

# HISTORICAL EVOLUTION OF THE WORLD'S LANGUAGES

**Ranko Matasović**

*University of Zagreb, Croatia*

**Keywords:** language diversity, language families, comparative linguistics, linguistic palaeontology, population genetics, language spread, wave of advance model, elite dominance model

## Contents

1. Introduction
  2. Models of language spread
    - 2.1. Wave of advance
    - 2.2. Elite dominance
  3. Language families in the Old World
  4. Language families in the New World
  5. Recent history
  6. Conclusion
- Glossary  
Bibliography  
Biographical Sketch

## Summary

Interdisciplinary research and cooperation of linguistics, anthropology, archaeology and population genetics have led to new insights about the prehistory of language families of the world. Several models of language spread are used to account for the current distribution of the world's languages. In some cases, this distribution reflects large-scale prehistoric migrations (the "wave of advance" model), while in other cases languages have spread without the actual movement of people, often because the idiom of a small, but dominant group acquired a great social prestige and was adopted by the majority of a given population (the "elite dominance" model).

## 1. Introduction

The subject of this chapter is the historical developments that have led to the current state and distribution of languages and language families in the world. This subject has been investigated from different points of view, and it is currently an area of interdisciplinary research.

The questions to be addressed are: why are some language families very small, in terms of the number of languages constituting them (e. g. the Kartvelian language family in the Caucasus, with only four languages), while others are extremely large (e. g. the Nigero-Kordofanian language family in Africa, with more than 1000 languages)? Why are there very few language families in some continents (or large geographical areas), while there are very many language families in other parts of the world? In what way is

the present-day geographic distribution of language families connected with prehistoric migrations of human populations?

The starting point of the historical development of the world's languages is still unknown. Attempts by a number of scholars to reconstruct the alleged protolanguage of humankind ("Proto-World"), from which all of today's languages have developed, are at best too speculative, at worst amateurish. The large majority of linguists agree that such a reconstruction is impossible, and that methods of comparative linguistics do not permit the reconstruction of proto-languages at time-depths earlier than, approximately, 10 000 years. Since humans are generally believed to have acquired language ability at least 100 000 years BP (and possibly much earlier), this means that the greatest part of the evolution of human languages and language families remains beyond the reach of comparative linguistics.

However, progress in our understanding of linguistic prehistory of the world is made possible through interdisciplinary research encompassing comparative linguistics, archaeology, and population genetics. Prehistoric population movements that left traces in archaeological record can often be correlated with the present-day distribution of language families, and methods of linguistic palaeontology allow us to conclude where proto-languages of particular families were spoken. For example, the existence of word for "beech" in the reconstructed lexicon of Proto-Indo-European was used to argue that the protolanguage had been spoken in an area where this arboreal species grows, or used to grow in prehistory. Moreover, the correlation of reconstructed vocabulary and archaeological data also permits linguists to deduce the approximate period in which the reconstructed protolanguage was spoken.

For example, the fact that we can reconstruct the noun denoting "copper" in Proto-Indo-European allows us to assume that the protolanguage from which Indo-European languages developed was spoken during the Eneolithic period (roughly the early fourth millennium BC). It is argued that Proto-Indo-European \*h<sub>2</sub>eyos, reconstructed on the basis of the regular correspondence of Latin *aes* "copper" and Sanskrit *ayas* "copper", could only have denoted copper, which means that the speakers of the protolanguage knew how to extract and use that metal. Of course, a careful application of this method requires of linguists to exclude the possibility that words for the reconstructed cultural items were actually loanwords, borrowed in different languages independently from some known or unknown source. For example, it would be absurd to conclude that speakers of Proto-Indo-European lived in the tropics because there is a shared word for "banana" in English, French (*banane*), Croatian (*banana*) and Albanian (*banane*). However, when applied with care, methods of comparative linguistics enable us to differentiate between likely borrowings and inherited words in the reconstructed lexicon of a protolanguage.

Another way of making assumptions (often just reasonable guesses) about prehistoric spreads of language families is by correlating the present-day distribution of genetically related languages with the results of population genetics. Thus, the comparison of genetic samples of the inhabitants of Oceania who speak Austronesian languages with those of various populations of Native Americans and Asians showed conclusively that Austronesians spread from South China and the neighboring areas. This conclusion is

independently confirmed by linguistic data (since three primary branches of Austronesian languages are still spoken on Taiwan), as well as by archaeological findings (since Oceania was uninhabited prior to the arrival of speakers of Austronesian languages, archaeologists can trace the direction and approximate dates of arrival of the first inhabitants). However, there are also notable instances of mismatches between population groupings arrived at by genetic methods and those established by linguistic classification of languages spoken by different populations. For example, analyses of genetic variation among the Armenians and the Azerbaijani suggest a very close relatedness between these two populations, although their languages are unrelated (Armenian being an Indo-European language, while Azerbaijani belongs to the Turkic family). Similarly, the Chechens and the Ingush speak closely related languages (belonging to the Nakh-Dagestanian family), but genetic evidence suggests that there was relatively little intermarriage between them until fairly recently. This suggests that one of these two groups underwent a language shift in prehistory, adopting the language of a different group with which it shared a relatively low number of common ancestors.

## **2. Models of language spread**

There are several ways in which a language may spread over a territory, the simplest one being the immigration of a population to a previously uninhabited land, whereby the language of that population is established in the given territory. This is what happened in Iceland, when it was colonized by the Norsemen in the Early Middle Ages. Modern Icelandic developed from the Old Norse language of the first colonists, with relatively little interference from any other language. Another possibility is the wholesale extermination of the speakers of one language, whereby the language of the population responsible for the extermination establishes itself in a given territory. This is what happened in Tasmania, where nearly all of the speakers of indigenous languages were killed by European colonists in the 19th century. However, most processes of language spread involve the sociolinguistic interaction of two populations, originally speaking different languages, only one of which eventually prevails in a given territory. In contemporary linguistics two models of language spread have been proposed. They are called "the wave of advance model" and "the elite dominance model".

### **2.1. Wave of advance**

The way a language spreads according to this model is by a large increase of the number of its speakers. Large population growth is usually made possible by an innovation in food production, e. g. when a hunter-gatherer population adopts agriculture. Population pressure is then likely to cause territorial expansion of that population, and its language will consequently spread. Neighboring populations, e. g. hunter-gatherers who fail to adopt the new food-producing technology, will slowly be pushed to areas less useful for agriculture, and their languages will either die out, or they will be preserved as tiny islands in the large area where the language of the agriculturalists predominate. That language will inevitably evolve and split into a number of genetically related languages, representing a language family.

British archaeologist Colin Renfrew has argued that the largest language family of Western Eurasia, the Indo-European family, expanded with the introduction of

agriculture to Europe from the Middle East in the 7th millennium BC. The Anatolian branch of Indo-European (languages such as Hittite and Luvian) would, according to this theory, be closest to the original homeland of Proto-Indo-Europeans, in present-day Turkey. However, this thesis is rejected by most linguists, who believe that Proto-Indo-European was spoken several millennia after the arrival of the earliest agriculturalists to Europe. Nevertheless, the wave of advance model can very probably be used to explain other spreads of large language families that were made possible by the spread of agriculture, e. g. the spread of the Sinitic languages, such as Mandarin, over much of today's China, or the spread of Bantu languages in central and southern Africa.

## **2.2. Elite dominance**

The population whose language prevails in a certain territory does not have to be more numerous, nor more technologically advanced. It can represent a minority in the given territory, but a minority whose language possesses prestige, so that the members of the majority strive to acquire it. In that case, only a small number of powerful individuals can migrate from one territory to another in order to effect, after a certain period of time, major language shifts. This is what happened in present-day Hungary, which was conquered by the Hungarians in the second half of the 9th century. Although the number of newcomers was relatively small in comparison with the original inhabitants of the country, who spoke Slavic, Hungarians established themselves as aristocracy, and the subjugated Slavs eventually adopted Hungarian. This model of language shift has been named "elite dominance model".

The distinction between the two models of language spread is not absolute, and most historical processes should probably be described as involving elements of both models. For example, the spread of Latin during Roman Empire included the colonization of large territories by Latin-speaking veterans, who expropriated the earlier inhabitants, and this process is best seen as a wave of advance. On the other hand, the Roman society was able to linguistically absorb a large part of the pre-Roman populations in certain provinces, and this was made possible through the social prestige of the Roman social institutions, and the relatively high level of social mobility for all free members of the society (including freed slaves). The fact that any free person could join the Roman army, in which the official language was Latin, contributed to the spread of the language in the provinces where the legions were stationed. This process can best be viewed as instating the elite dominance model, since the actual number of Latin-speaking immigrants was rather low in some areas where Latin subsequently prevailed (e. g. in present day Romania).

It must be stressed that both of the current models of language spread have been used to explain the origin and spread of the major language families of the world. However, nearly all of those families spread only within the last few millennia, i. e. in the periods when advanced techniques of food production and transportation allowed for population growth and mobility that were unthinkable in the palaeolithic period, when most of the Earth's surface was originally populated by humans. It is at present unknown to what extent are those two models applicable to situations of language spread obtaining during the palaeolithic period.

-  
-  
-

TO ACCESS ALL THE 11 PAGES OF THIS CHAPTER,  
Visit: <http://www.eolss.net/Eolss-sampleAllChapter.aspx>

### **Bibliography**

Cavalli-Sforza, L. L. (2001). *Genes, Peoples, and Languages*, Harmondsworth [This is a popular account of the results of population genetics and their correlations with comparative linguistics]

Diamond, J. (1999). *Guns, Germs, and Steel*, London.

Mallory, J. P. *In Search of the Indo-Europeans*, London [This book gives a readable and well-researched overview of the problem of Indo-European homeland, and also provides an useful introduction to the methods of linguistic palaeontology]

Nettle, D. (1999). *Linguistic Diversity*, Oxford.

Nichols, J. (1992). *Linguistic Diversity in Space and Time*, Chicago.

Renfrew, C. (1987) *Archaeology and Language*, London. [In this book the author argues that Indo-European languages spread as a result of the population boom caused by the adoption of agriculture in the Neolithic period.]

Ruhlen, M. (1991). *A Guide to the World's Languages*, London. [This is the most reliable source for the genetic classification of languages, especially at the lower levels of classification. It should be consulted critically when dealing with the problems of deep genetic relationships of languages]

Wells, S. (2003). *The Journey of Man. A Genetic Odyssey*, Harmondsworth. [This is a popular synthesis of the new revolutionary methods of research in population genetics]

### **Biographical Sketch**

**Ranko Matasović** was born in Zagreb in 1968. He studied linguistics and philosophy in the University of Zagreb, where he also got his BA (1992) and his Ph.D (1995).

He is presently Full Professor of comparative Indo-European linguistics in the University of Zagreb. He was a Fulbright scholar in the University of Wisconsin (Madison) in 1997-8, and a Humboldt Research Fellow in the University of Bonn in 2002-3. He published eight books, including "A Theory of Textual Reconstruction in Indo-European Linguistics" (Frankfurt a/M: Lang 1996), "Gender in Indo-European" (Heidelberg: Winter 2004), and "Jezična raznolikost svijeta (Language diversity of the World)" (Zagreb: Algoritam 2005). His main scientific interests are comparative and historical linguistics, Celtic languages and literatures, and language typology.

Prof. Matasović is a member of The Society of Indo-European Studies (Wiesbaden) and the Croatian Philological Association (Zagreb).