

COMPARATIVE AND HISTORICAL LINGUISTICS

Ranko Matasović

Department of Linguistics, University of Zagreb, Croatia

Keywords: Genetic linguistics, language typology, areal linguistics, language families, proto-language, language change, sound law, lexical diffusion, isogloss, implicational universal.

Contents

1. Introduction
 2. Historical Overview
 - 2.1. The Early History
 - 2.2. The Nineteenth Century
 - 2.3. The Twentieth Century
 3. Genetic Linguistics
 - 3.1. Principles of Language Change
 - 3.2. Models of Genetic Relatedness
 - 3.3. Proofs of Genetic Relatedness
 - 3.3. The Major Language Families of the World
 4. Typological and Areal Linguistics
 - 4.1. Typological Classification
 - 4.2. Areal Classification
- Glossary
Bibliography
Biographical Sketch

Summary

This chapter begins with a brief survey of the comparative research into language diversity and development. Historical linguistics is presented as the scientific study of language change. Language change affects all levels of language structure, and it eventually leads to language split, or creation of languages-descendants from common proto-languages.

The discovery of common proto-languages is the main object of genetic comparative linguistics, which classifies languages into language families. Several models of genetic relatedness of languages are discussed, as well as the methods of proof of genetic relatedness. A brief genetic classification of major language families of the world is included.

Typological study of language is concerned with assessing the structural features according to which languages may differ. Languages sharing several logically independent features constitute a language type. Finally, areal comparative linguistics classifies languages into language areas, sets of languages that influenced each other during periods of intensive language contact. Several language areas of the world are enumerated and briefly discussed.

1. Introduction

Comparative and historical linguistics are often treated as a single discipline, although they actually differ considerably with respect to their goals and methods. Comparative linguistics is the scientific study of language from a comparative point of view, which means that it is involved in comparing and classifying languages. To compare languages is to discover the features they share, while the classification of languages proceeds by discovering the relevant defining principles for various classes of languages. Languages can be compared and classified according to three different principles: genetic, typological, and areal. The basic unit of genetic classification is the *language family*, the set of languages for which it can be proved that they developed from a single ancestor, called the *proto-language* of that family. The notion of *proof* of genetic relatedness is crucial here, because all human languages might, or might not be ultimately derived from a single proto-language. The basic unit of areal classification is the *language area* (the German term *Sprachbund* is also sometimes used). It denotes the set of languages for which it can be shown that they developed a number of features as a consequence of mutual contacts. Finally, the basic unit of typological classification is the *language type*, which refers to the set of languages that share some typologically relevant set of features. What "typologically relevant" means here will be explained below.

Historical linguistics is the historical study of language change and development. Its results are directly relevant to comparative linguistics, because only by taking into account the history of languages can we understand why some of them share some of the features they do. This can be for one of the three following reasons: 1) because they stem from some common source, in which case we speak about *genetic relatedness* of languages; 2) because they influenced each other during periods of intensive language contact, in which case we speak of *areal affiliation* of languages, and 3) because their failure to share the features in question would violate some basic and non-obvious principles determining the structure of a possible human language; in that case we claim that languages are *typologically related*, or that they belong to the same linguistic type. In what follows, we shall consider all three of these instances of linguistic relatedness, and examine the methods for discovering them.

2. Historical Overview

2.1. The Early History

Although they made some interesting contrastive remarks about the grammars of Greek and Latin, classical grammarians did not show any interest in comparing languages systematically. The chief reason for this was the fact that for Greeks and Romans the study of language was not a theoretical discipline, concerned with explanations, but rather a practical one, whose primary task was to provide grammatical descriptions of the written language used by culturally important authors. Therefore, the study of barbarians' languages was not considered as a worthwhile objective. It was not until the interest in European vernaculars was aroused during the late middle ages that comparative approaches to language really took off. Dante Alighieri (1265-1321) was the first to attempt a classification of European languages of his time. In his work *De vulgari eloquentia* ("On the Vernacular Speech") he clearly distinguished between

Greek, on the one hand, and the Slavic, Germanic, and Romance languages, on the other; he was also fully aware of the fact that languages diverge over time and that dialectal differences arise because different changes occur in various areas in which a single language is spoken. While Dante used the words for "yes" in order to classify the European languages, Giuseppe Scaligero (1540-1609) used the word for "God", thereby classifying the languages of Europe into "deus-languages" (Latin and the Romance languages) "gott-languages" (the Germanic group), "boge-languages" (the Slavic group), and Greek, in which the word for "god" is *theos*. However, he thought that there was no relationship between these groups of languages, which he called "matrices". On the other hand, Gottfried Wilhelm Leibniz (1646-1716) came very close to recognizing the fundamental relatedness of (Indo-European) languages of Europe, most of which he classified as "Cello-Schytian".

During the Renaissance period and in the 17th and early 18th century, many scholars speculated about the "original language of humankind". Besides Hebrew, which was perhaps the obvious choice, several candidates for that status were advanced, including Chinese (by Webb, in 1669) and Dutch (by Goropius, in 1569). The positive impact of these speculations was that scholars became aware of the scale of language diversity and the ubiquity of linguistic change. The trend toward the accumulation of data about the languages of the world was enhanced by publications of grammars and dictionaries of many languages during the Reformation and Counter-Reformation periods. For example, the first grammar of Basque was published in 1587, the first Polish grammar in 1586, and the first grammars of the American Indian languages Nahuatl, Quechua, and Guaraní were published in 1547, 1560, and 1595, respectively. The encyclopedic movement in the 18th century also contributed to the availability of data about non-European languages. Basic data about several hundred of the world's languages were compiled in Johann Christoph Adelung's (1732-1806) compendium *Mithridates*.

In the eighteenth century information about Sanskrit, the learned language of India, became known among the learned circles in Europe. This was mostly due to the work of Christian missionaries in India, such as the French Pierre de Coeurdoux, or the Croat-Austrian Filip Vezdin (a. k. a. Paulinus a Sancto Bartholomaeo, 1748-1806), who published the first European grammar of Sanskrit. While many scholars had thought that the similarities of major European languages could be explained as the result of language contact, the obvious similarities of basic Sanskrit words with their synonyms in the classical languages required a different explanation. It was highly unlikely that the similarity between, e. g., Sanskrit *pitar-* "father", *mātar-* "mother", and *bhrātar-* "brother" with Latin *pater*, *mater*, and *frater* could have been the result of borrowing. It was not long before William Jones (1746-1794) proposed that Sanskrit, Greek, Latin, and several other languages we now call Indo-European, had "sprung from some common source, which, perhaps, no longer exists." In his programmatic lecture before the Asiatic Society in Calcutta in 1786, which became widely known in Europe, he also emphasized that the similarities between Sanskrit and the classical languages were not limited to the similar shapes of words, but also extended to grammar. In 1816 the German linguist Franz Bopp (1791-1867) used the correspondences between verbal systems of Sanskrit, Greek, Latin, and several other Indo-European languages to prove their genetic relatedness, and somewhat later Jakob Grimm (1785-1863) established the sound correspondences between the consonants of Germanic and those of the other Indo-European languages. These correspondences, which subsequently became known

as "Grimm's law", include the rule that voiced stops in Latin and Greek correspond to voiceless stops in Germanic, while the voiceless stops in the other Indo-European languages correspond to Germanic voiceless fricatives, hence, e. g., Latin *decem* and Greek *déka* "ten" fully match Gothic *taihun*. All of these words can be derived from Proto-Indo-European *dek'm (unattested forms are conventionally marked with an asterisk).

Even somewhat before the publication of the works of Grimm and Bopp, the genetic relatedness of the Uralic languages (Finno-Ugric and Samoyed) was proved by the Hungarian scholar Sámuel Gyarmathi (1751-1830). During the same period, the comparative study of several language families was established by using the same methods as those employed in Indo-European linguistics. These include the Semitic languages (now recognized as a branch of the Afro-Asiatic family), which was discovered and named by Friedrich von Schläzer in 1781, and Dravidian, suggested by Francis W. Ellis in 1816, but proved to be a valid genetic family in 1856 by Robert A. Caldwell. All of those scholars used the same methods as Bopp, Grimm, and the early Indo-Europeanists.

2.2. The Nineteenth Century

The search for the genetic relationships among the world's languages continued without interruption throughout the nineteenth century, and it is fair to say that by the middle of the 20th century, with Joseph Greenberg's masterly classification of the languages of Africa into just four genetic groupings (Afro-Asiatic, Niger-Kordofanian, Nilo-Saharan, and Khoisan languages), most of the now undisputed language families of the world were discovered. However, the major advances in the methodology of historical and comparative linguistics were developed in the field of Indo-European studies. During the 1860's August Schleicher (1821-1868), influenced by the evolutionary biology, introduced the genealogical tree-diagrams into comparative linguistics; in this model, genetically related languages are represented as nodes on a *genealogical tree*, in whose root is the common proto-language of that family. Schleicher also made the first attempts to reconstruct the Indo-European proto-language by applying the comparative method. The early optimism of this project can be seen in the fact that he even composed a fable in the reconstructed Proto-Indo-European language. However, most of his reconstructions are nowadays rejected, or thoroughly revised. Schleicher's tree-model of genetic relationships has also been criticized as simplifying too much the real complexities involved in the development of languages. An alternative model was proposed by Johannes Schmidt (1843-1901), who stressed that boundaries between descendants of a proto-languages are constantly shifting, because linguistic innovations spread like waves, never stopping at exactly the same limits. Schmidt's model was subsequently called the *wave-model* of genetic relationships.

A major breakthrough in the development of comparative and historical linguistics was achieved during the 1870s, when a group of young German scholars, gathered mostly at the University of Leipzig, began their systematic researches in the history of Indo-European languages and the reconstruction of Proto-Indo-European. They were called, somewhat mockingly, "Neogrammarians" (German *Junggrammatiker*), by their elder colleagues, but the name was soon accepted by the leaders of the movement: August

Leskien (1840-1916), Hermann Paul (1846-1921), Karl Brugmann (1849-1919), Berthold Delbrück (1842-1922), Hermann Osthoff (1847-1909), and others. Neogrammarians were profoundly influenced by the development of natural sciences in the second half of the 19th century, and the discovery of empirically established natural laws in sciences like physics and chemistry. The main methodological principle advocated by the Neogrammarians was that language development can be described by empirically founded, but refutable, "sound laws". A sound law is a rule which states that, if a sound A changes to B in one phonetic environment in one word, then the change of A to B will occur in all words of the language in question, in the same phonetic environment. For example, in Greek every word initial Proto-Indo-European *s became *h* before a vowel, so we have the regular correspondence sets: Greek *heptá* "seven" vs. Latin *septem*, Greek *háls* "salt" vs. Latin *sal*, Greek *hýlē* "wood" vs. Latin *silva*. According to a very influential paper by Leskien, "sound laws do not have exceptions"; apparent exceptions can always be explained as results of *analogy*, the principle by which irregular sound changes sometimes occur under the influence of some regular pattern. For example, the final consonant *-s* analogically changed to *-r* in the Latin word *arbos* "tree", which became *arbor*, under the influence of the genitive singular *arboris*, where the change of *-s-* to *-r-* is the outcome of a regular sound law (called "rhotacism"). The neogrammarian doctrine about the exceptionlessness of sound laws was reinforced by their discovery that many exceptions to sound correspondences, discovered by earlier generations of linguists, can be explained as instances of other sound laws operating in specific environments. For example, the apparent exception to Grimm's law seen in the Gothic word for "father", *fadar* vs. Greek *patér*, Latin *pater*, was explained by the Danish linguist Karl Verner (1846-1896), who proved that Proto-Indo-European voiceless stops (in this case *-t-) regularly developed into Germanic *-d-* word-medially, unless preceded by an accented syllable. This rule subsequently became known as Verner's law in Germanic linguistics.

Language typology was initiated as a linguistic discipline in the works of August Wilhelm Schlegel (1767-1845), who divided the world's languages into the following types: 1. *isolating languages*, such as Chinese, in which words do not change (take affixes); 2. *agglutinating languages*, such as Turkish, in which words contain a number of affixes, each of which has a single grammatical function, and 3. *inflectional languages* (such as Latin), in which words can take affixes expressing several grammatical functions (for example, the ending *-i* in the Latin form *vidi* "I saw" expresses the first person, singular, and perfect simultaneously). This morphological typology (so-called because it takes the morphological structure of words as the classifying feature) was subsequently refined by Wilhelm von Humboldt (1767-1835), who is also credited with the view that the structure of a language is influenced by the world view (*Weltanschauung*) of its speakers. He added the fourth morphological type to Schlegel's taxonomy, namely the *incorporating languages*, such as Inuit (Eskimo). In such languages the distinction between a clause and a word is blurred, since, e. g., direct objects can be "incorporated" into the verb. In the twentieth century the morphological typology of languages was thoroughly revised by Edward Sapir (1884-1939). In contrast to earlier language typologists, who ranked languages from "primitive" to "perfect" (whereby the Indo-European languages were almost always considered as the most perfect of all) Sapir freed linguistic typology from value judgments, treating all languages as equally valuable and revealing important aspects of the human mind.

Areal approach to linguistic comparison was still undeveloped in the nineteenth century, largely because of the neogrammarians' insistence on strictly genealogical models in historical linguistics. The predecessors of the areal approach are linguists such as the Austrian Paul Kretschmer (1866-1956) and the Italian Matteo Giulio Bartoli (1873-1946), who investigated the influence of pre-Roman and pre-Greek languages on Latin and Greek, respectively. The work of Bartoli and the Italian school of "neolinguistics" focused on the role of *substrates*, or languages previously spoken in some area, in the development of languages that replaced them in the given area (*superstrates*). Kretschmer contributed to the development of *linguistic palaeontology* by attempting to correlate the earliest historical and archaeological evidence for migrations of the speakers of Greek dialects with lexical evidence of the Greek language. This approach was also implemented in the search for the earliest homeland of the speakers of Proto-Indo-European, which continues until today. The principle of this line of research is to compare the meanings of culturally important words in reconstructed proto-languages with their probable referents, known from the archaeological and historical record, and search for their areal distribution. For example, it was argued that Proto-Indo-European homeland must have been situated in Europe, because we can reconstruct the Proto-Indo-European word for "beech", *b_heh₂g'os (Latin *fagus*, English *beech*, etc.), and beech does not grow to the east of the line connecting the Crimea with Königsberg. Although such arguments do not always yield conclusive results, they lead to interesting correlations between linguistic, historical, and archaeological data, which play an important role in contemporary inter-disciplinary approaches to language prehistory.

Finally, at the End of the 19th century, Hugo Schuchardt (1842-1928) gave an important impetus to the development of areal comparison of languages by his pioneering studies of pidgin and creole languages. He also argued against several doctrines proposed by Neogrammarians, including the exceptionlessness of sound laws, and stressed the importance of studying language change in its social and cultural context.

-
-
-

TO ACCESS ALL THE 26 PAGES OF THIS CHAPTER,
Visit: <http://www.eolss.net/Eolss-sampleAllChapter.aspx>

Bibliography

Anttilla, (1989) Raimo *Historical and Comparative Linguistics*, John Benjamins: Amsterdam [This is one of the best general introductions to historical and comparative linguistics]

Beekes, Robert S. P. (1995) *Comparative Indo-European Linguistics*, John Benjamins: Amsterdam [This is a good introduction to the comparative grammar of Indo-European languages]

Comrie, Bernard (1999) *Language Universals and Linguistic Typology. Syntax and Morphology*, Basil Blackwell: Oxford [This is a classic introduction to methods of linguistic typology]

Haspelmath, Martin, König, Ekkehard, Oesterreicher, Wulf, and Raible, Wolfgang (eds.) (2001) *Language Typology and Language Universals, Vols I-II*, Walter de Gruyter: Berlin & New York. [All of

the major issues in contemporary linguistic typology are discussed in various contributions to this handbook]

Dixon, Robert M. W. (1997) *The Life and Death of Languages*, Cambridge University Press: Cambridge [This book offers a challenge to the traditional methods of comparative and historical linguistics]

Hopper, Paul J. and Elizabeth C. Traugott (1993) *Grammaticalization*, Cambridge University Press: Cambridge [This is the introduction to grammaticalization, or the study of the change of lexical to grammatical elements in language]

Labov, William (1994) *Principles of Linguistic Change. Internal Factors*, Blackwell: Oxford [In this book the principles of sociolinguistic approach to language change are exposed]

Meillet, Antoine (1925) *La méthode comparative en linguistique historique (The Comparative Method in Historical Linguistics)*, Nygaard: Oslo (reprint: Paris 1970) [This is a classic work on the methods used by comparative and historical linguists]

Nichols, Johanna (1992) *Linguistic Diversity in Space and Time*, The University of Chicago Press: Chicago [This book marked the beginning of modern approach to areal typology and the problem of macroareas]

Ruhlen, Merritt (1987) *A Guide to the World's Languages*, Edward Arnold: London [This is the standard book about the genetic classification of the world's languages; however, some of its claims about deep-level genetic relationships are too speculative]

Sandfeld, Kristian (1930) *Linguistique balkanique. Problèmes et résultats (Balkan Linguistics: Problems and Results. In French)*, Société de Linguistique de Paris: Paris. [This is the classic work in which the features of the Balkan linguistic area are presented]

Trask, R. L. (1996) *Historical Linguistics*, Edward Arnold: London [This is a modern introduction to the subject, encompassing some recent developments in the field of historical linguistics.]

Biographical Sketch

Ranko Matasović was born in Zagreb in 1968, where he studied linguistics and philosophy. In 1995, he received his Ph. D. from the University of Zagreb, where he has been lecturing on comparative Indo-European linguistics and language typology. He is the Head of the Chair of the Comparative Indo-European Linguistics in the University of Zagreb since 1996, and he was the Head of the Department of Linguistics in the University of Zagreb from 1998 until 2000. In 2002 he received the Award of the Croatian Academy of Sciences and Arts for a lasting contribution to philology. He was a Fulbright Fellow in the University of Wisconsin (Madison) in 1997, and a Humboldt Research Fellow in the University of Bonn in 2002-2003. In 2005 he was granted full professorship in the Department of Linguistics in the University of Zagreb, and in 2006 he became an associate member of the Croatian Academy of Sciences and Arts.