Chamonikolasová, Jana

Proto-Indo-European

2 Proto-Indo-European

English belongs to the family of the Indo-European languages. This family includes most current languages spoken in the geographical area of Europe, the Indian Subcontinent, the Iranian plateau, and Asia Minor (Anatolia). Some of the Indo-European languages have spread to distant continents and are now spoken as a native language also in America and Australasia. The Indo-European language family consists of over four hundred languages and dialects, and it is probably the largest language family so far recognized in terms of number of native speakers (over 3 billion).

2.1 The common ancestor of Indo-European languages

Similarities between languages of the Indo-European family were noticed by several missionaries, merchants, travelers, and scholars in the 16th, 17th, and 18th centuries (Thomas Stephens, Filippo Sassetti, Marcus Zuerius van Boxhorn, Evliya Çelebi, and Mikhail Lomonosov). These observations, however, did not become widely known.

Indo-European studies were initiated and stimulated by a British orientalist, Sir William Jones, who was stationed as judge in Calcutta in the late 18th century. He propagated the observation of the resemblance between Sanskrit, classical Greek, and Latin. In his lecture "The Sanskrit Language" delivered in 1786 and published two years later, he suggested that all three languages developed from a common source. This common source was later named Proto-Indo-European.

In the 19th century, this common source became the focus of attention of a number of studies carried out by the representatives of Indo-European comparative linguistics – Franz Bopp, August Schleicher, Karl Brugmann, Rasmus Rask, Karl Verner, Jacob Grimm, and others. The languages under investigation were referred to as *Indo-European* or *Indo-Germanic* languages.

The Proto-Indo-European language, originally spoken by the Proto-Indo-Europeans, was reconstructed from later stages of development of Indo-European languages. This common ancestor of Indo-European languages was the first proto-language proposed and accepted by diachronic linguists. Scholars investigating the origins of Indo-European languages developed techniques of historical linguistics, which were later applied in the research of other language families (for example the comparative method and the

method of internal reconstruction). The reconstruction of Proto-Indo-European and its daughter languages was the focus of the majority of linguistic studies in the 19th century.

There are several hypotheses of the origin and spread of Proto-Indo-European (see Blažek 1993). According to the most popular model, the Kurgan hypothesis, the Proto-Indo-Europeans lived in the Pontic-Caspian steppe of Eastern Europe in the 4th millennium BC. Their language, Proto-Indo-European, probably split into different languages around 3 500 BC, when the Proto-Indo-Europeans expanded from their original settlement into different parts of Europe and Asia. However, time estimates vary by hundreds of years, and some Indo-European languages may have diverged from the common ancestor before the beginning of the 4th millennium BC.

The description of Proto-Indo-European was completed in the early 20th century; however, diachronic research continued and some refinements of the earlier reconstructions and hypotheses have been accepted more recently. The most important advancement in the field of Indo-European historical linguistics in the 20th century was the discovery of Anatolian and Tocharian languages, which resulted in the re-evaluation of some linguistic features shared by Indo-European languages, and led to the acceptance of the laryngeal theory proposing the existence of Proto-Indo-European laryngeal consonants that later disappeared completely from all Indo-European languages except the Anatolian languages.

A number of recent studies propose a relationship between Proto-Indo-European and other language families, for example the Uralic and Altaic languages. The proposed relationships, however, remain controversial.

Since there are no written records of Proto-Indo-European, all descriptions of the language are based only on reconstruction methods and identify hypothetical language features.

2.2 Branches of Indo-European languages

Scholars have presented different patterns of internal division of Indo-European languages (see Blažek 2012). One of the commonly accepted classification systems is a division into ten major branches. In the survey below, these branches are arranged according to the chronological order of their emergence as presented by Anthony (2007):

1. **Anatolian** (emerged around 4200 BC, located in Asia Minor, extinct)
2. **Tocharian (**emerged around 3700 BC, located in China, extinct)
3. **Germanic** (emerged around 3300 BC, earliest runic inscriptions from around the 2nd century AD, earliest coherent texts (the translation of the Bible into Gothic by Wulfila) from the 4th century AD. Old English manuscript tradition from about the 8th century AD)

4. **Italic** (including Latin and the Romance languages, emerged around 3000 BC)
5. **Celtic (**emerged around 3000 BC)
6. **Armenian** (emerged around 2800 BC)
7. **Balto-Slavic** (emerged around 2800 BC)
8. **Hellenic** (emerged around 2500 BC)
9. **Indo-Iranian (**emerged around 2200 BC)
10. **Albanian** (attested from the 14th century AD)

In addition to the languages listed above, several other Indo-European languages have existed: Illyrian, Venetic, Liburnian, Messapian, Phrygian, Paionian, Thracian, Dacian, Ancient Macedonian, Ligurian, Sicel, Lusitanian, and Cimmerian. These languages are all extinct and current knowledge of their origin and development is rather limited.

The table below presents a more detailed division of the Indo-European language branches listed above. Different classification systems, however, differ in the division of the Indo-European language family into branches and subgroups, as well as the lists of members of the subgroups. In the survey below, some extinct languages have been omitted.

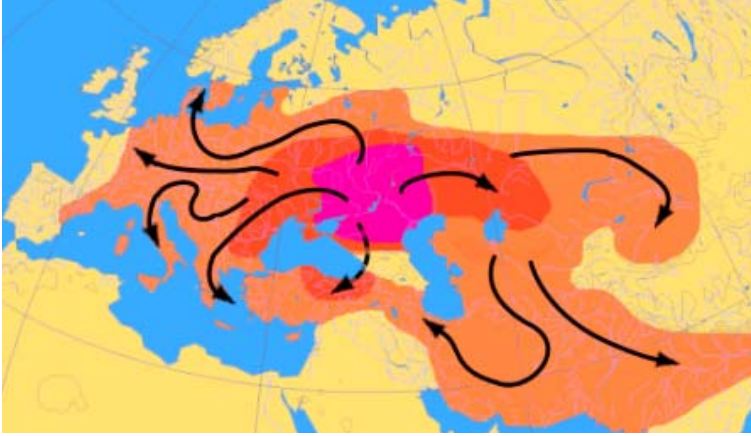**Branches of Indo-European languages**

	Indo-European Branches		**Indo-European Languages**
1.	Anatolian		Hittite, Luwian, Lydian
2.	Tocharian		Tocharian
3.	Germanic	East	Gothic
		North	Eastern Group: Swedish, Danish, Norwegian Bokmål Western Group: Icelandic, Norwegian Nynorsk
		West	Anglo-Frisian Group: English, Frisian Low Germanic Group: Flemish, Dutch, Afrikaans High Germanic Group: German, Yiddish
4.	Italic (< Latin)	East	Romanian
		Central	Italian, Sardinian
		West	French, Catalan, Spanish, Portuguese
5.	Celtic	Continental	Gaulish, Lepontic, Celtiberian
		Insular	Goidelic Group: Irish Gaelic, Scots Gaelic, Manx Brythonic Group: Cumbrian, Welsh, Cornish, Breton
6.	Armenian		Armenian

7.	Balto-Slavic	Baltic	Latvian, Lithuanian
		Slavic	Eastern Group: Belarusian, Russian, Ukrainian Southern Group: Bulgarian, Macedonian, Serbian, Croatian, Slovene Western Group: Polish, Slovak, Czech, Upper & Lower Sorbian
8.	Hellenic		Greek
9.	Indo-Iranian	Indo-Arian	Sanskrit, Bengali, Hindi, Marathi, Panjabi
		Iranian	Avestan, Pashto, Persian, Kurdish
10.	Albanian		Albanian

All the languages above are genetically related since they are all descendants of one parent language, Proto-Indo-European. The division into the individual branches and subgroups is based on genetic principles; however, an important criterion of the subdivision are innovations shared by several languages whose common ancestor originally split off from the other parent languages descending from Proto-Indo-European. Germanic languages, for example, share phonological and grammatical features determined by innovations that seem to have developed in their common parent language, Proto-Germanic, and that distinguish them from members of other branches.

Indo-European languages are traditionally divided into centum and satem languages according to the development of the palatal plosives. The terms centum and satem correspond to the expression *hundred* in Latin *(centum* [kentum]), and in Old Iranian/ Avestan (*satem* [satəm]). In centum languages, palatal plosives merged with velar plosives, therefore palatal *ḱ* changed into k. In satem languages, palatal plosives changed into sibilants, therefore *ḱ* changed into s. (Velar plosives in satem languages merged with labio-velar plosives.) English belongs to centum languages, together with other Germanic languages and with the Analolian, Tocharian, Italic, Celtic, and Hellenic language families; languages of the remaining Indo-European branches, i.e. Balto-Slavic, Armenian, Indo-Iranian, and Albanian, are satem languages.

The charts below illustrate Indo-European migration and the geographic distribution of Indo-European languages in different historical periods. Before the 16th century, Indo-European languages were located in Europe, and South, Central and Southwest Asia. Today, they are distributed worldwide.

The chart above illustrates the Migration of Indo-European tribes between ca. 4000 and 1000 BC according to the Kurgan hypothesis. The assumed original homeland (Urheimat) of the Indo-Europeans is the purple area of Samara/Sredny Stog culture north of the Black Sea and the Caspian Sea. By ca. 2500 BC, the Indo-Europeans probably spread to the area marked red; and by 1000 BC, they settled the orange area.

Source: Dbachmann. Indo-European expansion 4000–1000 BC, according to the Kurgan hypothesis. Wikimedia Commons [online] 3. 10. 2005. [accessed 2014-07-11]. Under the license CC BY-SA 3.0 (see http://creativecommons.org/licenses/by-sa/3.0/legalcode). Available at: http://commons.wikimedia.org/wiki/File:IE_expansion.png

Indo-European languages ca. 3500 BC

Source: Dbachmann. Indo-European languages ca. 3500 BCE. Wikimedia Commons [online] 9. 3. 2005. [accessed 2014-07-11]. Under the license CC BY-SA 3.0 (see http://creativecommons.org/licenses/by-sa/3.0/legalcode). Available at: http://commons.wikimedia.org/wiki/File:IE5500BP.png

Indo-European languages ca. 2500 BC

Source: Dbachmann. Indo-European languages ca. 2500 BCE. Wikimedia Commons [online] 8. 3. 2005. [accessed 2014-07-11]. Under the license CC BY-SA 3.0 (see http://creativecommons.org/licenses/by-sa/3.0/legalcode). Available at: http://commons.wikimedia.org/wiki/File:IE4500BP.png

Indo-European languages ca. 1500 BC

Source: Dbachmann. Indo-European languages ca. 1500 BCE. Wikimedia Commons [online] 8. 3. 2005. [accessed 2014-07-11]. Under the license CC BY-SA 3.0 (see http://creativecommons.org/licenses/by-sa/3.0/legalcode). Available at: http://commons.wikimedia.org/wiki/File:IE3500BP.png
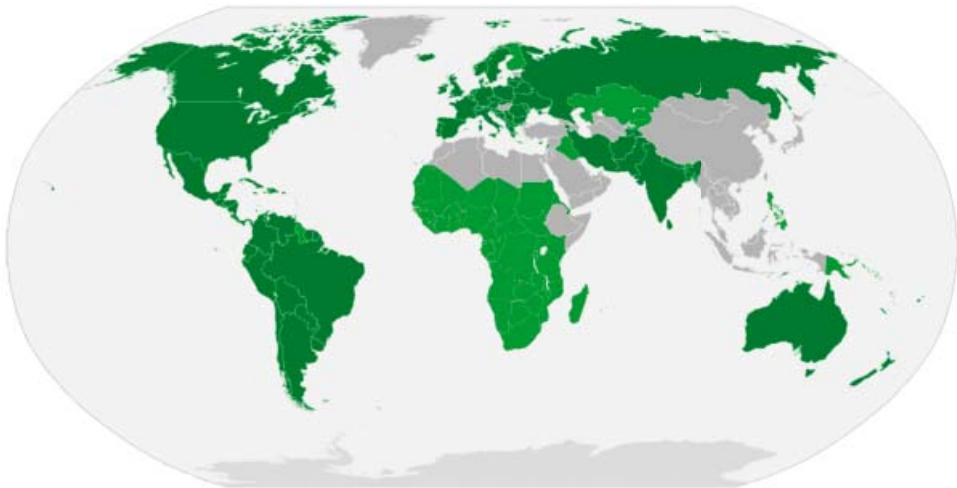

Indo-European languages ca. 500 BC

Source: Dbachmann. Indo-European languages ca. 500 BCE. Wikimedia Commons [online] 8. 3. 2005. [accessed 2014-07-11]. Under the license CC BY-SA 3.0 (see http://creativecommons.org/licenses/by-sa/3.0/legalcode). Available at: http://commons.wikimedia.org/wiki/File:IE2500BP.png

Indo-European languages ca. 500 AD

Source: Dbachmann – Spiridon Ion Cepleanu. Indo-European languages ca. 500 AD. Wikimedia Commons [online] 8. 12. 2013. [accessed 2014-07-11]. Under the license CC BY-SA 3.0 (see http://creativecommons.org/licenses/by-sa/3.0/legalcode). Available at: http://commons.wikimedia.org/wiki/File:IE1500BP.png

Present-day distribution of Indo-European languages
- ■ Countries with a majority of speakers of IE languages
- ■ Countries with an IE minority language with official status

Source: Brianski – Alphathon. Indo-European languages ca. 500 BCE. Wikimedia Commons [online] 6. 6. 2013. [accessed 2014-07-11]. Available at: http://commons.wikimedia.org/wiki/File:IE_countries.svg

Present-day distribution of Indo-European languages in Europe

Source: Iolaafi. Indo-European languages in Europe. Wikimedia Commons [online] 9. 11. 2013. [accessed 2014-07-11]. Under the license CC BY-SA 3.0 (see http://creativecommons.org/licenses/by-sa/3.0/legalcode). Available at: http://commons.wikimedia.org/wiki/File:Indo-European-languages.gif

2.3 Main grammatical features of the Proto-Indo-European language

The reconstructed Proto-Indo-European language was a synthetic language with a rich inflectional morphology. Its inflectional system consisted of a large number of different inflectional paradigms. Most words consisted of three parts: root + stem suffix + ending. The root, together with the stem suffix formed the *stem*. The stem suffix is often referred to as the *theme* and the presence or absence of the theme in a word paradigm is the basis for the distinction made between *thematic* and *athematic* nouns and verbs. Athematic words seem to belong to the oldest Proto-Indo-European word stock. Some of them can be recognized in modern Indo-European languages by certain types of irregularities in their inflectional paradigms. The majority of roots of Proto-Indo-European words were probably

monosyllabic and consisted of a sequence of three sounds: consonant + vowel + consonant (and possibly another consonant). An important phenomenon of Proto-Indo-European morphophonology, which later played an important role also in Germanic languages, is the variation in the root vowels usually referred to as *ablaut* or *vowel gradation*. The term vowel gradation, however, is wider than the term ablaut; it denotes all types of changes in the root of a word including *umlaut*. While ablaut was determined by the position of accent within different grammatical forms of a word, umlaut was caused by the influence of the vowel in the next syllable.

The Proto-Indo-European inflectional system of nouns included eight or nine cases and two basic types of declension (thematic and athematic). The nominal grammatical system included three genders: masculine, feminine, and neuter.

The main categories of Proto-Indo-European verbs distinguished according to their aspect were stative, imperfective, and perfective. The verbal system included several grammatical moods and voices and conjugation according to person, number and tense. By adding affixes to the base form of a verb (its root), new word could be created (especially new nouns, verbs, or adjectives). The Proto-Indo-European verbal grammatical system is probably best preserved in Ancient Greek and Vedic Sanskrit. Verbal conjugation paradigms included a variety of endings and were subject to ablaut. Ablaut is still visible in Germanic languages including English (e.g. in drive – drove – driven; ring – rang – rung).

Bibliography

The chapter above draws on the following sources:

Anthony, David W. (2007). *The Horse, the Wheel, and Language: How Bronze-Age Riders from the Eurasian Steppes Shaped the Modern World*. Princeton University Press.

Baugh, Albert C. and Cable Thomas (1993). *A history of the English language.* Routledge.

Beckwith, Christopher I. (2009). *Empires of the Silk Road*, Oxford University Press, p. 30.

Blažek, Václav (1993). Po stopách indoevropské pravlasti. *Slovo a slovesnost* 54, pp. 31–40.

Blažek, Václav (2012). On the Internal Classification of Indo-European Languages: A Survey. *Studia Indogermanica Lodziensia* 7, pp. 9–30.

Vachek, Josef (1994). *Historický vývoj angličtiny.* Brno: Masarykova univerzita.

Vachek, Josef (1991). *English past and present – an introductory course*. Praha: SPN.

Wikipedia:

Indo-European Languages. Wikipedia. The Free Encyclopedia [online] 27. 9. 2014. [accessed 2014-09-27]. Available at: http://en.wikipedia.org/wiki/Indo-European_languages

Indo-European Migrations. Wikipedia. The Free Encyclopedia [online] 9. 8. 2014. [accessed 2014-08-27]. Available at: http://en.wikipedia.org/wiki/Indo-European_migrations

Proto-Indo-European Language. Wikipedia. The Free Encyclopedia [online] 2. 10. 2014. [accessed 2014-08-27]. Available at: http://en.wikipedia.org/wiki/Proto-Indo-European_language

List of Indo-European Languages. Wikipedia. The Free Encyclopedia [online] 28. 9. 2014. [accessed 2014-10-02]. Available at: http://en.wikipedia.org/wiki/List_of_Indo-European_languages