The purpose of the present study is to confront most representative models of the internal classification of Indo-European languages and their daughter branches.

0. Indo-European

0.1. In the 19th century the tree-diagram of A. Schleicher (1860) was very popular:

After the discovery of the Indo-European affiliation of the Tocharian A & B languages and the languages of ancient Asia Minor, it is necessary to take them into account. The models of the recent time accept the Anatolian vs. non-Anatolian ('Indo-European' in the narrower sense) dichotomy, which was first formulated by E. Sturtevant (1942). Naturally, it is difficult to include the relic languages into the model of any classification, if they are known only from several inscriptions, glosses or even only from proper names. That is why there are so big differences in classification between these scantily recorded languages. For this reason some scholars omit them at all.

0.2. Gamkrelidze & Ivanov (1984, 415) developed the traditional ideas:

0.3. Vladimir Georgiev (1981, 363) included in his Indo-European classification some of the relic languages, plus the languages with a doubtful IE affiliation at all:
0.4. Eric Hamp proposed his original model of the Indo-European disintegration, including the relic idioms, based on specific isoglosses in phonology, morphology and lexicon (1990):
0.5. As the illustration of a realistic application of cladistics can serve D. Ringe, T. Warnow & A. Taylor (2002, 87):

0.6. The absolute chronology is available only thanks to glottochronology. The most recent result of Sergei Starostin (Workshop on the chronology in linguistics, Santa Fe 2004) applies his own model of the ‘recalibrated’ glottochronology, where all borrowings are excluded before any calculation and the coefficient of changes is empirically recounted to 5% per millennium instead of 14% postulated by Swadesh.
1. Indo-Iranian
The preliminary result of comparison of Avestan, Vedic, Dardic & Nuristani was proposed by Hegedűs & Blažek (2010):

<table>
<thead>
<tr>
<th></th>
<th>-2700</th>
<th>-2200</th>
<th>-1700</th>
<th>-1200</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indo-Nuristani</td>
<td></td>
<td></td>
<td>-1900</td>
<td></td>
</tr>
<tr>
<td>Indo-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iranian</td>
<td>-2700</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1a. Indo-Aryan
The only attempt to apply glottochronology for several modern Indo-Aryan languages in confrontation with Sanskrit was realized by S. Starostin and his team (database 2004):

<table>
<thead>
<tr>
<th></th>
<th>-1400</th>
<th>-1000</th>
<th>-600</th>
<th>-200</th>
<th>200</th>
<th>600</th>
<th>1000</th>
<th>1400</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vedic Sanskrit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cl. Sanskrit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1b. Dardic
The only tree-diagram was constructed by S. Starostin’s team (database 2004):

<table>
<thead>
<tr>
<th></th>
<th>-1000</th>
<th>-600</th>
<th>-200</th>
<th>200</th>
<th>600</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kashmiri</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shina</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maiya</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bashkarik</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torwali</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wotapuri</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phalura</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sava</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tirahi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Khowar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pashai</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kalasha</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gawar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shumashhti</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2. Nuristani
The first application of glottochronology to Nuristani by Hegedűs & Blažek (2010):

3. Iranian
The only attempt to construct the tree-diagram for the Iranian languages was realized by S. Starostin and his team (Santa Fe 2004).

4. Anatolian
With exception of Hittite no Anatolian language allows to apply glottochronology for our limited knowledge of their lexical corpora. That is why the existing classifications are based on combinations of phonological, morphological and lexical isoglosses. In the recent time these three alternative models of the internal classification of the Anatolian languages were proposed.

4.1. N. Oettinger 1978, 92 (supplemented on the basis of personal communication in 2001):
4.2. R. Werner (1991, 17)

Hieroglyphic Luwian
Cuneiform Luvian
Milyan
Lycian
Palaic
Hittite
Lydian

South Anatolian
Anatolian

4.3. C. aan de Wiel <http://iiasnt.leidenuniv.nl/pie/ielangs/anatolian.html>

Hittite
Palaic

Anatolian
Hittite-Palaic

Luwian
Sidetic
Pisidic
Lycian, Milyan
Carian
Lydian

Southeast Anatolian

5. Greek

The most detailed scheme classifying the Greek dialects was proposed by A. Bartoněk on the basis of phonology and morphology (1987, 104; 2003, 494):

<table>
<thead>
<tr>
<th>2000</th>
<th>1800</th>
<th>1600</th>
<th>1400</th>
<th>1200</th>
<th>1000</th>
<th>800</th>
<th>600 B.C.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>dialects</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Doric proper</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Saronic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Northwest</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Aeolic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Arcado-Cypriote</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ionic-Attic</td>
</tr>
</tbody>
</table>

West Greek

Proto-Doric

Protoaeolic

Achaean

Mycenaean

Protoionic

Elis
Laconia
C. Crete
E. Crete islands
W. Argolis
E. Argolis
Megaris
Corinth
Phocis
Locris
Aetolia
Boeotia
W. Thessalia
E. Thessalia
Lesbos
Arcadia

The application of the ‘recalibrated’ glottochronology gives almost the same result, dating the disintegration of West and East Greek to the 38th cent. BP (Blažek 2010).
Note: Greek can be classified as one of the Hellenic languages, together with Phrygian / Brygian, ancient Macedonian, and perhaps also Messapic, if the hypothesis of M. Huld (1995, 147-55; cf. also Blažek 2009b) is accepted. Unfortunately, the lexical corpora do not allow any quantification.

6. Paleo-Balkanian
Extremely poor data and their ambiguous interpretations lead to various hypotheses. The present author finds as probable following: In Prehellenic = Pelasgian the Lautverschiebung operated; the language was of the centum-type (Hamp) rather than of satem-type (Georgiev). If Thracian & Bithynian were satem-languages with Lautverschiebung, their closer relation with Armenian is expectable (so Kortlandt 1988). Albanian is a descendant of Illyrian, both the satem-languages. The change *gʷ > b (Pisani 1957) in Dacian indicates more probably the centum-type, regarding the complementarity of the *kʷ : *k and *k : *k’ distinction (similarly Witczak, p.c.).

7. Italic
7.0. The ancient Italic languages are only fragmentarily recorded, naturally with exception of Latin. For this reason, their classification cannot use the lexicostatistic methods and so only the qualitative analysis of phonology bring some results. The present model does not reflect any grade of a mutual relationship.

Wallace (1984, 123-51) discusses five scenarios of classification of the Osco-Umbrian languages:

(A) Planta 1892; Sommer 1948; Krahe 1966.
(B) Palmer 1954.
(C) Buck 1928; Kent 1945.
(D) Poultney 1951; Durante 1978.
(E) Conway 1897.

Synthesis of contemporary views on the classification of the Italic languages:
The following model (Urbanová & Blažek 2008, 34) represents an attempt to find a synthesis based on conclusions of various scholars, who are in agreement in a close position of Venetic to the Latino-Faliscan branch, with the idea of H. Rix (2002, 3-9) about an internal structure of the Osco-Umbrian branch. The constitution of a special Ausonian-Sicilian branch is based on a witness of the antique authors, documenting that in past the Sicilians-Ausonians were pushed away from Italy for Sicily (Thukydides; Dionysius Halicarnassensis referring to Hellanicus; see Schmoll 1958, 96). The change *kʷ > p assumed for Siculic connects the hypothetical Ausonian-Sicilian branch with the Osco-Umbrian languages.
7A. Romance

On the other hand, the lexical material of the Romance languages served for determination of the basic constants in glottochronology. Let us confront several models of their disintegration:


Note: Rix (2002) differentiated the following local varieties of Oscan: from Capua, Pompeie, Cetera Campania, Central Oscan, Lucania & Bruttium, Messina.

7A.3. Merritt Ruhlen 1987, 326:

7A.4. It is natural that glottochronology was also applied for Romance languages. Let us confront two attempts from the recent time: Embleton (1986, 142):
8. In the area between Italic and Celtic there were at least two relic languages which could form a closer unit in the genealogical sense: **Ligurian** and **Lusitanian**, former reconstructed on the basis of proper names attested by classical authors in northern Italy, latter known from several inscriptions written in the Latin alphabet, discovered in south Portugal and Spain (cf. Urbanová & Blažek 2008, 178-81; Witczak 2005; Blažek 2006).

9. Celtic
There are two alternative models of disintegration of the Celtic languages.
9.1. The first model has to reflect the opposition between the insular and continental languages. It is defended e.g. by W. Cowgill (1975) or P. Schrijver (1995, 463).

9.2. The alternative and more traditional model is based on the \(\phi/p\)-isogloss in the reflexes of the Indo-European labiovelar \(*k^\prime\). The figures for living languages (plus Cornish) and the age of the divergence of Goidelic vs. Brythonic were calculated by S. Starostin and his team (Santa Fe 2004). The positions of other nodes indicated by question marks represent only rough assessments:
9.3. Applying the Starostin’s approach with one modification, namely the systematic inclusion of synonyms for all sufficiently described languages, i.e. all Goidelic, Brythonic, plus Gaulish, the internal structure of the Celtic tree diagram is in principle the same, the differences occur only in details. Let us mention that the time depth of the divergence of Gaulish vs. Brythonic (1000 BC) is practically the same as Goidelic vs. Gaulish-Brythonic (Novotná & Blažek 2006, 91; Blažek 2009a):

10. Germanic
The best summarization of various ideas concerning the classification of the Germanic languages is the study of W. Mańczak (1992; cf. also Blažek & Pirochta 2004).
10.1. J.Ch. Adelung (1806) divided the Germanic languages into two branches:

10.2. Similarly J. Grimm (1819) operated with the binary classification. For some of the tribal dialects he supposed the transit character, viz. Frisian & Anglian (1-2), Frankish (2-3), Quadic & Marcomanic (3-4). Alternatively he assumed the opposition of East Germanic vs. others.
10.3. Applying his original method based on the lexicostatistic analysis of parallel texts, Mańczak (1992) formulated a similar conclusion. He ordered the languages decliningly according to their relationship with Gothic: the closest has to be Old High German, further Old Saxon, finally Scandinavian languages.

10.4. Another model of the binary classification was presented by K. Müllenhoff (1898):

10.5. The most frequent model divides the Germanic languages into three branches: East, North and West. The author of the following classification is J. Schmidt (1860):

10.6. F. Maurer (1943) tried to depict the development from the tribal Germanic dialects to the languages of the late middle age and present time, including the convergent processes:
10.7. E. Schwarz (1951) assumed that c. 200 B.C. the Germanic language continuum was already divided into the North zone, generating the later Scandinavian languages and Gothic, and the South zone, where the later German dialects were formed. About 4 cent. later the third, transit zone, cristalized, developing in the languages of Angels and Frisians.

10.8. The most detailed scheme of the development of the Germanic languages was proposed by T.V. Toporova (2000), inspired by Maurer and Schwartz:

10.9. E. Antonsen (1975) assumed the opposition of the east and northwest branches:

```
Germanic

Northwest
Old Runic

East

North
West

Gothic
```

10.10. H. F. Nielsen (2000) returned to the traditional idea, identifying in Old Runic a direct ancestor only of the Scandinavian languages:

```
Northwest Germanic

Old Runic

North Sea
Old Frisian

North Sea
Old English

Old Runic

Old Norse

A.D. 100 200 300 400 500 600

West Germanic

Old High German

Old Saxon

Old English

Old Norse
```

10.11. For her classification of the Germanic languages Sheila Embleton (1986, 117) used her modification of glottochronology:

```
100 n. l. 300 500 700 900 1100 1300 1500 1700

Swedish
Danish
Norwegian

Faeroese
Islandic

English

Frisian
Vlamiish
Afrikaans
Dutch

Yidish
Low
German
High

Gothic

189 264 143 873 1047 1236 1425 1664 1224 1379 1531 1812

```

10.12. The most recent attempt to classify the Germanic languages was published by Starostin & Burlak (2001, 82-105). They applied Starostin’s recalibrated glottochronology for 7 literary living languages and Gothic.
11. Baltic

According to tradition, the Baltic languages are divided into the west part represented by Old Prussian, from c. 1700 extinct, and eastern part, represented by the living languages, Lithuanian and Latvian. But the Baltic dialectology was much more complex a millennium ago. The following model was proposed by V. Mažiulis (1981):

```
North periphery
  Zemgalian
  Selian
  Couronian
Baltic
  Latvian
  Lithuanian
  Yatvingian
  Prussian
  Galindian
South periphery
```

The first serious application of the classical glottochronology was used by Lanszweert (1984, xxxii-xxxvii), who has found 58.6% for Prussian vs. Lithuanian and 55.2% for Prussian vs. Latvian. The results of Girdenis & Mažiulis (1994, 9) are lower: 68% Lithuanian vs. Latvian, 49% Lithuanian vs. Prussian, 44% Latvian vs. Prussian. Starostin (Santa Fe 2004 and p.c., June 2005) dated the separation of Lithuanian and Latvian to 80 B.C., Lithuanian and the ‘Dialect of Narew’ to 30 B.C., Latvian and the ‘Dialect of Narew’ to 230 B.C. The position of Prussian in his calculations is rather strange, it has to be closer to Slavic than to Baltic. Novotná & Blažek (2007, 205), calculating the synonyms too, have reached the following results:

```
-1400 -1000 -600 -200 +200 +600
  -1400
  -830 / -730
  56% / 58%
  46.7%
  +190
  76.3%
  +600
  84.8%
  +1200
  94%
  +1550
  +1000
  89-91%
  +700
  94%
  +1200
  89-91%

Swedish
Danish
Nynorsk
Icelandic
English
Dutch
High German

Gothic
```

Common Slavic
The double-result 58/56% for Prussian vs. other Baltic languages reflects the calculation without / with the ‘Dialect of Narew’. The score 43% between Prussian and the ‘Dialect of Narew’ (Pogańskie gwary z Narewu; see Zinkevičius 1984) in confrontation with 62% and 55.2% for Prussian vs. Lithuanian and Prussian vs. Latvian respectively, excludes the identification of the ‘Dialect of Narew’ with the historical Yatwingians, known from the Middle Ages, if their language had to be connected with the other Baltic idioms of the southern periphery, including Prussian. Regarding this big difference, it seems better to accept the explanation of Schmid (1986) who identified in the ‘Dialect of Narew’ a strong influence of Northeast Yiddish, spoken in the big cities of Lithuania and Latvia, hence the hybrid East Baltic - German idiom.

12. Slavic

12.1. According to the traditional model the Slavic languages are divided into three parts (cf. e.g. J. & B. Grimes 1996, 58):

12.2. The classification of the Slavic languages by Starostin (Santa Fe 2004), using his recalibrated glottochronology, is revolutionary in both topology and chronology:
12.3. Using the principles of Starostin’s recalibration of glottochronology, Novotná & Blažek (2007, 201) proposed another model of the internal grouping of the Slavic languages which seems to be in better agreement with historical data and archaeological research:

![Graph showing internal grouping of Slavic languages](image)

13. Tocharian
The beginning of the disintegration between Tocharian A and B can be dated to c. 400 BC according to the glottochronological test (Blažek & Schwarz 2008: §8, Appendix 1).

References:
Blažek, Václav & Schwarz, Michal. 2008. Tocharians. Who they were, where they came from and where they lived. Lingua Posnaniensis 50, 47-74.


Václav Blažek
Ryneček 148
26101 Příbram III
Czech Republic
blazek@phil.muni.cz