Language Resources

Online Resources to accompany *A History of Humanity: The Evolution of the Human System*
By Patrick Manning

Introduction

The purpose of this online resource collection is to interpret the place of language in human history. In a simplified presentation of a complex issue, this Introduction begins with concise definitions and descriptions. It traces the logical order of language divergence and displays the major phyla or families going back more than 15,000 years. After summarizing the history of language divergence and movement in six periods, we turn to the problems and debates in language history. These include the effects of "language overlays" as one replaces another, efforts to define "macro-phyla" for very early times, and early maritime migrations.

The accompanying files for 14 individual phyla provide descriptions of each Homeland, language migrations over time, maps (which are also available as separate image files), concise spreadsheets showing major subgroups in each phylum, and citations of works on each phylum. In a separate Excel file, the 14 individual sheets each give a restatement of the concise spreadsheet at top and, below, a full spreadsheet showing many of the languages in each phylum.

Definitions

The **elements of language**, as understood by linguists, include *lexicon* (the meanings of words), *morphology* (the pieces of words and how they are fit together), *phonology* (the sounds made in any language), and *syntax* (the organization of lexicon, morphology, and phonology into meaningful sentences). Language **phyla** (singular **phylum**) are ancestral families of languages, defined here as families likely to have existed for more than 15,000 years. This analysis traces the history of 14 known phyla, with attention to their most prominent subfamilies.

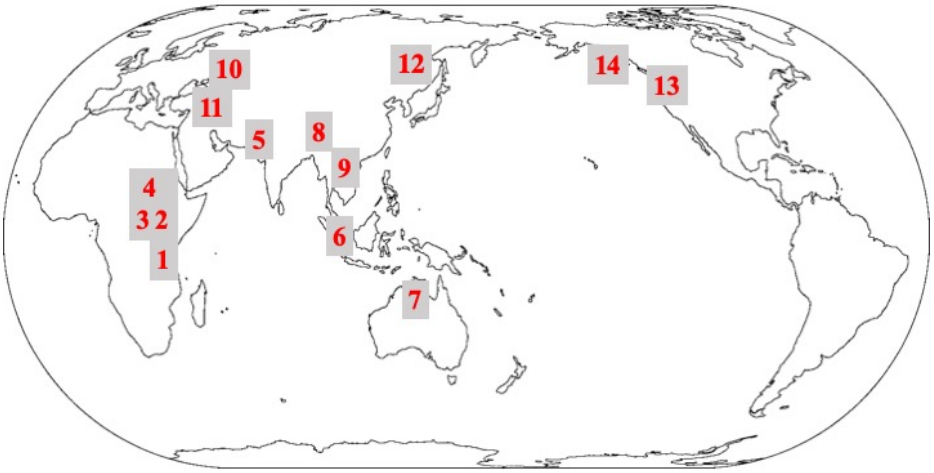
The order of language change is the principal emphasis in this historical summary. Languages change through an orderly divergence over time of their lexicon, phonology, morphology, and syntax. In the "tree model" that is applied here, descent of languages is unilineal, with a genealogy such that each language has a single parent but may have several sisters. For each phylum, a concise spreadsheet displays the top four levels at which languages diverged and gave way to daughter languages. Language families displayed in spreadsheets, along with locations of current languages, make it possible to estimate the places in which the parent languages were spoken. Relying on this approach of working from current to past languages, the materials here show estimations of the geographic homeland in which the founding language of each phylum was spoken. (Using more complex techniques, linguists also work toward reconstruction of the lexicon and syntax of the founding language of the phylum.)

Language Phyla: Homeland and Distribution ca. 1500 CE

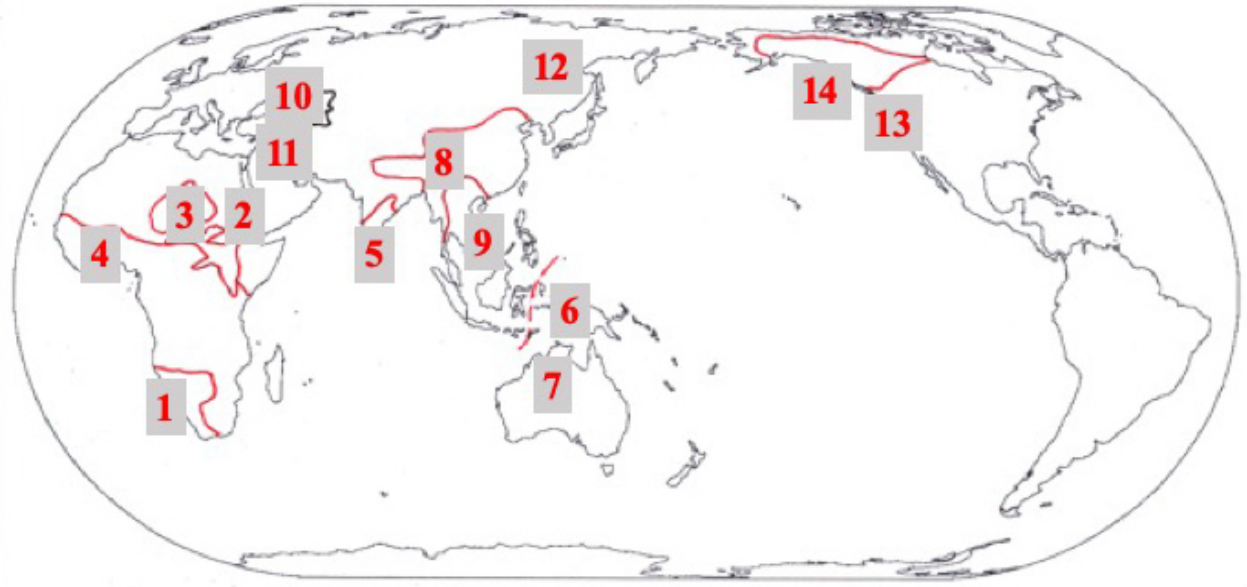
The accompanying maps show the estimated geographic homeland of each phylum and the geographic distribution of each phylum in ca. 1500 CE, before the migrations of the oceanic age. (Not shown are the many subfamilies within each phylum).

Language phyla, showing estimated homeland (ordered in estimated order of emergence):

1.Khoesan
2.Afroasiatic
3.Nilo-Saharan
4.Niger–Kordofanian
5.Elamo-Dravidian
6.Indo-Pacific
7.Australian
8.Trans-Himalayan
9.Austric
10.North Caucasian
11.Kartvelian
12.Eurasiatic
13.Amerind
14.Na–Dene



Language phyla, showing homeland and boundaries of phyla ca. 1500 CE:



History: Eras of Expansion in Language Groups

Processes of change are fundamental to languages. Some direct or indirect descendants of early language families survive today, enabling us to estimate the process of their initial migration and settlement. On the other hand, more recent migrations and settlements have overlaid or

imposed new languages that have replaced earlier languages, making it difficult to recover old language history. Here is a concise summary of the migration and divergence of languages and their speaking populations in six periods from 70,000 years ago to the recent past.

Origin of Spoken Language. It is assumed here that the process began with the rather sudden invention of spoken language – "proto-human" language – some 70,000 years ago, among people in the Northeast African region, who were a subgroup of the *Homo sapiens* population spread throughout the African continent. Further, it is assumed that the original language gave rise to all other human languages by processes of differentiation and divergence. While we have no direct linguistic evidence on the place of origin within this region, it is striking that four major phyla, the source of virtually all of Africa's languages, have their homeland in this general region. The Language Homeland has a very diverse ecology, supporting many sorts of plants and animals.

Pleistocene Tropical Expansion, 65,000 to 45,000 Years Ago*.* The initial language phyla whose descendants are now known as Khoesan, Afroasiatic, Nilo-Saharan, and perhaps Niger–Kordofanian took form within a few thousand years of the rise of a "proto-human" language. Early groups developed in the middle Nile Valley, later groups migrated to some distance away, and an alternation of large and small moves persisted over time. Migration eastward, into Asia, proceeded along the Indian Ocean littoral, relying on boats in significant measure. These migrations led to communities that formed the Elamo-Dravidian, Indo-Pacific, and Australian phyla. Similarly, and further inland, settlers formed the Trans-Himalayan and Austric phyla. In both Africa and Asia, the speaking human migrants encountered other hominins who lacked language: it is known that they interbred; those incorporated into the speaking *Homo sapiens* community may have learned to speak.

Pleistocene Temperate Expansion, 45,000 to 21,000 Years Ago*.* Humans entered the temperate zone 45,000 years ago, most likely by moving north from the Indus Valley or the Persian Gulf. Once they reached the grasslands that lay beyond the intervening deserts and mountains, the migrants moved both westward into Europe and eastward into northeastern Asia. Three phyla survive to reveal the story of the settlement of temperate lands—North Caucasian, Kartvelian, and Eurasiatic—but also the Basque family. Of these, Eurasiatic has since grown the most and thus preserves the best record of its origin and expansion. In Europe and Central Asia, the speaking human migrants encountered Neanderthal communities; again, it is known that they interbred.

Terminal Pleistocene Expansion, 21,000 to 12,000 Years Ago. Soon after the most extreme moments of the Glacial Maximum, temperatures began to rise at a rapid rate. This terminal Pleistocene era of warming encouraged migration in many parts of the world, and it also brought the development of housing and other new productive techniques by humans everywhere. Maritime migrants from northeast Asia traveled eastward to North America, relying on the "kelp highway" of offshore resources to travel beyond the Laurentide ice sheet until they settled at the Salish Sea (the region of today's Vancouver and Seattle) and created the Amerind phylum. From there, other voyagers moved southward along the western coast of the Americas, spreading the Amerind phylum across most of North and South America. Soon thereafter, a subsequent group of mariners arrived from northeast Asia, launching the Na–Dene phylum along the north Pacific coast and up the Yukon Valley. By 15,000 years ago, therefore, humans had settled most of the lands of both hemispheres.

Early Holocene, 12,000 to 6000 Years Ago. Though with fluctuations, the warming of Earth continued for 4000 years into the Holocene era. In this age, productive techniques expanded especially with the rise of agriculture and animal husbandry in several regions. The migrations of this era were no longer the occupation of lands empty of humans, but the entry of new settlers into existing settlements. Migrations overlapping the Terminal Pleistocene and Early Holocene included a dramatic expansion of Afroasiatic speakers from the Nile Valley to the west, north and northeast making their languages dominant in much of the then-grassy Sahara, North Africa, the Levant, and Arabia. In a parallel movement, speakers of Chinese languages expanded from southern highlands into the lowlands of what are now south China and the Yellow River Valley of north China. Far to the north, speakers of Yukhagir family of Eurasiatic languages occupied the Arctic littoral, moving to the west. Agriculture developed especially in the Levant, and from there it seems to have spread eastward to India, westward to Europe, and northward to the Caucasus, though the language groups of these early farmers are not yet known.

Middle and Late Holocene, 6000 to 1000 Years Ago. Agricultural technology continued to advance. Rice farmers developed paddy rice, relying on water buffalo to build terraces and irrigate their crops, settling lands on the Asian mainland and throughout the Philippine and Indonesian archipelagoes. Ox-plow-driven cultivation of wheat and barley, similarly, led to expansion of Indo-European speakers throughout Europe. Yam-farming Bantu speakers, beginning near Mt. Cameroon, spread and diversified their production, eventually occupying most of the southern third of Africa. As horses became domesticated in the steppes of west and central Asia, the resulting war chariots enabled Indo-European and Altaic-speaking warriors to conquer widely until others adopted the weaponry. Yukhagir-speaking migrants moved to the west of the Urals, now relying on domesticated reindeer for traction and milk as well as meat. In North America, Algonkian speakers spread from their homeland near the Pacific to become dominant throughout the St. Lawrence Valley, while Penutian speakers migrated southeast from California, ultimately becoming the main population of the Maya city-states of Central America. More widely known as "civilizations" but emerging in the same era were the urban populations of Mesopotamia, the Nile Valley, the Yellow River Valley, and the Indus Valley.

Within the most recent thousand years, wars brought substantial loss of life and widespread migration in the Mediterranean, West Asia, and East Asia. Most spectacular was the rise of the Mongol Empire, which was the primary force for Eurasian change for nearly two centuries. From 1500, languages of western Europe reached coasts around the world, as Russian language reached across Siberia. From 1800, western European languages overlaid languages of the Americas and, by 1900, became second languages in Asia, Africa, and the Pacific. Nevertheless, Chinese, Japanese, Arabic, and South Asian languages also spread through migration.

Major Issues in the History of Language Change

Language Overlays in History of the Holocene Epoch. Where a new language and new settlers enter a region, they *overlay* and may eliminate the previously spoken language. Examples of overlays by expanding migratory groups include the Boreafrasan and Semitic families within Afroasiatic; the Bantu and Adamawa–Ubangi families within Niger–Kordofanian; the Austronesian family within Austric; the Altaic and especially Indo-European families within Eurasiatic; the Chinese and Burmese families within Trans-Himalayan; and the Algonkian and Penutian families within Amerind. Evidence on these migrant overlays is greatly valuable for documenting patterns of movement and social and technical change during the Holocene epoch.

In consequence, however, these overlays have obscured knowledge of earlier inhabitants and languages. Language phyla that are heavily overlaid include Khoesan, Indo-Pacific, Australian, Elamo-Dravidian, North Caucasian, and Kartvelian. Perhaps some entire phyla were obscured completely so that we retain no knowledge of them. The Australian case is unusual in that one of the subgroups, Pama–Nyungan, expanded so greatly that it obscured major groups within the phylum. The world regions in which the overlays are most substantial are central Africa, maritime Southeast Asia, Australia, South Asia, West and Central Asia, Europe, and China. In sum, the world's most central and most densely populated regions are those with the most overlays. In those regions it is most difficult to dig back and document the history of language before the mid-Holocene.

Because languages change with time, it becomes progressively more difficult to trace relationships among them in the distant past. Historical linguists have been reluctant to trace relations among languages as far back as the start of the Holocene epoch, 10–12,000 years ago. On the other hand, the hypothesis that spoken language began some 70,000 years ago provides a base from which one can hope to trace the differentiation among languages. That is, efforts to link the known Holocene patterns of language with the hypothesized proto-human language of Northeast Africa may enable researchers to fill in the gaps and develop a fuller picture of Pleistocene-era language change.

Certain phyla sustain evidence on the long-term path of their evolution, from the emergence of their homeland up to their recent divergences. Phyla that are little overlaid include Afroasiatic, Nilo-Saharan, Niger–Kordofanian, Eurasiatic, Amerind, and Na–Dene. Therefore, we can turn to these phyla in search of evidence on the earliest human migrations. All of these were surely founded in the Pleistocene epoch and leave evidence on their evolution and migration in times

before the Holocene. Amerind and Australian are the cases in which the interactions of phyla were almost entirely among subgroups within the phylum rather than with other phyla. In the case of Eurasiatic, the westward spread of settlers appears to have encountered one after another competing groups yet continued to expand.

Within the past 500 years, the Indo-European family of Eurasiatic has overlaid most of the previous territory of Amerind languages, yet study of Amerind languages continues. Similarly, the Trans-Himalayan and Austric phyla appear to be intermediate cases – their histories include Holocene-era migration and overlays, but their homelands and nearby areas may yet reveal information on Pleistocene migrations.

Theses on Macro-phyla. Three groups of linguists have proposed "macro-phyla," based on groupings of language families, that are argued to have descended from a common ancestor beginning in Pleistocene times. None of these language-familial structures has been verified: all have their strengths and weaknesses and all will necessarily be part of future discussions on the history of Pleistocene-epoch language distribution and divergence.

"Dene–Caucasian" (John Bengtson). The hypothesis is that the Trans-Himalayan (Sino-Tibetan) phylum is the ancestor of the language families of North Caucasian, Basque, Yenisei, and Na–Dene, assuming migration from the Trans-Himalayan zone to temperate Eurasia ca. 45,000 years ago. It also assumes a later migration of Yeniseans from the Trans-Himalayan zone to Siberia and a migration from there to found the Na–Dene family. The argument implicitly assumes substantial overlay of temperate-zone Dene–Caucasian speakers by later migrations of Eurasiatic speakers.

"Southern" (Joseph Greenberg in Crofts) along with "Kongo–Saharan" (Edgar Gregersen). The hypothesis implicitly assumes Nilo-Saharan as the source of populations migrating eastward and forming the Elamo-Dravidian, Indo-Pacific, and Australian phyla, in the initial human occupation of tropical Asia ca. 65,000 years ago. The thesis raises but leaves unresolved the question of what language group would have been the basis of Trans-Himalayan, Austric, and other Asian-based phyla. In a related thesis, Gregerson proposed that Nilo-Saharan and Niger–Kordofanian phyla are both descended from a "Kongo–Saharan" macro-phylum, well before the Last Glacial Maximum.

"Nostratic" (Aharon Dolgoposky). The hypothesis for this macro-phylum centers on a presumed homeland in the Levant or Caucasus, ca. 20,000 years ago. From this homeland, divergence and migration created most but not all of the families within Eurasiatic, as well as Dravidian, Kartvelian, Sumerian, and Afroasiatic. But the analysis of Afroasiatic argues that its homeland is in the Nile Valley, not the Levant or Caucasus; the analysis of Eurasiatic argues that its homeland is in northeast Asia not the Caucasus or Levant. Ehret argues that the eclectic identifications of common roots in "Nostratic" is more likely to be traces of proto-human vocabulary.

Maritime Migration. Maritime migrations have been central at several stages of human expansion. First was the crossing of the Bab el-Mendeb from Africa to Asia more than 60,000 years ago, followed by the maritime dimension of eastward migration along the Indian Ocean

littoral. Second was the crossing of oceanic spaces of 60 to 100 km from Sunda to Sahul as settlers reached what are now Australia and New Guinea, somewhat later than 60,000 years ago. Third was the migration some 30,000 years ago, documented indirectly, of Indo-Pacific speakers from the Indonesian or Philippine archipelagoes, along the islands of the western Pacific to Japan or even to the Amur River estuary. Fourth was the migration, relying on the "kelp highway," of settlers from northeast Asia past the Laurentide ice sheet to their settlement south of the ice at the Salish Sea, to form the Amerind-language homeland about 19,000 years ago. This migration continued as more settlers moved along the Pacific littoral all the way to Monte Verde in South America. Fifth, a separate group of maritime settlers also relied on the kelp highway, settling on the ice-free Haida Gwaii island to found the Na–Dene language homeland about 17,000 years ago. Later maritime migrations included Indo-Pacific migration from New Guinea to the Solomon Islands and beyond. About 4000 years ago, Austronesian settlement of the Philippine and Malay archipelagoes was followed by interplay of Austronesian and Indo-Pacific communities in New Guinea, and later by the settlement of Micronesian and Polynesian islands of the Pacific, then by Indonesian migration to East Africa and settlement of Madagascar roughly 2000 years ago. In addition, Arawak migrants from South America occupied much of the Caribbean archipelago some 5000 years ago, followed by Carib settlers from South America. Finally, the newly opened global sea lanes after 1500 allowed the migration of western Indo-European speakers – and, indeed, speakers of all languages, to settle along all the coasts of the world.

Theoretical Debates in Historical Linguistics

Debates in the field of historical linguistics arise especially because of the inherent difficulty of developing consistent theory. Nevertheless, it is important that debates in this field be pursued in order to yield a consistent view of the role of language in the human past.

- The process of unique descent of all spoken languages from an original language is not accepted by all linguists. Even in accepting this single-origin thesis for language, languages are known to change rapidly enough that relations among them are difficult to discern after several thousand years.
- In the basic model of language change, one finds debate between the proponents of tree models, proponents of wave models, and those seeking a mix of the two.
- In tree classification, one finds differences between "lumpers" who identify very large phyla (e.g., Nostratic) and "splitters" who find that phyla cannot be confirmed (e.g., rejecting Amerind, Indo-Pacific, and Altaic).
- Debates over what aspects of language are most fundamental: lexicon, morphology, phonology, syntax, and combinations of these.
- Contrasting views of language by world region. Since ca. 1950, social science and humanities scholarship has relied on an area studies framework, with multidisciplinary approaches to discrete world regions (Africa, East Asia, South Asia, etc.). This academic structure encourages divergence in linguistic theory and practice by region.
- Language universals. The exploration of universal patterns in language, led by Joseph Greenberg from the 1960s to 1990s.

- Lexicostatistics. Efforts to make quantitative estimates of rates of language change, pioneered by Morris Swadesh in 1950. This approach has been extended to Bayesian methods of linguistic analysis.
- The hope for global theory and analysis of language classification. At a time when analysis of human evolution is racing ahead on many disciplinary fronts, the historical study of language is adding very little that is new. As a result, language evidence is largely ignored in studies human history and human social evolution. While various research groups are definitely pursuing their own local objectives, there is no sign of a set of congresses to define an overall research agenda – one intended to link up to current advances in paleontology, archaeology, biology, genomics, anthropology, and history. It may be that the key new insights have already been developed. But what is needed is public discussion, hypothesis testing, and practical contrast of alternative approaches until the field of historical linguistics begins again to generate the advances that will contribute to fuller understanding of the place of language in human evolution.

Language Spreadsheets: A Guide to Details on Language Phyla

The language spreadsheets are shown as separate sheets on one large Excel file that may be downloaded. Each sheet lists the phylum by name in the upper left cell. The columns show the succeeding “generations” or levels of subfamilies, with family members listed below and to the right of the parent language; the rows show succeeding family groups within each generation.

- The spreadsheet summary is a simplification of language reality: it has rigid columns and rows into which language descent is fit. Each spreadsheet column can be seen to represent one or several generations of language change.
- The spreadsheet model relies on ordinal but not chronological time. That is, each step downward or to the right corresponds to a later time, but the time cannot be quantified without additional information.
- All languages are listed in terms of their relationship to other languages in a unilineal descent system. The spreadsheet includes no explicit indication as to the number of speakers of any language.
- Evidence for certain languages and language families may be missing. When there is no evidence for a language or group, the analysis may have to carry on as if the languages in question did not exist.
- When a parent language has several daughter languages, the spreadsheet displays them as if all the daughter languages diverged at the same moment.
- When several daughter languages appear at much the same time, it may be that the divergence happened at a time favorable to divergence and migration. Three such time periods were the early expansion of speaking humans in Africa and Asia, the late Pleistocene era of rapid warming, and the mid-Holocene era of agricultural expansion.
- In sum, the combination of spreadsheet and maps can give good indications on the homelands of language families and of the migrations of language families within a phylum.

References

(In addition, see the references cited for each phylum)

Bengtson, John. “Notes on Sino-Caucasian.” In Vitaly Shevoroshkin (ed.), *Dene-Sino-Caucasian Languages*. . Bochum, Germany: Brockmeyer, 1991.

Croft, William. *Joseph Harold Greenberg, 1915 – 2001: A Biographical Memoir.* Washington, DC: National Academy of Sciences, 2007.

Dolgopolsky, Aharon, with an introduction by Colin Renfrew. *The Nostratic Macrofamily and Linguistic Paleontology.* Cambridge: McDonald Institute for Archaeological Research, 1998.

Ehret, “Nostratic—or Proto-Human?” In Colin Renfrew and Daniel Nettle (eds.), *Nostratic: Examining a Linguistic Macrofamily*, 93–112. Cambridge: The McDonald Institute for Archaeological Research, 1999.

Erlandson, Jon M., Michael H. Graham, Bruce J. Bourque, et al. “The Kelp Highway Hypothesis: Marine Ecology, the Coastal Migration Theory, and the Peopling of the Americas.” *The Journal of Island and Coastal Archaeology* 2 (2007): 161–74.

Greenberg, Joseph H. *Genetic Linguistics: Essays on Theory and Method*, edited and introduced by William Croft. Oxford: Oxford University Press, 2005.

Greenberg, Joseph H. “Appendix: A generalization of glottochronology to *n* languages (1987).” In Greenberg, *Genetic Linguistics*, 108–13.

Gregersen, Edgar. "Kongo-Saharan." *Journal of African Languages* 11 (1972): 69–89.

Ruhlen, Merritt. *The Origin of Language: Tracing the Evolution of the Mother Tongue.* New York: John Wiley & Sons, 1994.

Swadesh, Morris. “Towards greater accuracy in lexicostatistic dating.” *International Journal of American Linguistics*, 21 (1955): 121–37.