Austric, Indo-Pacific, and Australian. Of these I think Austric is the most likely parent or affiliate of Eurasiatic, but that assertion is based so far on geographic proximity rather than on any detailed linguistic comparison.

---

[81] For a good survey of research and debates on megafaunal extinctions, see Alfred W. Crosby, *Throwing Fire: Projectile Technology Through History* (Cambridge: Cambridge University Press, 2002), pp. 52–69.

[82] See nn. 65 and 66.

[83] It is of interest that Amerind speakers appear not to have had skin boats. Nevertheless the bark canoes built around wooden frameworks, so widely used in North America and also used in Siberia, relied on a principle similar to that of skin boats.

One important issue that I have skimmed over is the interaction of *Homo sapiens* and other hominids.[84] The linguistic evidence discussed above, while it does not give a definitive answer to how temperate Eurasia was occupied, provides important background for understanding the ways in which *Homo sapiens* encountered and displaced previous hominids. Especially for Europe, we have evidence to help clarify the story of the competition of *Homo sapiens* for space with hominid predecessors, especially with *Homo neanderthalensis* in Europe. The genetic evidence so far indicates little interbreeding of the two closely related hominid populations. A likely scenario is that the incoming *Homo sapiens* occupied the best lands, grew in population, and reduced the preceding populations to marginal life and then to disappearance. Some intermixture could have occurred within this scenario.

Conclusion

This interpretation of human migration in the period up to the end of the last Ice Age has focused principally on the benefits of adding linguistic analysis to the recent advances in study of genetics, archaeology, paleontology, and earth sciences. Systematic consideration of linguistic evidence, along with that of genetics and archaeology, can give us more detail and can resolve some of the ambiguities in present interpretations. Both genetic composition and languages evolve, but they evolve in different fashions, and a detailed reconstruction of both sorts of evolution can add substantial new information on the paths and the timing of early human movements.

Available linguistic information, as interpreted here, is more specific on the paths of human migrants than are the data from genetics and archaeology. The patterns of language suggest a gradual human occupation of the Old World tropics, reaching its geographical limits about 50,000 B.P. Then, after a pause, humans accommodated to life in temperate and even arctic zones and achieved a rapid occupation (though perhaps in two stages) of northern Eurasia; the occupation of North America took place as part of the same movement northward. Occupying the remainder of the Americas, however, was a daunting task that involved adaptation to a succession of montane, arid, and tropical environments.

---

[84] On the intriguing discovery of remains of diminutive hominids on the island of Flores, 18,000 years ago, see P. Brown et al., "A New Small-Bodied Hominid from the Late Pleistocene of Indonesia," *Nature* 431 (2004): 1,055–1,061.

The evidence of language provides essential clues on the timing and direction of early human migrations. This use of linguistic data to support long-term interpretations appears to make it fit well with available genetic and archaeological data, and even to fill blanks in genetic and archaeological analysis. Such a use of linguistic data, however, involves stretching of the interpretation of language phyla to much longer time frames than has been conventional. Therefore, the linguistic analysis I have presented above can neither be confirmed nor refuted at present because of the inconsistency of methods and standards in historical linguistics. Those with training in linguistics and especially in historical linguistics need to take leadership in debating the inconsistencies in their classification of languages and in their assessment of the historical depth of language groups. At the same time, world historians, who reach habitually across disciplinary boundaries, should not hesitate to involve themselves in the research and debate on language and early human history, as the linkage to genetic and archaeological data may help resolve some of the linguistic debates.

Appendix: Language Families According to *Ethnologue*

Afroasiatic (372)
Alacalufan (2)
Algic (40)
Altaic (65)
Amto-Musan (2)
Andamanese (13)
Arauan (8)
Araucanian (2)
Arawakan (60)
Artificial language (3)
Arutani-Sape (2)
Australian (258)
Austroasiatic (168)
Austronesian (1262)
Aymaran (3)
Barbacoan (7)
Basque (3)
Bayono-Awbono (2)
Caddoan (5)
Cahuapanan (2)

Cant (1)
Carib (29)
Chapacura-Wanham (5)
Chibchan (22)
Chimakuan (1)
Choco (10)
Chon (2)
Chukotko-Kamchatkan (5)
Chumash (7)
Coahuiltecan (1)
Creole (81)
Deaf sign language (114)
Dravidian (75)
East Bird's Head (3)
East Papuan (36)
Eskimo-Aleut (11)
Geelvink Bay (33)
Guahiban (5)
Gulf (4)
Harakmbet (2)

Hmong-Mien (32)
Hokan (28)
Huavean (4)
Indo-European (443)
Iroquoian (10)
Japanese (12)
Jivaroan (4)
Katukinan (3)
Keres (2)
Khoisan (29)
Kiowa Tanoan (6)
Kwomtari-Baibai (6)
Language Isolate (30)
Left May (7)
Lower Mamberamo (2)
Lule-Vilela (1)
Macro-Ge (32)
Maku (6)
Mascoian (5)
Mataco-Guaicuru (11)
Mayan (69)
Misumalpan (4)
Mixe-Zoque (16)
Mixed language (8)
Mosetenan (1)
Mura (1)
Muskogean (6)
Na-Dene (47)
Nambiquaran (5)
Niger-Congo (1489)
Nilo-Saharan (199)
North Caucasian (34)
Oto-Manguean (172)
Paezan (1)

Panoan (30)
Peba-Yaguan (2)
Penutian (33)
Pidgin (17)
Quechuan (46)
Salishan (27)
Salivan (2)
Sepik-Ramu (104)
Sign language (2)
Sino-Tibetan (365)
Siouan (17)
Sko (7)
South Caucasian (5)
Subtiaba-Tlapanec (4)
Tacanan (6)
Tai-Kadai (70)
Torricelli (48)
Totonacan (11)
Trans-New Guinea (552)
Tucanoan (25)
Tupi (70)
Unclassified (96)
Uralic (38)
Uru-Chipaya (2)
Uto-Aztecan (62)
Wakashan (5)
West Papuan (26)
Witotoan (6)
Yanomam (4)
Yenisei Ostyak (2)
Yukaghir (2)
Yuki (2)
Zamucoan (2)
Zaparoan (7)

Source: Ethnologue, www.ethnologue.org/family_index.asp.