Transactions of the Philological Society Volume 110 (2012) 1–28

AREAL TYPOLOGY OF PROTO-INDO-EUROPEAN: THE CASE FOR CAUCASIAN CONNECTIONS

By Ranko Matasović
University of Zagreb

Abstract

This paper re-examines the evidence for early contacts between Proto-Indo-European (PIE) and the languages of the Caucasus. Although we were not able to find certain proofs of lexical borrowing between PIE and North Caucasian, there are a few undeniable areal-typological parallels in phonology and grammar. Some features generally attributed to PIE are not found in the majority of languages of North and Northeastern Eurasia, while they are common, or universally present, in the languages of the Caucasus (especially North Caucasus). Those features include the high consonant-to-vowel ratio, tonal accent, number suppletion in personal pronouns, the presence of gender and the morphological optative and, possibly, the presence of glottalized consonants and ergativity.

1. Introduction[1]

More than a century ago, Dutch linguist Christian Uhlenbeck suggested that Proto-Indo-European (PIE) was substantially influenced by Caucasian languages. He believed that PIE was a language genetically related to Uralic, but that it underwent areal influences from unknown, but presumably Caucasian languages (Uhlenbeck 1901). Like most of Uhlenbeck's output, his hypothesis about PIE and Caucasian connections remained largely ignored, but many of his arguments have been revived over the last couple of decades (for a re-evaluation of Uhlenbeck's work see Kortlandt 2009). Much discussed at the time were the proposals by Trubetzkoy, who in the 1930s claimed that PIE was, in a sense, a mixed language, one component of which came from the Caucasus (Trubetzkoy 1939). Although his denying of the 'purity' of PIE was understandable in the period of intensive ideological debates about the significance of 'Aryan' languages and 'races', Trubetzkoy's ideas were not well received by many linguists. In the 1960's and 1970's, a number of typological parallels between PIE and Kartvelian, chiefly in the verbal system, was investigated by Schmidt (1967). Some of his rather inconclusive findings were gathered by one of his pupils, Shimomiya (1978). In 1984, Gamkrelidze & Ivanov (1984) argued that the PIE homeland was located to the south of the Caucasus mountains, and tried to support this theory by adducing a number of arguments from areal typology, showing structural convergences between PIE and the Caucasian languages (especially Kartvelian). They also discussed a set of lexical isoglosses connecting PIE and Semitic, Kartvelian and ancient Mesopotamian languages, which were, in their opinion, only possible if PIE was spoken south of the Caucasus. However, the majority of

[1] Earlier versions of this paper were read in the Russian State University for the Humanities in January 2011, at the meeting of the Philological Society in Oxford (in March 2011), and in the École Normale Supérieure in Paris (in April 2011). I thank all the audiences who commented on it and helped me express my ideas more clearly, as well as two anonymous reviewers of *TPhS*.

Indo-Europeanists have not found their arguments persuasive, and their main thesis about a PIE homeland somewhere in present-day Armenia or Eastern Anatolia was not supported by archaeological data (Mallory 1989; Anthony 2007). Likewise, Renfrew's (1987) theory, which assumes that PIE spread with the expansion of agriculture from Anatolia in the neolithic period, found little acceptance among linguists. It is almost generally acknowledged that the reconstructed lexicon of PIE reflects the eneolithic, rather than neolithic culture (see the papers in Markey & Greppin 1990; and the recent assessment of the palaeolinguistic evidence in Anthony 2007). Moreover, Renfrew's Anatolian homeland of Indo-Europeans is difficult to square with the linguistic and toponymic evidence for non-IE languages (Hattic, Hurrian) which were spoken there in the Bronze Age.

Some typological parallels between PIE and the NW Caucasian languages were proposed by Colarusso (1981), who believes, moreover, in their deep genetic relationship.[2] Even more recently, Kortlandt (1989, 2009) reviewed the original arguments in favour of PIE–Caucasian language contacts quite favourably. He claimed that some of the features that are, in his opinion, reconstructable to PIE (e.g. the minimal vowel system and a very rich consonantism, the presence of gender, and the ergative clause alignment), should be attributed to adstratum or substratum influence of Caucasian languages. However, Kortlandt exposed this thesis without presenting any new evidence, and without attempting a thorough critical re-examination of the problem. In this paper, we attempt to do precisely that: we re-examine the existing evidence for early contacts between PIE and the Caucasian languages in the light of the growing consensus about the PIE homeland to the north and northwest of the Caucasus mountains. In Section 2, we look at the archaeological data relevant to the issue. Section 3 discusses the lexical evidence for language contact between PIE and the indigenous language families of the Caucasus, and Section 4 examines the structural-typological parallels between them. Section 5 presents some areal features that PIE does not share with the languages of the Caucasus, and the conclusions are presented in Section 6.

2. On the historical context of Caucasian–Indo-European contacts

The problem of the PIE homeland is as old as it is unresolved. However, there is little doubt that PIE was spoken during eneolithic, or early Bronze Age, by a population familiar with metals, agriculture and animal husbandry, wheeled transport and other elements of what Sherratt (1983) has called 'secondary products revolution'. The theory that best fits the linguistic evidence about the PIE community has a long history – it was known as the 'Kurgan hypothesis', first proposed by Gimbutas (1970), but subsequently supported by strong evidence by Mallory (1989) and Anthony (2007).[3]

According to this line of thinking, speakers of Early PIE probably included, but were not necessarily limited to, the people associated with eneolithic Sredni Stog culture in Ukraine (late fifth and early fourth millennium BC, e.g. the famous Dereivka site),[4] while the late PIE community was represented by the so-called Yamnaya or 'Pit-grave' horizon, stretching from Moldova to Kazakhstan from the late fourth millennium BC. Some would prefer to regard the Yamnaya horizon as covering only a part of the PIE area, excluding at least the dialects that would later evolve into Tocharian and Anatolian languages, but there is little doubt

[2] His proposed 'Proto-Pontic' family has not, however, been well received in the scholarly literature.

[3] For a review and critique of Anthony's book, see Kaiser (2010).

[4] Some archaeologists doubt the unity of the Sredni Stog culture (e.g. Rassamakin 1999, 2004) and also deny that there is any evidence for large-scale migrations from the east to the west in the early eneolithic period in the Pontic area. This is not crucial for our purposes, since the eastward expansion of Indo-Europeans – say, during the late Cucuteni-Tripolye culture (3800–3600 BC) – would have brought them into contact with the Maikop, or Maikop-influenced cultures (on which see below). According to Korenovskij (2004), Maikop influences reached the Don in the period between 3500–3000 BC.

among both linguists and archaeologists that the Yamnaya horizon *was* Indo-European. It remains unclear whether its sources lay to the east, in the earlier Khvalynsk culture on the middle Volga (Carpelan & Parpola 2001), in the Sredni Stog culture, which is sometimes considered to be the first to domesticate horses in the steppes (Anthony 2007),[5] or further west, in the Cucuteni-Tripolye culture of Romania, Moldova, and Western Ukraine (Middle fifth – Middle fourth millennium BC). Note, however, that the rise of Yamnaya culture, the evidence for the deep exploitation of the steppes by means of wheeled transport (and, probably, horseback riding), as well as the rise of new metallurgical centres on the northern slopes of the Caucasus, are simultaneous with the demise of the Cucuteni-Tripolye complex around 3500 BC, which is a plausible date for the period when PIE was spoken. According to Kohl (2007), there is evidence that sedentary populations of western Ukraine (the area of Cucuteni-Tripolye expansion) adopted a more mobile lifestyle at that time, and that the economy shifted from agriculture to herding. It is well known that such periods of crisis and change in subsistence patterns are well suited for massive language shifts. For example, the population of Illyricum shifted from Vulgar Latin to Slavic in the sixth and seventh centuries, during the collapse of the Roman economic system, and a similar process occurred in the eastern parts of Britain, where British Celtic was replaced by Anglo-Saxon, roughly a century earlier.

As can be seen from Figure 1, the regions where PIE was spoken were near both the Urals – which was the old and persistent border between neolithic and hunter-gatherer societies during the fourth and third millennia BC – and the Caucasus. There is little doubt that PIE influenced Proto-Uralic, and loanwords from PIE into that proto-language are counted by the dozens (see e.g. Haarmann 1994; Carpelan & Parpola 2001; Kallio 2004). Likewise, it is very probable, if not strictly proved, that both PIE and Proto-Uralic were descended from a common proto-language usually called Indo-Uralic (see, e.g. Kortlandt 2001). On the other hand, possible relationships between speakers of PIE and their southeastern neighbours on the Caucasus are by far less clear.

The Caucasus has been inhabited by Neolithic populations since the sixth millennium BC. However, the earliest neolithic farmers of the Kura and Araxes river valleys (ca. 5800–5600 BC according to Anthony 2007: 186), were isolated from the steppes by the impassable mountain peaks, and no direct contacts were possible at that time. Eneolithic spread to the area to the north of the Caucasus range somewhat later, in the early fifth millennium; inhabitants of the area around Nalchik (the capital of the present-day Kabardino-Balkar Republic in Russia) had at least some domesticated sheep, and the use of copper tools is also documented, but metals were still uncommon. There is some evidence that the material culture of North Caucasus was influenced by the steppe cultures at that time, and that the domesticated animals were imported there from the Dnieper region, rather than from Mesopotamia (Anthony 2007: 187–8).

The most important archaeological culture of the early bronze age in the North Caucasus is the so-called Maikop culture (Korenovskij 2004, 2010). Its beginnings are slightly earlier than the Yamnaya horizon, and it is roughly contemporaneous with the collapse of the 'Carpatho-Balkan' metallurgical complex of the early eneolithic (Kohl 2007). It is also contemporaneous with the expansion of the trade in copper and other cultural items from Mesopotamia, and the beginnings of the use of arsenic bronze in the Caucasus. It is generally agreed that Maikop played the pivotal role in the transmission of the cultural artefacts from Mesopotamia to the

[5] Note that the evidence for the horse cult at Dereivka (a famous Sredni Stog site) is very scant. Moreover, there is no evidence for the use of horse-bits in Sredni Stog sites which Anthony himself admits. Therefore, most of the evidence for the use of horses in this early period in the steppes is circumstantial (the interpretation of certain objects as cheekpieces and the interpretation of the bone remains as indicative of horse breeding rather than horse hunting).

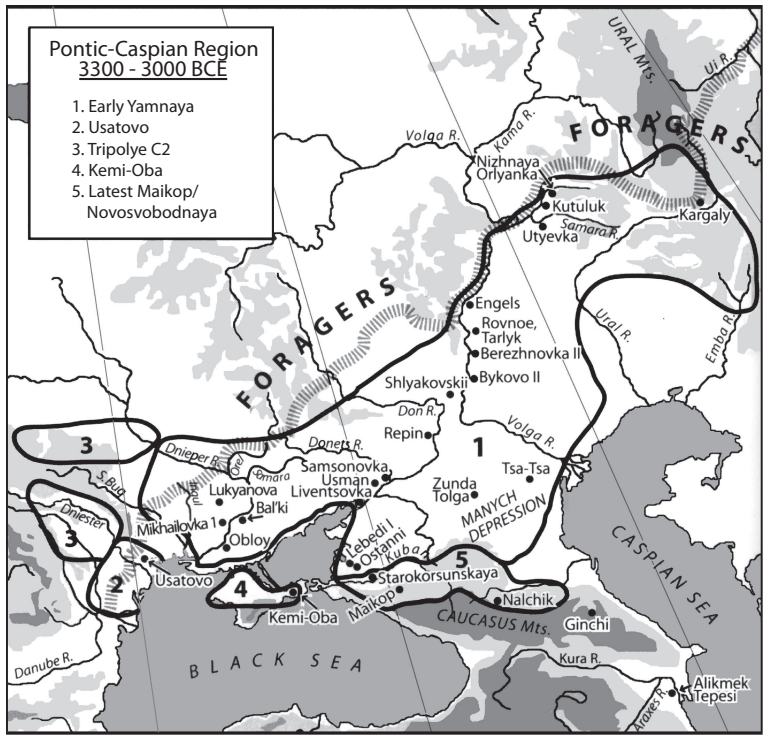

Figure 1. Culture areas of the early Bronze Age (from Anthony 2007: 302, reprinted with permission)

steppes in the second half of the fourth millennium BC. The following quotation by Andrew Sherratt testifies to this claim:

> Maikop is famous, of course, as the first 'barbarian' burial of Eurasia, with fine silver drinking-vessels and other metalwork in a provincial variant of the Uruk style of the period; and its position between the Transcaucasian area, linked by contacts along the Euphrates to Arslantepe and the Uruk colonies, and the Pontic steppe area where the Pit-grave culture crystallised around innovations such as wheeled vehicles transmitted to Mesopotamia, give it a pivotal role in mediating links between the two. (Sherratt 1998: 271)

The Maikop culture is known chiefly through its burials. The well-known kurgan grave from Maikop was full of precious objects made of gold and silver, in large parts imports from Mesopotamia, including figures of lions and other non-native wild animals. Archeological discoveries from Maikop and other contemporary sites testify that bearers of that culture were familiar with woollen fabric and wheeled transport, as well as with rather developed forms of metallurgy. Many archaeologists connect the development of Maikop culture with the expansion of trade with Mesopotamia, and some would go so far as to identify its bearers with migrants from the south of the Caucasus (Korenovskij 2004). However, there is a sharp contrast between the house structures at Maikop sites and those in the Transcaucasia, for example in the contemporary Kura-Araxes culture. Also, the economy at Maikop seems to have been based chiefly on animal husbandry, while south of the Caucasus agriculture played

a much more prominent role. A trade network connected Maikop with sites to the north and northwest, and many pots and artifacts of recognisably Maikop provenance were found in the Steppes. The Konstantinovka 'Kurgan-type' culture along the lower Don and in the Azov Steppes owes a lot to the Maikop culture, and metal artefacts from Maikop are found at the sites attributed to Kemi Oba culture (3700–3200 BC), which was connected with the steppe cultures and located chiefly on the Crimea (Mallory & Adams 1997: 327). Some archaeologists argue that Lower Mikhailovka (3600–3000 BC), which is traditionally viewed as a typical 'Kurgan-type' culture, also has its origins in the Caucasus (Manzura 2005: 327). Some of the earliest evidence for wheeled transport outside Mesopotamia is found at Maikop, and it appears very likely that carts and wagons were introduced into the Steppes via the Maikop culture (Kohl 2007).

If many cultural items were indeed transmitted from the Middle East to the Steppes via the Caucasus, we would expect at least some linguistic traces of these early cultural and trade contacts. It is almost universally assumed that peoples of the North Caucasus are indigenous to the region. Both NE and NW Caucasian languages[6] share rich terminology connected with neolithic economy, for example, the words for cereals such as 'wheat' (Kabardian *g*^{w}*adz*, Tsez *qiči*, Dargwa Chiragh *q:Iuč:* 'bread', according to NCED (*North Caucasian Etymological Dictionary*; Nikolayev & Starostin 1994) 895)[7] from PNC *q:HwōžĀ), and 'millet' (Kabardian *x*^{w}*ə*, Abkhaz *a-šə́*, Lak *š*^{w}*a* 'mown crops', Lezgian *fu* 'bread', according to NCED 763 from PNC *ʃwɨ ʔwV). We also find no less than six terms for metals in NCED, and almost all of them are attested in both NW and NE Caucasian languages, for example, *rĕwc:wi 'red copper, gold' (Chechen *deši* 'gold', Lak *duwsi:* 'red copper', Tabassaran *jišur* 'gold', Kabardian *dəśa* 'gold', etc., NCED 948), *t'VtV(wV) 'silver, gold' (Chechen *deti* 'silver', Ubykh *t*^{w}*at*^{w}*a* 'gold', NCED 1005), etc. The North Caucasian languages, whether they are genetically connected or not, may have spread in their present-day homeland during the early Neolithic from the area to the south of the Caucasus range, in which case it would be easier to understand their probable genetic affiliations: the NW Caucasian languages are possibly related to Hattic, the extinct language of Central Anatolia, while the NE Caucasian languages have been connected to Hurro-Urartian, the extinct language family spoken in present day Armenia and Northern Mesopotamia (Starostin & D'jakonov 2007).[8] Alternatively, both of these language families, or only the NW Caucasian languages, may be associated with the expansion of the Cucuteni-Tripolye culture westwards in the fifth and early fourth millennium BC (Manzura 2005), which would help us better understand the undeniable Balkan influences in the Maikop culture.

After ca. 3000 BC the trade with Mesopotamia via the Caucasus was diminished, and Mesopotamian trade routes shifted southward and eastward, towards India through Iran and the Elam civilisation. Consequently, sites like Maikop lost their importance, and there is little reason to believe that areas on the northern slopes of the Caucasus were in any sense culturally influential in the third millennium BC.

[6] The genetic relationship between NE Caucasian (or Nakh-Daghestanian) and NW Caucasian (or Abkhaz-Adyghean) appears probable, but it is still not strictly proved. The hypothesis that these two families are related is accepted, for example by Chirikba (1996) and Starostin (2007a, 2007b), but doubted by several western linguists (e.g. Nichols 2003).

[7] The examples quoted from NCED in this paper follow the transcription used in that dictionary. Although the reconstructions in NCED are quite speculative, the primary data are generally very reliable.

[8] D'jakonov and Starostin originally proposed that Hurro-Urartian was just another branch of Northeast Caucasian, but today many linguists would prefer to consider it as a branch of North Caucasian, or, rather, as related to North Caucasian but not one of the North Caucasian branches themselves. For a critique of the hypothesis that Hattic is specifically NW Caucasian see Kassian (2009) (the author believes that Hattic and NW Caucasian are related at a deeper level). All of these hypotheses are speculative and need more research.

There is a potential chronological problem for the claim that Maikop culture influenced the PIE speech community. Many linguists now believe that Proto-Anatolian separated from the rest of PIE dialects before 3500 BC, that is, before the Maikop influence on PIE could have been exerted (according to the archaeological data presented above). If that is correct, we would not expect any Caucasian loanwords or contact-induced structural traits in Anatolian languages, or in Proto-Indo-Hittite (the putative ancestor of Proto-Anatolian and the other IE languages). On the other hand, there is no firm evidence that Proto-Anatolian was geographically separated from the rest of Indo-European in the late fourth millennium. Especially if the Sredni Stog culture proves to be a mirage, and the 'Kurgan intrusions' into the Balkans during the period from 4200–3800 BC never occurred (Rassamakin 1999, 2004), there is no objection to the claim that the separation of Proto-Anatolian from PIE occurred later, in the middle of the third millennium BC.

3. Possible loanwords from Caucasian languages in PIE

If the early contacts between Caucasian languages and PIE occurred at a time when the important trade routes linked the steppes and Mesopotamia via the Caucasus, we would expect at least some loanwords from Caucasian languages in PIE. The lexical correspondences between Kartvelian and PIE have been studied extensively by Klimov (1981, 1986: 195–7, 2010), Gamkrelizde and Ivanov (1984), Dolgopolsky (1987), Gippert (1995), and others. There appears to be a general consensus that Kartvelian was on the receiving side of loanwords, while there do not seem to be any demonstrable Kartvelian loanwords in PIE. PIE loanwords in Kartvelian may have spread through trade networks in the eneolithic period; interestingly, they are said to include two numerals (Proto-Kartvelian *otxo- 'four' (Georgian *otx-*) < PIE *ok'toh$_1$, cf. Lat. *octō,* Gr. *oktṓ*, etc.,[9] and Proto-Kartvelian *eks$_1$w- 'six' (Georgian *ekws-*) < PIE *s(w)ek's (Lat. *sex*, Gr. *héks*, etc.).[10] However, the exact number of those loanwords is disputed,[11] especially since, for linguists adhering to 'Nostratic theory', Kartvelian and PIE also share a part of their vocabulary inherited from PIE. Many of the aforementioned etymologies could turn out to be based on no more than chance similarities. Moreover, it is highly likely that several words sometimes considered as Indo-Europeanisms were actually borrowed from Proto-Armenian (Gippert 1995). This may be the case with the word for 'wine' (Georgian *γvino*), where Georgian *γv-* clearly points to a Proto-Armenian origin (Dolgopolsky 1987, Martirosyan 2010: 214–215), cf. Arm. *gini*, with the regular change of PIE *w to Arm. *g* (via *γ^{w}).[12] Another early loanword which might have been borrowed through Armenian is the verb meaning 'to weave', Proto-Kartvelian *γweb- (Georgian *γob-*) < PIE *web^{h}- 'id.' (Gr. *hyphaínō,* OHG *weban*, TochB *wāp-*, etc.). Note, however, that its reflex is unattested in Armenian. The word for 'daughter-in-law' or 'bride' (Megrelian *nosa,*

[9] This is taken to be the dual of an earlier, but unattested, numeral meaning 'four', which is, of course, rather suspicious. Moreover, the canonical PIE roots began with consonants, so the reconstruction *h$_3$ek'toh$_1$ is preferred by many Indo-Europeanists.

[10] In the case of the numeral 'six', the uncertainties connected with the reconstruction of the initial consonants (Szemerényi 1989: 236) might imply that this word is of foreign origin in PIE: W *chwech* and Av. *xšvaš* point to *sw-, Lat. *sex* and Goth. *saihs* point to *s-, while OPr. *uschts* seems to point to initial *w-. And, of course, the similarity with the Kartvelian word could be accidental.

[11] Klimov (2010) finds 14 PIE loanwords in Common Kartvelian, and 40 more in Georgian-Mingrelian. Since Proto-Kartvelian is generally believed to be a younger proto-language than PIE (its time-depth need not be more than 4000 years), it is very likely that it borrowed words from already divergent IE dialects, especially Proto-Armenian and Indo-Iranian.

[12] The fact that winemaking is native to Georgia is not crucial here. Irish *beoir* is an English loanword, although there was an OIr. word for 'beer' (OIr. *cuirm*). Besides, there are reasons to believe that the word for wine, *woyh$_1$no-, is original in PIE, since it can be derived from the root *weyh$_1$- 'to weave, to twist', cf. Lat. *vieo* 'to plait, weave', Lith. *výti* 'to twist, wind'.

nisa) is usually derived from PIE *snuso- (Gr. *nyós,* Arm. *nu, nuoy*, OCS *snъxa*, etc., IEW (*Indogermanisches etymologisches Wörterbuch*; Pokorny 1959) 978). This word was also borrowed into many NE Caucasian languages, cf. Chechen *nus*, Avar *nus*, Akhvakh *nusa*, etc., and we also find its reflexes in NW Caucasian (Kabardian *nəsa*). It may also belong to the layer of Proto-Armenian loanwords (Martirosyan 2010: 506–7), but Ossetic *nos-tae* is also a possible source.[13]

The possible loanwords from North Caucasian languages are a much less studied area. So far, there have been two attempts to identify possible North Caucasian loanwords in PIE; the first is Starostin (2007a), which identifies as many as 82 possible lexical parallels between PIE and North Caucasian.[14] Starostin relies heavily on his and Nikolaev's (Nikolayev & Starostin 1994) reconstruction of Proto-North Caucasian, which is rather speculative, and is not accepted by all specialists (Schulze 1997; Nichols 2003). On the basis of sound correspondences between PIE words and their alleged Caucasian cognates, Starostin concludes that the direction of the borrowing was exclusively from Proto-North Caucasian (or one of its dialects) into PIE, and not vice versa.[15] The other attempt, by Colarusso (1997, 2003), is based on his own peculiar reconstruction of Proto-NW Caucasian, as well as on his assumption that PIE and Proto-NW Caucasian are descended from a putative Proto-Pontic language, which has not gained much acceptance. In what follows we concentrate on Starostin's (1997a) list, which has to be constrained and re-evaluated with clear methodological principles. It is not our task here to determine all of the loanwords from North Caucasian into PIE, but rather to establish whether there are any. In doing that, we must note the following problems with Starostin's etymologies:

1. Loanwords usually belong to semantic fields that are especially prone to borrowing (e.g. technological and cultural terminology, names of plants and animals); thus, Starostin's (2007a) derivation of PIE *$b^heh_2g'^hu$- 'arm' (Skr. *bāhú-,* OIc. *bōgr*, Gr. *pēkhys,* IEW 108) from PNC *p:üggV 'hip' (Kryz *beg*, Khinalug *buγru,* Kabardian *bγə*, etc., not mentioned in NCED) is *a priori* improbable in light of the general resistance of body part terms to borrowing. Although it is not impossible, one would have to have very strong additional reasons to accept such an etymology, but there are none. Note also that the reconstructed meanings in PIE and PNC differ.
2. The loanwords in PIE should be identifiable by their unusual phonological characteristics (i.e. unusual root structure, the presence of the rare vowel *a, etc.); they should be isolated in the lexicon and not easily derivable from verbal roots, like the majority of inherited PIE words. Several of Starostin's alleged NC loanwords are patently derived from PIE roots. For example, PIE *$kleh_2u$- 'key' (Lat. *clāvis*, OCS *ključь*, etc.) is easily derived from the PIE root *$kleh_2u$- 'to close, obstruct' (Lith. *kliūti* 'to obstruct', IEW 605), so there is no need to assume that it has anything to do with PNC *k'uɬe 'lock, bolt, key' (Lak *k'ula*, Avar *k'ul*, etc., NCED 726f.). Similarly, the PIE word for 'wheel' reconstructed on the basis of Hitt. *hurki-* 'wheel', ToA *wärkänt-* 'id.' has been related to PNC *hwəlkwē- 'wheel' (Lak *hark^w*-, Lezgian *akur*, Avar *hokó*, Kabardian *g^{w}ə* 'carriage', Šagirov 1977, I: 113, NCED 498). However, the PIE word may be derived from the root *$h_2$werg'- 'to turn' (Lat. *vergo* 'slope down, sink', OE *wrencan* 'turn, wring', Skr. *vṛṇákti*

[13] NCED 856f. attributes this word to PNC (*nusA) and claims that the Kartvelian forms were borrowed, ultimately, from West Caucasian. See also Dolgopolsky (1987).

[14] Dolgopolsky (1987) discusses only some possible loanwords in NE Caucasian and concludes that all of them stem from Indo-Iranian rather than PIE.

[15] This is not altogether certain; Starostin does not discuss such possible loanwords as the word for 'mead' (Lezg. *med* 'syrup', Tindi *med* 'beer', Godoberi *medi* 'beverage'), which may be from PIE *med^hu- 'mead' (Skr. *mádhu*, Russ. *mëd*, Lith. *medùs*, OIr. *mid*, etc.), or the word for 'bride, girl' (PNC *ś^{w}asa > Chechen *sēsag*, Lezgian *s^{w}as*, Agul *susa-*, Kabardian *pśāśa* 'girl'), which might be connected with PIE *swesōr 'sister' (Skr. *svasā*, Lat. *soror*, OCS *sestra*, etc.).

'turn, ward off', perf. *vāvrjé,* LIV (*Lexikon der indogermanischen Verben*; Helmut Rix et al. 2002.) 259, IEW 1154), so there is no need to consider it a loanword.

3. The reflexes of the PIE etymon should not be limited to European languages only, because in that case a local substratum is much more probable than any Caucasian source. For example, Albanian *dardhë* 'pear' and Gr. *ákherdos* (attested only as a gloss in Hesychius, so somewhat doubtful) were more probably borrowed from some Balkan substratum language than from PNC *qHü:re 'pear' (Archi *χ̄er-t*, Dargwa *qaʔr*, Chechen *qor*, Kabardian *q*^{w}*əź*, etc., NCED 893). Likewise, PIE *(H)rug^{h}yo- 'rye' (OE *ryge*, Lith. *rugỹs*, Proto-Slavic *rъžь) could, in principle, be from a North Caucasian source (cf. Avar *rok:* 'wheat', Lezgian *gerg* (with reduplication) 'id.'), but reflexes in Indo-European are limited to Germanic and Balto-Slavic, so it seems more probable that the source is ultimately a North European substratum.
4. Finally, we should avoid root etymologies. This means that it is not enough to establish the correspondence of roots between PIE and North Caucasian, questions of word formation must be addressed as well. In almost all of Starostin's etymologies, including the ones discussed above, we do not find any explanation for the formation of the reconstructed PIE words, which otherwise look just like inherited lexical items. For example, Starostin (2007a) derives PIE *g^{w}reH(w)ōn 'quern' (Goth. *qaírnus*, Lith. *gìrnos*, Arm. *erkan,* OIr. *bró,* etc.) from PNC *χ:ʔ*werV*- 'mill, quern' (Chechen *ħajra*, Lak *hara*, Tabassaran *raRʔ*, Lezgian *reR*^{w}, etc., not mentioned in NCED). It is true that this PIE etymon is not easily connectible with any verbal root, so it may be a loanword; on the other hand, it looks like a regular n-stem with Ablaut (Nom. sg. *g^{w}reH(w)ōn/Gen. sg. *g^{w}rHn-os), which we expect in inherited vocabulary.

However, among Starostin's 82 PIE roots, we do find a few that *could* represent early loanwords from North Caucasian, or some sub-branch of that phylum. Let us adduce just the best candidates:

1. PIE *ayg'- 'goat' (Gr. *aíks, aigós,* Arm. *ayc*, IEW 13) and *ag'- 'goat' (Skr. *ajá-*, Lith. *ožỹs*, Alb. *edh*, OCS *azьno* 'goat-skin', IEW 6).[16] If the two etymons for 'goat' in PIE do belong together, the irregular alternation between *a and *ay might indicate that we are dealing with a loanword. NCED reconstructs more than a dozen NC words for 'goat' (including 'kid', 'mountain goat', etc.), and one of them might be the source of the PIE words, cf. Adygh *āča*, Kabardian *āža* (NWC *aǩ:^{w}a), Dargwa Chiragh ʕač:*a* and perhaps Chechen *awst* (all derived by NCED 245 from Proto-NC *ʔ̄ējʒwē). Note also PNC *ʒ:ikV- 'goat, kid' (Tsez *ceḳi,* Lak *ʂuku,* Akhvakh *ǯeḳer*, etc. NCED 1094, not mentioned by Starostin 2007a), which may have been the source of PIE *dig^{h}- 'goat' OHG *ziga*, Gr. (< ?Thracian) *díza*, perhaps also Alb. *dhi*.[17]
2. PIE *wlHneh$_2$ 'wool' (Lat. *lāna*, W *gwlan*, OCS *vlьna*, Lith. *vìlna*, Skr. *ū́rṇā-*, Hitt. *hulana-,* etc., IEW 1139); Starostin and Nikolaev reconstruct seven words for 'wool' in PNC, and one of them, reconstructed as PNC *ƛ:wāħni (NCED 769), might be the source of the PIE word. Its reflexes are preserved only in NE Caucasian (cf., e.g., Chechen *kan*, Bezhta *ƙũ*, Aghul *xej*, Archi *oƙ*, etc., Khinalug *ka*). The PIE root for 'wool' is attested only in the latest stratum of the vocabulary (it is an ā-stem, a-stem in Anatolian). It does not show Ablaut and it is not obviously connected to any verbal root, although some connect it to Lat. *vello, vellere* 'to pull out' < 'to pluck' (de Vaan 2008: 325).[18] The correspondences in NE Caucasian are regular according to NCED

[16] These words would be reconstructed as *h$_2$eyg' and *h$_2$eg', respectively, if one does not believe in PIE *a, but there is no independent evidence for a word-initial laryngeal in either case. If *ag'- and *ayg'- are indeed related, the alternation between *a and *ay could be due to alloglotic adaptation of a foreign phoneme.

[17] Note that this Albanian word has also been derived from *ayg'- above.

[18] Lat. *vello* is more plausibly derived from *welh$_3$- 'to strike' (Hitt. *walhzi*).

(though not obvious) so this word could be a serious candidate for a Caucasian loanword. However, note that Nichols (2003: 262) offers a different reconstruction, positing Proto-NE Caucasian *(D=)Vła(-j) (where D- is the gender prefix, and = marks the morpheme boundary), which makes the loanword hypothesis somewhat less plausible.

3. PIE *b^{h}ares- 'a kind of grain, flour' (Lat. *far*, OIr. *bairgen* 'bread', OIc. *barr* 'grain', Croat. *brašno* 'flour', IEW 111); Starostin (2007a) reconstructs PNC *bVrč'- 'a cereal, grain' (Chechen *bažan* 'rye', Avar *purč'ina* 'barley', Godoberi *beč'in* 'barley', not mentioned in NCED). Note that the IE reflexes of this word are limited to Europe, but the a-vocalism points to the conclusion that this etymon was borrowed. Caucasian languages are a possible source.

An interesting case is the PIE word for 'horse', *h$_1$ek'wo- (Lat. *equus,* Skr. *áśva-,* etc., IEW 301). This word looks suspiciously like the words for 'horse' in various North Caucasian languages, cf. Kabardian *šə*, Abkhaz *a-čə́* (Proto-NWC *č^{w}ə), Avar *ču*, Karata *ič^{w}a* 'mare', Lezgian *šiw* 'horse', etc., NCED 520). In PIE, the root of *h$_1$ek'wo- may be connected with *h$_1$oh$_1$k'u- 'swift' (lat. *ōcior* 'faster', Gr. *ōkýs*, Skr. *āśú-*).[19] In this case the direction of borrowing is probably from PIE into North Caucasian, and I do not see why PIE could not have been the source language: steppe horses were probably traded for Mesopotamian imports on the North Caucasus in the eneolithic period. The fact that we find fricatives and affricates in the Caucasian reflexes of this word indicates that the source could have been an IE dialect of the satem type.[20]

To conclude this section, evidence for borrowing between PIE and the languages of the Caucasus is scant and inconclusive. In light of tremendous difficulties in reconstructing Proto-NE Caucasian and Proto-NW Caucasian (not to mention putative Proto-North Caucasian), this is perhaps unsurprising. The lexical parallels with Kartvelian are easier to demonstrate, but they mostly involve cultural items that may have been borrowed indirectly, via the trade route that connected the steppes with Mesopotamia in the fourth millennium BC. It is quite possible that there are loanwords that cannot be recognised as such, and we should not forget that intensive language contacts are possible even without massive lexical borrowing. Such patterns of language contact have been documented among the Native American languages of the Pacific Northwest (Thomason 2001: 11, 63), in the Amazon (Aikhenvald 2001) and among the Turkic, Uralic and Yeniseian languages of South and Central Siberia (Skribnik 2011). In cases where the structure of the language makes it difficult for it to borrow new lexical items (e.g. if compounding is the default strategy for deriving new meanings), languages can co-exist side by side for centuries, and this will not be visible in their vocabularies. However, mutual influences can exert themselves in grammatical structure, especially if there is a widespread pattern of bilingualism, and if exogamy is the norm between two ethnically and linguistically different communities. Such a sociolinguistic situation could have been rather long-lasting (say, if speakers of PIE had extensive trade relations with the bearers of Maikop culture during the eneolithic), but one can also imagine a situation of relatively rapid language shift (say, if parts of the non-IE speaking Maikop population shifted to PIE at a time of social and/or economic crisis). The sociolinguistic typology of language change (Guy 1990; Ross 1991; Yakubovich 2009: 410–3), simply does not allow us to decide which is the most likely scenario. Finally, languages can be parts of the same *Sprachbund* (language area) and share a

[19] If this etymology is correct, the adjective would have to be derived from the root by reduplication, which is somewhat unusual in PIE word formation. Many linguists reject it as popular etymology (e.g. Kulanda 2008, who argues that the PIE word is indeed borrowed from North Caucasian, since there are no known Nostratic cognates). For another example of a reduplicated adjective in PIE cf. Hitt. *āššu-* 'good' < PIE *h$_1$o-h$_1$su- (Gr. *eu-* 'good' < *h$_1$su-).

[20] Dolgopolsky (1987) considers only the Daghestanian forms and claims they were borowed from Indo-Iranian.

number of structural features if they are not spoken in areas adjacent to each other, that is, if other languages belonging to the same *Sprachbund* intervene. We see this, for example, in the Balkans, where we do not find many loanwords from Romanian in Albanian (or vice versa), because a belt of South Slavic languages separates them.[21] On the other hand, both languages are exemplary members of the Balkans *Sprachbund*, sharing such features as the vowel 'shwa' (ə), the lack of infinitive, postposed definite articles, clitic object pronouns. It is perfectly possible that PIE and the North Caucasian languages were not spoken in immediate vicinity, but that one or several now extinct languages, belonging to the same *Sprachbund*, were spoken between them. In that case, we would expect the linguistic contacts between them to be visible only in shared structural features, to which we now turn.

4. Areal-typological features shared by PIE and Caucasian languages

We will consider only those linguistic features of PIE that show areally non-trivial distribution, that is, features that are not omnipresent in languages of Eurasia. For example, it would be useless to try to conclude anything from the fact that both PIE and Proto-Northeast Caucasian had case inflection, since languages with case inflection are abundant in the whole of Eurasia, and are reasonably common in other parts of the world. Therefore, we shall focus on those features that PIE shares with languages of the Caucasus, but not with the languages of North and Northeastern Eurasia, especially with those that are historically attested closest to the putative PIE homeland, namely, Uralic and Altaic.

Moreover, it would be methodologically improper to compare typological features of PIE, which was spoken several millennia ago, with the areal distribution of features found in contemporary languages, unless there are additional reasons to believe that these features are likewise inherited from their respective proto-languages (Nichols 2007). A good illustration of the case in point is provided by the distribution of the rhotic consonant *r.

PIE had only one rhotic consonant (*r), apparently showing restricted distribution: it could not occur word-initially. All of the examples with word-initial *r-reconstructed in the classical Indo-Europeanist handbooks (e.g. IEW) may have contained a word-initial laryngeal, reflected as the 'prothetic vowel' in Greek, for example, in PIE *$h_1$rud^{h}ros 'red' (Gr. *erythrós*, Lat. *ruber*, OCS *rъdrъ*, etc., IEW 872).[22] We cannot be sure that all prothetic vowels in Greek are indeed from PIE laryngeals, especially in light of the fact that initial *e*- before *r* (from PIE *h_1) is more common than *a*- (from *h_2), and *o*- (from *h_3), but it is fair to say that the reconstruction of initial *r- in PIE is very uncertain, while there were no restrictions against the occurrence of other consonants in word-initial position. The restriction against a word-initial rhotic is typologically well attested in the world's languages, but the areal distribution of languages with this restriction is not random: in Eurasia, they tend to cluster in and around the Caucasus.

The restriction against word-initial rhotic ([r]) is quite common in the languages of the Caucasus (Matasović 1992). This restriction is found, for example, in Aghul, Kabardian, Lezgian, Chechen, and Khinalug. Modern Kartvelian languages have words beginning with [r], but it seems that this resonant could not occur word-initially in Proto-Kartvelian[23]

[21] Of course, there is a layer of extremely old and numerous Latin loanwords in Albanian, and Romanian also contains a number of non-Slavic and non-Romance substratum words, some of which are also attested in Albanian (Solta 1980).

[22] Greek words beginning with *rh*- always come from PIE *sr- or *wr-, e.g. *rheûma* 'stream' < *srewmn. Armenian also has prothetic vowels before word-initial *r-, but these are not always derived from PIE laryngeals, for example, in Arm. *erkar* 'quern' (rather than the expected *rkar) < PIE *g^{w}reh$_2$wōn (OCS *žьrny*, OIr. *bró*, etc.).

[23] Words beginning with *r*- in Modern Georgian are the result of rhotacism, namely, of the change #s- > #r-. However, Klimov in his etymological dictionary of Kartvelian posits a number of Proto-Karvelian etyma with initial *r*- mostly in consonant clusters.

(Melikišvili 1980; see Matasović 1992, with references). On the other hand, the restriction against word-initial [r] is certainly not universal in the Caucasus, and it can be shown that ancestors of some modern languages had it. NCED reconstructs word-initial *r- in quite a number of etyma, e.g. PNC *rakwǐ- 'axe' (NCED 944f.), PNC *rīhV- 'time, day' (NCED 952), *rV:xwV- 'cattle' (NCED 956). Moreover, although Kabardian and Adyghe do not have word-initial *r-*, it occurs in that position in Abkhaz and Abaza, and it seems probable that the Abkhaz-Abaza situation represents an archaism. Likewise, the absence of word-initial *r- in Chechen and Ingush seems to be the result of the change of original *r to Proto-Nakh *d-, while in languages such as Lezgian, PNEC *r changed to *j- (cf. Lezgian *jaḳ*^{w}, Chechen *dig* 'axe'). Therefore, we cannot conclude anything but that the languages of the Caucasus offer some typological parallels to the distribution of the rhotic consonant, but these parallels are not probative of early contacts.

In what follows, we attempt to compare the features reconstructed for PIE with features in other proto-languages, rather than in their historically attested descendants. The problem is, of course, that PIE is much more reliably reconstructed than the other families of Eurasia, with the possible exceptions of Uralic and Kartvelian. For North West Caucasian, we have at least Chirikba's reconstruction of Proto-NW Caucasian (Chirikba 1996) and many helpful insights can be found in Abdokov's monograph (1981) and Šagirov's etymological dictionary (1977). For Proto-NE Caucasian, we have the important contributions to comparative phonology by Bokarev (1981) and Nichols (2003), as well as the rather speculative, but impressive etymological dictionary by Nikolayev and Starostin (1994). Comparative grammars of individual branches of NE Caucasian are the subject of Alekseev (1985, 1988) and Giginejšvili (1977).

4.1. Phonological features

4.1.1. Consonant-to-vowel ratio

It has been claimed (e.g. by Kortlandt 2009), that the minimal vowel system of early PIE is an areal feature shared by PIE and Caucasian. Minimal vowel system can indeed be posited for Proto-NW Caucasian (Starostin 2007a; Chirikba 1996), but it is very probable that labialisations and palatalisations of consonants in that language family reflect earlier vowel timbre distinctions. As for NE Caucasian, languages of that family never have minimal vowel systems; on the contrary, some of them, for example, Chechen, have some of the richest vowel inventories of all the languages of the world. And, although Starostin and Nikolayev's reconstruction of Proto-North Caucasian is far from being generally accepted, it does not posit a minimal vowel system for that protolanguage. Thus it is better to compare the PIE phonological system with those of the Caucasian languages with respect to another parameter, the consonant-to-vowel (c/v) ratio.

The consonant-to-vowel ratio is the number obtained by dividing the number of consonants in a language with the number of its vowels (this applies to the inventory of the phonemes, not to the occurrence of sounds in actual use). Consonant-to-vowel ratio is high[24] for PIE (12.5), if one accepts that the proto-language had only two vowels in the earliest reconstructable stage (*e and *o), as assumed by Kortlandt and the Leiden school (e.g. Beekes 1995). Such systems are rare cross-linguistically, but common in the Caucasus

[24] WALS (World Atlas of Language Structures; Haspelmath et al. 2006) classifies languages as having a 'high' c/v ratio if the number of consonants divided by the number of vowels is above 6.5. Only ten out of 563 languages in WALS have the ratio of twelve or higher, and all of the Eurasian languages of this kind are spoken on the Caucasus. We have assumed the standard reconstruction of the PIE consonant system with three series of stops (aspirated, voiced and voiceless), three types of gutturals (plain velar, palatalised and labiovelar), and three laryngeals. By this count, the total number of consonants in PIE was 25.

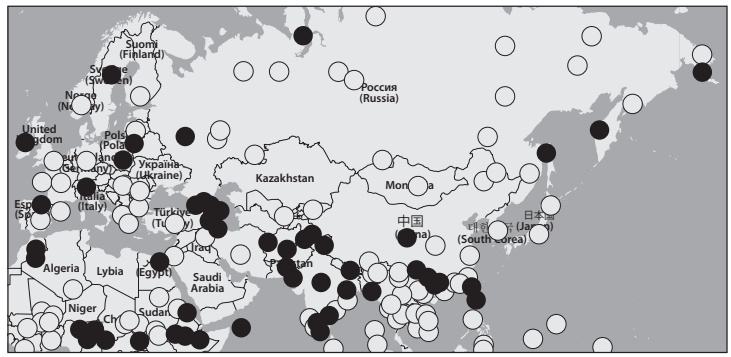

Figure 2. Languages with high and moderately high consonant-to-vowel ratio in Eurasia (black dots)[25]

(especially in NW Caucasian); both the reconstructed NW and NE Caucasian systems were certainly characterised by high c/v ratios. If four vowels are reconstructed for PIE (*e, *o, *ē and *ō, which is the system posited for Late PIE by the Leiden school, and closer to the vowel system posited for PIE by the majority of Indo-Europeanists), then it should be classified as a language with 'moderately high' c/v ratio (4.5–6.5). Even if short *a is accepted for PIE (if it existed, it was certainly rare), the c/v ratio for PIE would remain moderately high. Figure 2 shows the distribution of languages with high and moderately high c/v ratios in Eurasia; if modern IE languages in Europe (in which high c/v ratio may be inherited from PIE) are not taken into account, it is easy to see that languages with high and moderately high c/v ratios predominate in SW Eurasia, while languages with low c/v ratios are found chiefly in N and NE Eurasia, with very few exceptions (e.g. Nenets and Chukchi in Siberia both have moderately high c/v ratios). Languages of the Caucasus all have high or moderately high c/v ratios.

The presence of a high or moderately high c/v ratio can be used as an areally significant feature of PIE only if one accepts the Leiden school's reconstruction of the PIE vowel system, which the present writer does. If one sticks to the traditional reconstruction (presented, e.g. in Szemerényi 1989), PIE had a more complex vowel system, perhaps with up to five short and five long vowels (*a, *e, *i, *o, *u, *ā, *ē, *ī, *ō, *ū), and its c/v ratio was not similar to the c/v ratios found in the languages of the Caucasus.

4.1.2. Glottalised consonants

Clearly, the Caucasus is the only area in the whole of Eurasia, with the exception of Indochina, where glottalised consonants are at all common, hence the distribution of the feature is non-random (Figure 3). Moreover, the feature spread to non-indigenous languages of the Caucasus (e.g. to Ossetic and to Armenian dialects), which shows its areal nature. Glottalised consonants are safely posited for the proto-languages of all three Caucasian families, and they are not found in either Uralic or Altaic (with the possible exception of Korean). However, the attribution of glottalic consonants to PIE depends on whether one accepts the glottalic theory or not. The glottalic theory (see Gamkrelidze & Ivanov 1984), presupposes that PIE phonemes usually reconstructed as voiced stops (*b, *d, *g, etc.) should be reinterpreted as ejectives (*p', *t', *k', etc.). It is argued that this interpretation explains the

[25] Figures 2–12, with the single exception of Figure 10, are based on the WALS database available online at: http://www.wals.info.

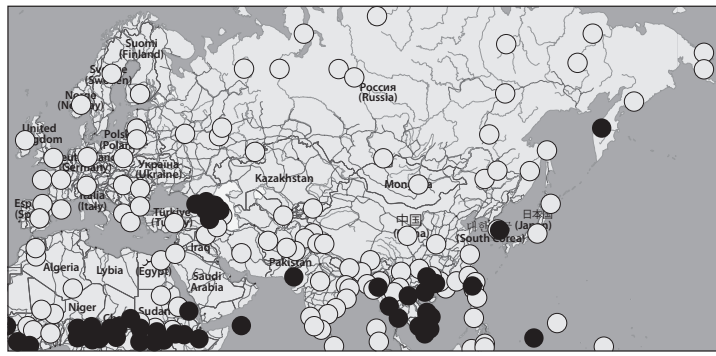

Figure 3. Languages with glottalised consonants in Eurasia (black dots)[26]

rareness, or non-existence of PIE *b (while *p and *b^h are well attested), and it is also in line with the typological improbability of the system with voiced, voiceless and voiced aspirated stops (but without voiceless aspirated stops), which is assumed by traditional IE linguistics. Arguments supporting the glottalic theory are disputed, and it would seem that today the theory is rejected by the majority of Indo-Europeanists (with a number of notable exceptions, like the Leiden school). However, no typologically plausible alternative interpretation of the PIE system of stops has been offered so far. Although it cannot be considered as proven, the hypothesis that PIE had ejective stops should not be dismissed out of hand, and if it is accepted, this has important consequences for the areal connections of the protolanguage. In our opinion, as far as the phonetic interpretation of the three PIE series of stops is concerned, no better alternative has been proposed (see Salmons 1994 for a thorough discussion of the evidence).

4.1.3. Tone distinctions

Proto-Indo-European was a language with tone accent; judging by Vedic, the tone system was minimal, with only two distinctive tones. However, there are serious reasons to assume that the Vedic system developed from a slightly more complex system, and that differences between various PIE accentual patterns in declension and conjugation resulted from earlier tonal differences (Dybo 2003; Pronk 2010).

On the Caucasus, tone systems are found in both NE and NW Caucasian families. Among NW Caucasian languages, tone accent is found in the Tapanta dialect of Abaza, and experts believe that this system might be inherited from Proto-NWC (Chirikba 1996: 94; Dybo 1989). In NE Caucasian, tone accent is found, for example, in Tindi, Chamalal, Ingush, Kryz, Budukh, Andi, Akhvakh, and the Inokhwari dialect of Khwarshi (Van den Berg 2005: 154, cf. also Hewitt 2004: 59). Most of the cases studied represent minimal tone systems, rather similar to the PIE system. It is at present unclear whether Proto-NE Caucasian and Proto-NW Caucasian were tone languages, but this is a possibility to be reckoned with. Kartvelian languages, on the other hand, have only dynamic stress systems. Just like Kartvelian, Semitic, Uralic and Altaic languages[27] also have only stress accentual systems, which makes the presence of tone accent in PIE and North Caucasian more significant from the areal point of

[26] Included are both languages with ejectives and those with implosives from the WALS database.

[27] Starostin (2007a: 597ff.) reconstructs tones for Proto-Altaic (on the basis of correlations between vowel length in core Altaic languages and tones in Korean and Japanese), but this hypothesis has not gained wide acceptance.

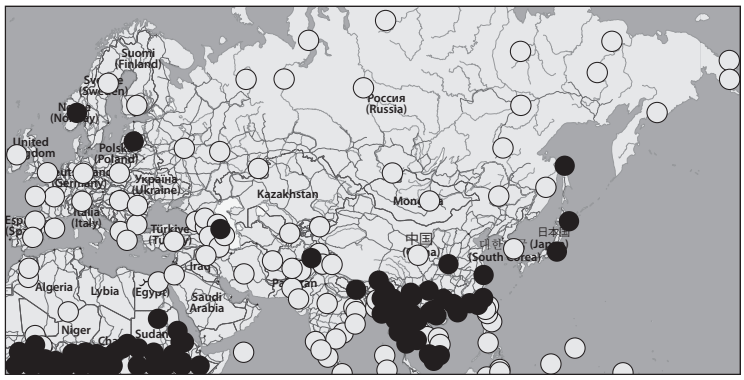

Figure 4. Languages with tone accent in Eurasia (black dots).[28]

view. Tone is a known areal feature, and SE Asia is, besides the Caucasus, the only other area where tone languages are found abundantly in Eurasia (Figure 4).

Since it appears that tone languages are quite widespread on the Caucasus and rare elsewhere in Eurasia (except in SE Asia), and tone oppositions can be reliably posited for PIE, this is a relevant feature for the areal typology of PIE.

4.1.4. Productive Ablaut

PIE used morphologically conditioned vowel alternations very abundantly, both in nominal and in verbal inflection. This type of morphology is rather uncharacteristic of languages of Northern Eurasia, but it characterises all indigenous families of the Caucasus, although in Caucasian languages, Ablaut is usually limited to the verbal system. Typological parallels between PIE and Kartvelian Ablaut patterns were noted long ago by Gamkrelidze (1966). For example, in Old Georgian, Ablaut can be used to distinguish transitive from intransitive verbal forms, for example, *v-cwet* 'I wear it out' vs. *v-cwt-eb-i* 'I wear out, am worn out' (Gamkrelidze & Mačavariani 1982: 197). Moreover, the other two Caucasian language families have productive Ablaut patterns in many ways comparable to PIE. For example, in Kabardian, we find the alternation between transitive and intransitive verbal stems expressed by Ablaut (Matasović 2011b: 11–12), e.g. *txan* 'read' (intransitive) vs. *txən* 'read' (transitive), and in Chechen, verbal tenses and aspects may be expressed by Ablaut for example, *mal* 'drink' vs. *mīl* 'drink repeatedly' (iterative), or *xottu* 'ask' (present) vs. *xätt-ira* 'asked' (past), Van den Berg (2005: 168–9). In many other verbs aspectual distinctions are expressed by consonant distinctions. In a few NE Caucasian languages, Ablaut is also used in nominal morphology, for example, in Lak, where it is used to distinguish absolutive and oblique stems of the noun, for example, *barc'* 'wolf' (absolutive) vs. *burc'-il* (genitive) (Klimov 1978: 110). Similar patterns are found in Archi and Bagvalal, and Ablaut is generally reconstructed for Proto-NW Caucasian and for Proto-NE Caucasian (Chirikba 1996; Nichols 2003). Although the extensive use of Ablaut in PIE and in all the indigenous families of the Caucasus can be viewed as an areal-typological parallel, it is certainly not exclusive, because productive Ablaut also exists in other language families of Eurasia, most notably in Afro-Asiatic. Like gender agreement (on which see below), it is rather one of the features which is generally present in the languages of SW Eurasia, and absent in those of NE Eurasia. Moreover, Ablaut, as such,

[28] For the languages of the Caucasus the WALS map is not quite correct.

is unlikely to be borrowed in situations of language contact. Rather, one would have to show that the presence of Ablaut in languages in contact is due to the spread of some other feature which caused the parallel development of Ablaut in neighbouring languages, for example, a dynamic accent causing regular syncope in certain gramatical categories, or a free tonal accent affecting the colouring of vowels. However, the origin of Ablaut in PIE is just as unclear as it is in the language families of the Caucasus.

4.2. Morphological and syntactic traits

4.2.1. Ergativity

The hypothesis that PIE was an ergative language was proposed early in the twentieth century (cf. Uhlenbeck 1901; Vaillant 1936),[29] but it still hasn't gained general acceptance. It is based on the fact that the accusative singular of masculine thematic nouns has the same ending *-m as the nominative and accusative singular of the 'thematic' class of neuters, which does not distinguish these two cases, cf. Lat. *iug-um* 'yoke' (NOM/ACC SG N) vs. *lup-um* 'wolf' (ACC SG M). Other neuters had zero-marked nominative and accusative singular, cf. Lat. *nomen-0* 'name' (NOM/ACC SG N). Since neuters are non-canonical agents (or transitive subjects), it has been assumed that, at some stage in PIE, they were not used as transitive subjects, and that their zero-marked 'nominative' case is actually the old absolutive. The original case of transitive subject might be preserved as the genitive/ablative ending in *-os (Beekes 1995: 194), identical with the nominative of the thematic stems, cf. Gr. *lýk-os* 'wolf' (NOM SG M) vs. *pod-ós* 'foot' (GEN SG). Note also that the stem of the nominative sg. of the PIE M and F demonstrative pronoun is *so- (Skr. *sá* M Gr. *ho* M etc.), while the oblique cases are built from the stem *to-, which is also the stem of the neuter demonstrative in all cases (Skr. *tád* N, Gr. *to* N, etc.). This fact has also been interpreted as a remnant of an earlier ergative pattern, in which the stem *so- was used only in the ergative.[30] From the areal-typological point of view, it must be admitted that languages of Eurasia with case systems and nominative-accusative clause alignment tend to have unmarked nominatives and marked accusatives (like Turkish, Finnish or Hungarian), rather than marked nominatives. The curious coincidence that the nominative singular ending has the same form as the genitive singular can be explained if one assumes that the genitive and the nominative endings in *-s of the animate noun classes originally expressed the ergative case (the pattern in which the ergative and the genitive (or possessive) case have the same morphology is well attested typologically, for example, in the Eskimo-Aleut languages and in the NE Caucasian language Lak (Van den Berg 2005: 162). The original absolutive of the animate class would have been preserved in the asigmatic nominatives of the stems in resonant, for example, *$ph_2tēr$ 'father' (Lat. *patēr,* Gr. *patḗr*, etc., IEW 829), although many linguists prefer to derive such nominatives from earlier sigmatic endings, with compensatory lengthening of the vowel before the resonant (*ph_2ters > *ph_2terr > *$ph_2tēr$). Moreover, the opposition between the nominative and the oblique stems in personal pronouns is expressed through supletion, so it is probably very old (e.g. PIE *eg'-h_2om 'I' > Skr. *ahám,* OCS *azъ* vs. *h_1me 'me' > Gr. *emé,* Lat. *mē*, Skr. *mām*). Of course, the proto-language could have had nominative-accusative pronominal inflection, and ergative noun inflection, a pattern which is found in many aboriginal languages of Australia. If only nouns were inflected according to the ergative pattern, it would be easier to understand why only some inanimates belong to the neuter gender in Late PIE (Matasović 2004). Those inanimates that are masculine or feminine (e.g. PIE *pōd-s M 'foot', Lat. *pēs*, Gr. *poús*, Skr.

[29] For the history of the proposal, see Kortlandt (1983), Rumsey (1987). For more recent versions of the hypothesis, see Schmalstieg (1981), Beekes (1995).

[30] Note that Kartvelian and many NE Caucasian languages also have suppletive demonstrative stems, one of which is used in the absolutive case, and the other in oblique cases, including the ergative.

pāt, IEW 790) simply generalised their old ergative case form as the new nominative. Those that became neuters, like, foe example, *$h_3$nomn N 'name' (Lat. *nōmen*, Gr. *ónoma*, Skr. *nā́ma*, IEW 321) generalised the old absolutive form which became their nominative-accusative. This would not contradict Silverstein's (1976) case-marking hierarchy (as assumed by Rumsey 1987, among others), because it would imply that originally all nouns (not just the animates) had ergative case inflection. In this scenario, the original ergative case ending is preserved as the sigmatic genitive of all Late PIE nouns. Needless to say, this is a highly speculative hypothesis in need of further elaboration (Matasović 2011a).

On the whole, the case for ergativity in PIE is not conclusive; the evidence is consistent with the hypothesis that PIE was nominative-accusative since time immemorial, but it also does not contradict the hypothesis that it was ergative at some earlier stage of its development.[31]

In the whole of Eurasia, ergative languages are found only on the extreme northeastern periphery (Chukchi-Kamchatkan and Eskimo-Aleut), the extreme western periphery (Basque), and in the two residual zones, namely the Himalayas-Hindukush region and the Caucasus, where ergative languages are the norm (Figure 5). There are no ergative Uralic and Altaic languages, and there is a general consensus that the respective proto-languages were nominative-accusative.

However, the claim that ergativity is an areal feature of Caucasian languages is somewhat misleading (Tuite 1999). All indigenous Caucasian languages are ergative in the sense that in at least some constructions, the single arguments of intransitive verbs are morphosyntactically treated like direct objects of transitives, rather than as subjects of intransitive verbs. However, there are substantial differences in the clause alignment of NW Caucasian, NE Caucasian and Kartvelian languages. In very broad terms, Late Proto-Kartvelian had a split ergative system (Harris 1985), and a kind of split ergative system is found in the NE Caucasian language Batsbi, which is spoken on the southern periphery of the NE Caucasian area (Holisky 1987). Early Proto-Kartvelian may have had true ergativity with no split. The two northern Caucasian families also differ in the expression of ergativity. In NE Caucasian, languages generally have the ergative-absolutive case-marking pattern, but they usually code the

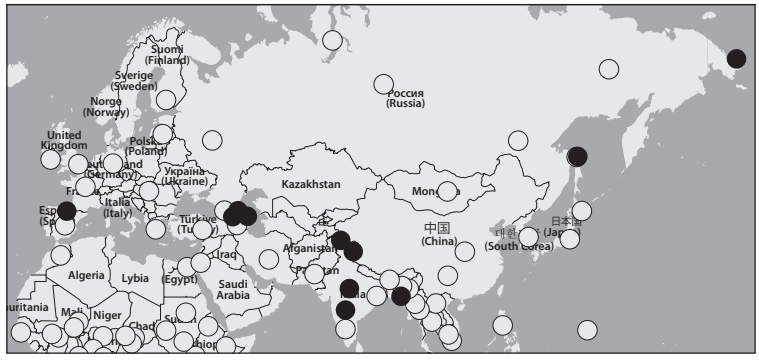

Figure 5. Languages without nominative-accusative case marking of nouns in Eurasia (black dots)[32]

[31] For limitations of space we shall not discuss here the hypothesis that PIE had active clause alignment, as advocated by Lehmann (1993) and Gamkrelidze & Ivanov (1984), among others. For a critique of this view see Matasović (2000).

[32] Languages with active and tripartite systems from the WALS database are grouped together with those with ergative case marking.

absolutive argument on the verb by gender-agreement affixes, while in NW Caucasian, the most archaic pattern is found in Abkhaz-Abaza, where ergativity is expressed only by personal prefixes on the verb. The ergative case marking on nouns in Ubykh and Circassian is a clear innovation. Thus, if the PIE case inflection really was ergative at some stage, it shows typological affinities with NE Caucasian, but not necessarily with the syntactic patterns found in other Caucasian languages. Note that many extinct languages of the ancient Near East were also ergative, including Sumerian, Hattic and Hurrian.

4.2.2. Inclusive-exclusive opposition in pronouns

The presence of inclusive-exclusive opposition in pronouns is not generally assumed for PIE, and none of the early dialects have that opposition. However, it is beyond doubt that at least two different stems for the first person plural pronoun can be reconstructed in PIE, namely *we- (Skr. *vayam,* ToB *wes,* Goth. *weis*) and *mes/*ns- (Lat. *nōs*, Lith. *mes*, OCS *my*, Arm. *mek*^{c}, etc.).[33] The curious thing is that the stem *we- can also be posited for the second person plural (Lat. *vōs,* Gr. *hymẽ̃is*, Skr. clitic Acc. *vas*), and it also occurs in the first and second person dual (cf. OCS 1. DU *vě* and 2. du. *va*, ToB *wene,* Skr. 1 DU *āvām,* 2 DU *yuvām*). In order to explain this unusual homonymy, it has been claimed that this stem reflects the PIE inclusive pronoun, while the other stem, *mes/*ns-, reflects the exclusive pronoun (Gamkrelidze & Ivanov 1984).[34] Although this cannot be proved, the reconstruction of inclusive/exclusive opposition in PIE elegantly explains the awkward homonymy of the stem *we- (which refers to both the first and the second person plural). The opposition between inclusive and exclusive pronouns is common in North Caucasian, and it may be posited for Proto-NE Caucasian (Nikolayev& Starostin 1994; Nichols 2003). Today we find it in the Nakh languages (e.g. Chechen *way* 'we (incl.)' vs. *txuo* 'we (excl.)'), Avar, the Andic languages and some Lezgic languages. It appears that some languages have generalised the original exclusive pronoun, and others the inclusive. In West Caucasian, the opposition between inclusive and exclusive exists in Abkhaz-Abaza, but not in Kabardian and Adyghe. In Kartvelian, Svan has an inclusive vs. exclusive distinction, and it has been claimed that this distinction goes back to Proto-Kartvelian (Oniani 1965), but this is not universally accepted. Since this opposition is otherwise rather rare in Eurasia (except on its eastern fringes), its co-occurrence in PIE and North Caucasian would certainly be significant, especially since this feature is demonstrably areally, but not genetically stable (it may spread through areal diffusion, cf. Nichols 1992: 123f.). The opposition is not attested in Uralic or Turkic, but it exists in Mongolic (cf. Mongol *bida* 'we (incl.)', *ba* 'we (excl.)') and in some Tungusic languages (Evenki *mit* incl. vs. *bu* excl.). The areal distribution of the inclusive-exclusive opposition in Eurasia is shown in Figure 6.

However, it must be emphasised that there are other explanations for the origin of the two different stems of the first person plural pronoun in PIE. It is also possible that *we- is originally the dual stem which spread to the plural only in some languages, while the original plural stem is *mes/*ns-. This would be in accordance with the fact that *-we is also used as the dual ending in the verbal inflection (cf. OCS *jes-vě* 'we two are', Skr. *svás*); the original dual pronoun meaning 'me and you (SG)' could have been re-interpreted as 'we' in one group of languages, and as 'you' in another. If this is correct, it is easy to understand the 'elliptic'

[33] It is possible that the original stem was only *nV-, while *me- was independently introduced in the nominative plural (in Balto-Slavic and Armenian) on the analogy with the verbal ending *-mes (Skr. *–mas*, Gr. Dor. *–mes*, etc.).

[34] This was already suggested by Jensen (1930) and subsequently by Watkins (1969: 47). For a critique, see Babaev (2009) and Yakubovich (2010: 209–10). Note also that, according to Cysouw (2003: 84), 'in almost all cases, the exclusive "we" is marked by the same morpheme that is used for the first person singular', which is not the case in PIE (the 1 SG pronoun is *eg'h$_2$om > Lat. *egō*, Skr. *ahám*, etc.).

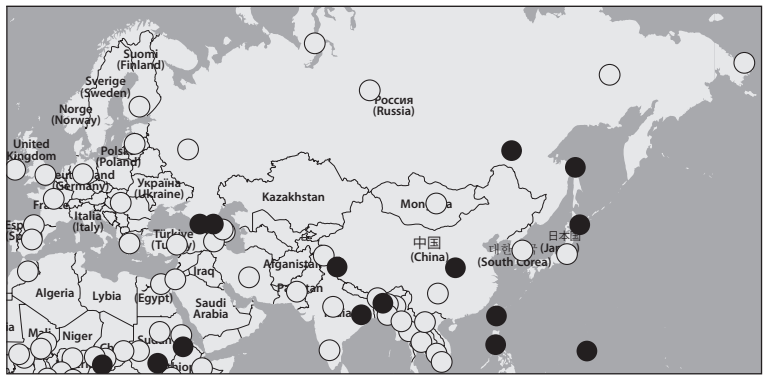

Figure 6. Inclusive-exclusive opposition in Eurasia

use of the dual, attested in the early IE dialects, e.g. Gr. (Hom.) DU *Kástore* 'Castor and the other one (sc. Polydeuces)', Skr. (Rig-Veda. 1, 14, 3) *Mitrā́* 'Mitra and the other one (sc. Varuna)'. Thus, *we- may originally have been interpreted either as 'you (and the other one)' > 'you (PL)', or 'me (and the other one)' > 'we'.[35] Hence, the frequency of the inclusive/exclusive opposition in the Caucasus is of limited importance for the areal typology of PIE.

4.2.3. Number suppletion in personal pronouns

As is well known, PIE had number suppletion in personal pronouns, that is, the plural (and dual) forms of personal pronouns were formed from a different stem from the singular forms, cf. PIE *eg'-$h_2$om 'I' (Skr. *ahám,* OCS *azъ*) vs. *wey-/ne- 'we' (Skr. *vayám,* Lat. *nōs*), PIE *tuH 'thou' (Lat. *tū,* OCS *ty*) vs. *yuH- (Skr. *yūyam*, Lith. *jūs*). This pattern is not at all characteristic of Northern Eurasia, and it cannot be posited for either Proto-Uralic or for Proto-Altaic.[36] It is found, however, with great frequency, in the Caucasian languages (data from Vinogradov 1967).

There are a few exceptions to this pattern in NE Caucasian languages, for example, in Dargwa, where we find the plural forms with the suffix *–ša* added to the singular (*nu* 'I' vs. *nuša* 'we', *Gu* 'thou' vs. *Guša* 'you'). Nikolaev and Starostin (1994) posit the number suppletion of personal pronouns in Proto-NEC, and the same suppletion can be posited for Proto-NWC (Chirikba 1996). In Kartvelian, Klimov (1964) reconstructs suppletive stems in the singular and plural (*me(n)- 'I', *sen- 'thou', *čwen- 'we', and *(s_1)tkwen- 'you'). Note that the number suppletion in personal pronouns is not an exclusive PIE-Caucasian feature, since it is also found in other language families (e.g. Dravidian, cf. Andronov 1970, Steever 1998), though rarely in Northern Eurasia (Figure 7).

In light of its relative rareness in Northern Eurasia, we believe that this parallelism between PIE and the Caucasian language families is areally significant.

[35] The form of the pronoun should perhaps be reconstructed as *$h_1$we- and compared with *$h_1$wi(d)k'mtih$_1$ 'twenty' (Gr. *eíkosi*, Lat. *vigintī*, etc.). If so, this may represent an instance of the alternation of *d and *h_1 (as, e.g., in *mēd and *meh$_1$ 'me', cf. Lat. *mēd* vs. Skr. *mām*), or in *med- 'measure' (Lat. *medeor* 'cure', OIr. *midithir* 'judges') and *meh$_1$- 'measure' (OCS *měra* 'measure', Skr. *mā́ti* 'measures').

[36] The suppletion is found in some Uralic languages, but it is secondary, as in Hungarian *én* 'I' vs. *mi* 'we' (cf. *te* 'thou' vs. *ti* 'you' without suppletion). For Altaic see Poppe (1964: 194). There was also no suppletion in Proto-Semitic (Moscati 1964).

Table 1. Number suppletion in personal pronouns in the Caucasus

Language	I	thou	we (incl./excl.)	you
Chechen	suo	ho	way/txuo	šu
Batsbi	so	ho	way/txo	šu
Avar	dun	mun	niž/nil	nuž
Botlikh	den(i)	mun	il'i/iśi	bišti
Godoberi	den	min	il'e/iśe	bitte
Akhvakh	dene	mene	il'i/issi	ušdi
Tindi	de	me	il'a/iśa	bissa
Khwarshi	do	mo	ilo	mižo
Hunzib	da	mə	ile	miže
Lak	na	ina	žu	zu
Lezgian	zun	vun	čun	kin
Agul	zun	vun	hin/čin	kin
Tsakhur	zy	vu	ši	šu
Kryz	zin	vun	jin/žin	vin
Khinalug	za	va	jir	zur
Udi	zu	un	jan	vaˤn
Abkhaz	sara	wara	hara	šʷara
Kabardian	sa	wa	da	fa
Georgian	me	šen	čven	tkven
Svan	mi-	si	šgvej-	sgäj-

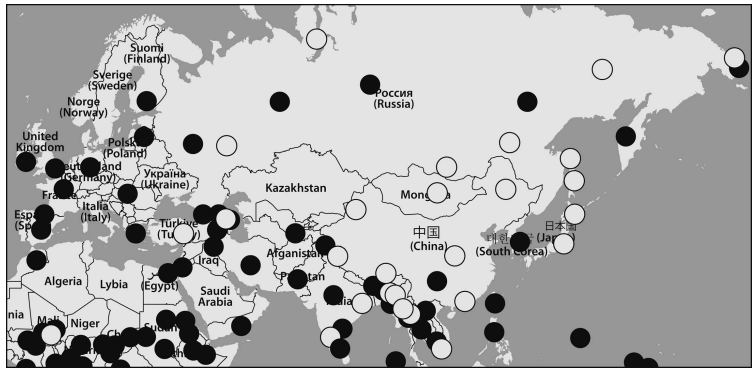

Figure 7. Languages with number suppletion in personal pronouns (black dots)

4.2.4. Two-stem inflection

One of the most salient features of NE Caucasian inflection, attested in all branches of that family, is the two-stem inflection of nouns. One stem is used for the nominative (absolutive) case, whereas the other is used for the ergative and other oblique cases (Kibrik 1991). This is strongly reminiscent of the 'heteroclitic' inflection of the PIE neuters, which form the nominative and accusative singular with the stem ending in *-r, and the oblique cases with the stem ending in *-n, cf., e.g., PIE *yēkʷr (NOM and ACC SG) vs. *yekʷns (GEN SG) 'liver', cf. Gr. *hẽpar, hếpatos*, Lat. *iecur, iecinis*, IEW 504. Although heteroclitic stems are an archaism in most IE languages, in Anatolian they are quite productive, which testifies that, at least in Early PIE, they were quite common.

This type of inflection is otherwise rare in the languages of North and Northeastern Eurasia, so its occurrence in PIE and NE Caucasian appears even more important. Note that it is at present impossible to reconstruct complete nominal paradigms in Proto-NE Caucasian, and that formal means of expression of the two-stem opposition differ in various languages

(Alekseev 2003), but several different suffixes used to form the oblique stem can be posited; in Chechen, for example, the oblique stem can be formed with the nasal suffix. Thus, we have Chechen *buhʕa* 'owl' (Absolutive SG) vs. *buhʕ-an-uo* (Ergative), *buhʕ-an-ash* (Absolutive Plural), or Dargi *neš* 'mother', oblique *neš-li* (dative *neš-li-s*), plural *neš-ani*, oblique plural *neš-an-a-* (dative *neš-an-a-s*). After discussing the evidence, Alekseev (2003: 34) concludes that the heteroclitic inflection of this type is original in the NE Caucasian family.

The fact that two-stem inflection is a trait shared exclusively by PIE and NE Caucasian is areally highly significant. However, one has to bear in mind that in PIE only neuters showed this feature, while in NE Caucasian it is attested across the lexicon.

4.2.5. Gender and nominal agreement

Nominal agreement may be defined as agreement with grammatical features of the noun, usually gender, number and case, within the noun phrase;[37] the PIE system of gender agreement is typologically highly unusual in that its sole domain is the NP (Matasović 2004: 204); in all other language families in Eurasia that have gender, the clause is also the domain of gender agreement, which means that at least some verb forms agree in gender with the subject (and sometimes with the object) of the clause. In PIE, the verb agreed with its subject in number and person,[38] and only the nominal modifiers in the NP agreed with the head noun. PIE was traditionally reconstructed with three genders (masculine, feminine and neuter), but it appears probable that the earlier system, which may be preserved in the Anatolian languages, had only two genders (Matasović 2004): common (for all animates and a large number of inanimates) and neuter (for the remaining inanimates).

Just like PIE, North Caucasian languages have grammatical gender. In NE Caucasian, it is more usual to talk about 'noun classes', but the phenomenon is essentially the same as the gender system in Indo-European.[39] The number of genders in NE Caucasian languages varies from two (in Tabasaran) to perhaps eight (in Batsbi, depending on the analysis adopted), but most languages have three (e.g. Avar) or four genders (e.g. Archi), and the system with four genders could be original. In contrast to PIE, in NE Caucasian verb is the main locus of gender agreement (it agrees with the absolutive argument in the clause), but in some languages adnominal adjectives and/or pronouns also agree (e.g. in Khwarshi, Khalilova 2009: 278–82). The Abkhaz-Adyghe languages of the NW Caucasus do not show much nominal agreement today, but there are clear indications that it had been more widespread. Abdokov (1981) argues that the NW Caucasian languages lost the inherited North Caucasian gender agreement, and the contemporary gender system of Abkhaz and Abaza is an innovation. In present-day Abkhaz-Abaza, demonstratives and numerals agree with nouns within the NP (the verbs also agree within the clause). Finally, the Kartvelian languages do not have gender and, despite the opinion of some Georgian scholars (see, e.g. Vinogradov 1967: 21, with references), there is no evidence that they ever had it.

[37] There are also other types of nominal agreement, involving definiteness (e.g. in Modern Hebrew), honorific agreement (e.g. in Korean), etc., but all of these are rather rare cross-linguistically and will not be treated in this paper.

[38] It is dubious whether 'person agreement' is really agreement in the same sense as gender and number agreement; there are certainly some theoretical and areal-typological reasons not to treat it as such: languages with gender, number, and case agreement have an areally much more restricted distribution than languages with person 'agreement', which are clearly the default type cross-linguistically. Moreover, it could be argued that the verb does not 'agree' with its arguments in person, but that it rather indexes one of the arguments, especially in pro-drop languages where free NPs need not be expressed, and there need be no NPs for the verb to agree with in person. For discussion see Corbett (2006).

[39] Three NE Caucasian languages (Udi, Agul, Lezgian) lost gender, presumably under the influence of the neighbouring Turkic languages.

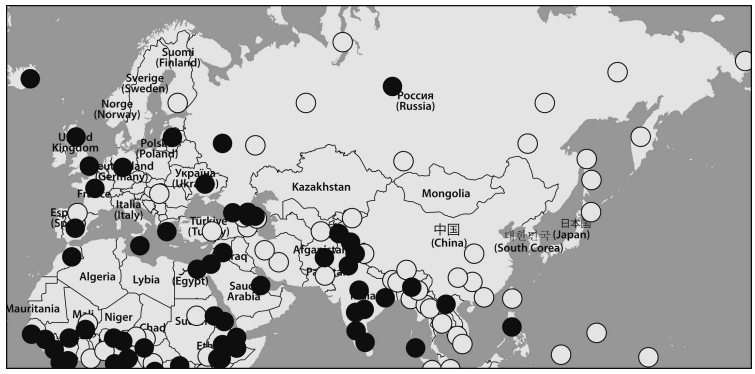

Figure 8. Languages with gender in Eurasia (black dots)

As argued by Matasović (2004) the distribution of languages with and without gender is not random within Eurasia: language families of the Northeastern half of Eurasia are generally genderless, while those of the Southwestern half have gender (Figure 8). The same distribution holds for other types of nominal agreement (chiefly case and number agreement, cf. Matasović 2010). Moreover, this applies to both ancient and modern languages, so the distribution appears to be diachronically stable. It is generally assumed that there was no nominal agreement in Proto-Uralic, though some modern Uralic languages in Europe (notably Finnish, Estonian and Saami)[40] have developed limited nominal agreement, presumably under the influence of the neighbouring IE languages. The same holds for the Altaic languages in general, where nominal agreement is generally non-existent.[41] On the other hand, it would not be fair to say that gender is typical only of Caucasian languages: it is a larger areal phenomenon in SW Eurasia, found also in Dravidian, Semitic and the extinct languages of the Near East (with the exception of Hurrian).

It is certainly important for the areal typology of PIE that it had gender, just like the NW and NE Caucasian languages, but unlike the overwhelming majority of languages of Northern Eurasia. Note, however, that in NE Caucasian, where the category of gender is certainly old, the domain of gender agreement is rather unlike the one found in PIE.

4.2.6. The presence of (inflectional) optative mood

The function of the optative was the expression of wish or desire, at least in the main clauses (in dependent clauses it marked the subordinated status of the verb). It is safely reconstructed in PIE as a category formed with the suffix *-yeh$_1$-/-ih$_1$- (e.g. Skr. *syāt* 'may (s)he be'). However, as Figure 9 shows, the inflectional optative is extremely rare in the languages of Eurasia, but it is quite common in the Caucasus, occurring in both the indigenous languages (NE and NW Caucasian, Kartvelian), and in the 'intruders' (Azerbaijani, Kumyk, Noghai). Note that the optative is lost in the history of most IE languages, but it is preserved in the early stages of Greek, Sanskrit and Gothic. In some languages, the form of the optative is preserved, but its function changed (e.g. in Slavic the optative became a plain imperative, and in Latin some optative forms became subjunctives). It is at present unknown whether optative

[40] Hungarian has number agreement with only one demonstrative pronoun.

[41] Some languages, like Even, have optional case/number agreement, but it seems to be motivated by pragmatic factors.

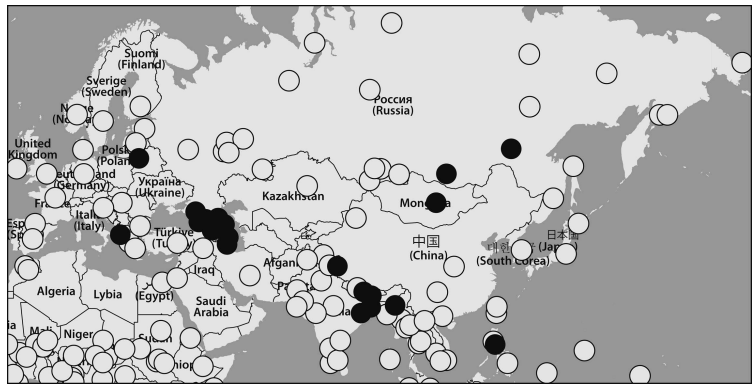

Figure 9. Languages with a inflectional optative in Eurasia (black dots)

should be reconstructed for Proto-NE Caucasian, but it may be safely posited for Proto-NW Caucasian.

The evidence discussed in this chapter suffices to show, in our opinion, that the areal-typological parallels between PIE and the languages of the Caucasus are not accidental. This particularly applies to NE Caucasian, which shares at least five of the ten areally diagnostic traits discussed in this section (Table 2). If the reconstruction of PIE favoured by the present writer were accepted, the agreement in typological traits would be even more probative of early contacts (PIE would agree with the North Caucasian languages in all traits but the presence of clusivity in personal pronouns). However, this would not reflect the *communis opinio* of Indo-European scholars.

5. Diagnostic areal features of PIE without parallels in the Caucasus

The picture of the areal typology of PIE would not be complete if we ommitted a few areally diagnostic traits of PIE grammar that are, as a rule, not found in Caucasian languages. One such feature is the presence of the dual in both nominal, pronominal and verbal morphology. The existence of PIE dual is beyond reasonable doubt. It was formed with the suffix *-h_1 (for masculines) namely, *-ih_1 (for feminines and neuters), for example, PIE *$h_3$ek^w-ih_1 'eyes' (neuter) > Gr. (Hom.) *ósse*, OCS *oči*, etc. Although the dual is unattested in Anatolian, it is quite probable that it was lost in that branch of Indo-European.[42] The Hittite first-person plural present ending *–weni* is usually derived from the PIE dual ending with the formant *-we- (rather than plural *-me-, as in Gr. *–men*, Skr. *–mas*, etc.). As shown in Figure 10, the distribution of languages with the dual is highly restricted in Eurasia. There are, to my knowledge, no languages with the dual spoken on the Caucasus.[43]

One areal feature of Caucasian languages that is notably absent from PIE is the vigesimal numeral system. Although languages with the decimal system predominate in the whole of Eurasia, languages with the vigesimal system are the norm in the Caucasus (Figure 11). In the

[42] For the dual in PIE see especially Fritz (2011), where it is argued that the category may be recent in PIE. However, the PIE dual marker *-h_1 can be compared with the Proto-Uralic dual marker *-k (Kortlandt 2001), and the dual seems to be an old category in Uralic (cf. Collinder 1960, Abondolo 1997).

[43] The so-called 'dual' in the Khevsurian and Pshavian dialects of Georgian is actually a paucal, reserved for groups of two-three instances (Boeder 2005). The suffix expressing this category comes from one of the Proto-Kartvelian plural suffixes. Alekseev (2003: 178) discusses the possibility that some affixes expressing the plural in Dagestanian languages developed from earlier dual markers, but rejects it.

Table 2. Summary of the structural-typological parallels

Trait	PIE	NE Caucasian	NW Caucasian	Kartvelian
(moderately) high c/v ratio	+/–	+	+	+
glottalised consonants	+/–	+	+	+
tone distinctions	+	+/–	+/–	–
productive Ablaut	+	+	+	+
ergativity	+/–	+	+	+/–
clusivity	+/–	+	+/–	+/–
number suppletion in pronouns	+	+	+	+
two-stem inflection	+	+	–	–
gender	+	+	+/–	–
morphological optative	+	+	+/–	+

Note: +/– denotes that there is no *communis opinio* among scholars with respect to the feature in question.

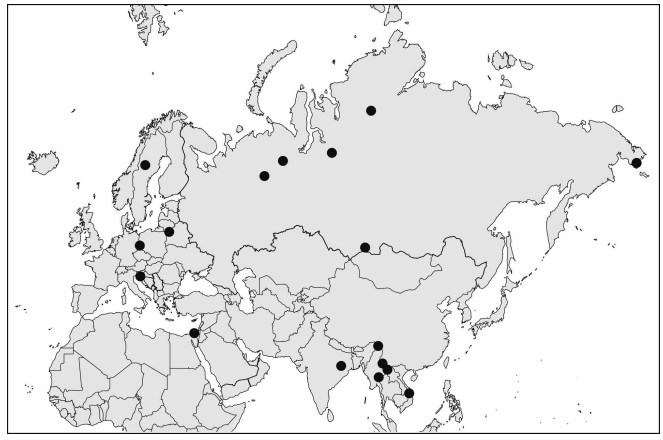

Figure 10. Languages with the dual number in Eurasia (from a sample of 150 languages)

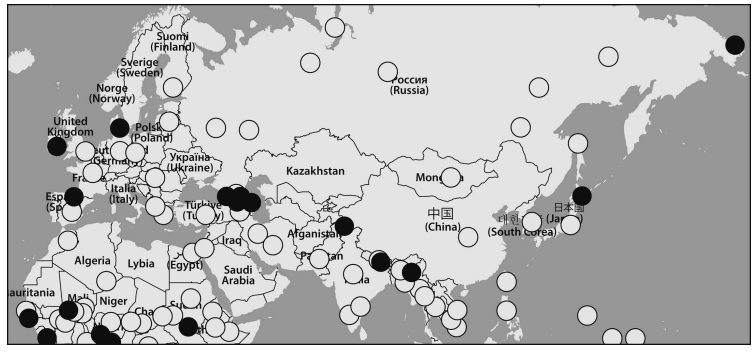

Figure 11. Languages with the vigesimal numeral base (black dots)[44]

WALS database the majority of indigenous Caucasian languages have the vigesimal system, and for some (e.g. Kabardian) it can be shown that the system was originally vigesimal, but that the language adopted the decimal base only recently.

[44] Languages with hybrid vigesimal-decimal and those with pure vigesimal number bases are grouped together.

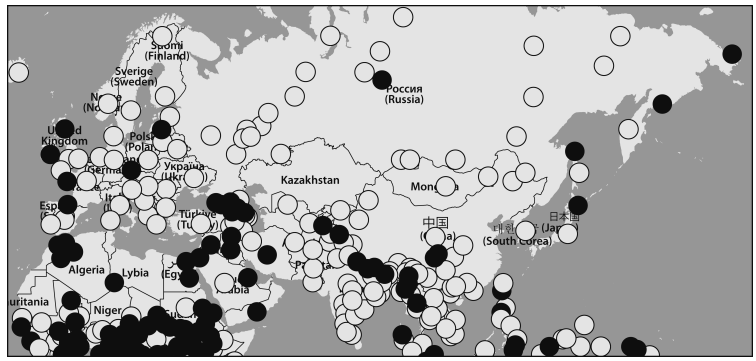

Figure 12. Prefixing and prefixing-suffixing languages in Eurasia (black dots)

Finally, it should be noted that PIE was an exclusively suffixal language (see, e.g., Szemerényi 1989).[45] The same applies to the majority of the language families of North and Northeastern Eurasia, including Uralic and Altaic (Figure 12). On the other hand, all three of the language families of the Caucasus have at least some prefixation, and in NW Caucasian prefixes actually predominate over suffixes. Prefixal morphology can be posited for all three Caucasian proto-languages.

The fact that we find certain typological traits in the Caucasus that are very different from those in PIE had to be mentioned for completeness' sake. It does not contradict early contacts between those languages. In language areas not all typological traits need be shared among languages belonging to those areas (i.e., forming a *Sprachbund*). What we do find is rather a kind of 'family resemblance' pattern, where most features are shared by most languages, but it is not necessary for all languages to have all the features characteristic of the *Sprachbund*. After all, Modern Greek is a very typical Balkan language, but it did not develop a postponed definite article and the vowel 'shwa' (ə) like Albanian, Bulgarian and Romanian.

6. Conclusions

This paper attempted to re-examine the evidence for early contacts between PIE and the languages of the Caucasus. Although we were not able to find certain proofs of lexical borrowing between PIE and North Caucasian, there are a few undeniable areal-typological parallels in phonology and grammar. Some features generally attributed to PIE are not found in the majority of languages of North and Northeastern Eurasia, while they are common, or universally present, in the languages of the Caucasus (especially North Caucasus). Those features of PIE that find their parallels only in North Caucasian languages could have been much more widespread in the fourth millennium BC, when language contacts between PIE and the Caucasian languages were likely. Of course, the Caucasian languages show a great deal of internal typological variability, and one must beware of taking random typological similarities as probative of pre-conceived hypotheses about prehistoric contacts; there is so much typological diversity in the Caucasus that we would probably be able to find at least some typological similarities between any language in the world and one, or a few, languages of the Caucasus. It is only when such typological similarities occur in bundles that they cannot be dismissed lightly.

The adduced typological parallels between PIE and Caucasian languages make it likely that PIE was, indeed, in contact with languages of the northern Caucasus. However, these contacts

[45] Of course, all early IE languages have prefixes, but these go back to PIE adpositions or nominal roots occurring in compounds.

could also have been of indirect nature, since there are no demonstrable loanwords from North Caucasian languages in PIE, or vice versa. If such loanwords exist, their number is certainly not high. If direct contacts did exist, we cannot determine their nature: both long-term bilingualism due to exogamy and trade networks, as well as rather rapid language shift appear equally possible. The evidence presented in this paper is in line with the widely-held theory that the PIE homeland was somewhere in the region to the north of the Black Sea (Anthony 2007). We believe that our arguments make it likely that it was in the eastern part of that region.

Finally, the weight of the evidence discussed in this paper clearly depends on the view of PIE one adopts: the reconstruction of several areally diagnostic features, such as glottalised consonants, ergative clause alignment, and the opposition between inclusive and exclusive pronouns, remains disputed. For those who accept that PIE had these features, or a selection of them, prehistoric contacts between PIE and the indigenous languages of the Caucasus will be a much more plausible hypothesis.

Department of Linguistics
University of Zagreb
I. Lučića 3, HR-10000
Zagreb, Croatia
Email: rmatasov@ffzg.hr

References

Abaev, Vasily I., 1973. *Istoriko-ètimologičeskij slovar' ossetinskogo jazyka*, Leningrad: Nauka.
Abdokov, Aues I., 1981. *Vvedenie v sravnitel'no-istoričeskuju morfologiju abxazsko-adygskix i naxsko-dagestanskix jazykov*, Nal'čik: Kabardino-balkarskij Gosudarstvennyj Universitet.
Abondolo, Daniel (ed.), 1997. *The Uralic Languages*, London: Routledge.
Aikhenvald, Alexandra, 2001. 'Areal diffusion and language contact in the Içana-Vaupés basin, north-west Amazonia', in R. M. W. Dixon & Alexandra Aikhenvald (eds), *The Amazonian Languages*, Cambridge: Cambridge University Press, 385–415.
Alekseev, M. E., 1985. *Voprosy sravnitel'no-istoričeskoj grammatiki lezginskix jazykov: morfologija, sintaksis*, Moscow: Nauka.
Alekseev, M. E., 1988. *Sravnitel'no-istoričeskaja morfologija avaro-andijskix jazykov*, Moscow: Nauka.
Alekseev, M. E., 2003. *Sravnitel'no-istoričeskaja grammatika naxsko-dagestanskix jazykov. Kategorii imeni*, Moscow: Academia.
Andronov, M. S., 1970. *Dravidian Languages*, Moscow: Nauka.
Anthony, David, 2007. *The Horse, the Wheel, and Language*, Princeton, NJ: Princeton University Press.
Babaev, Kyrill, 2009. 'The problem of clusivity in Indo-European vs. Nostratic', *Aspekty komparativistiki* 3, 112–24.
Beekes, Robert S. P., 1995. *Comparative Indo-European Linguistics*, Amsterdam: Benjamins.
Boeder, Winfred, 2005. 'The South Caucasian languages', *Lingua* 115, 5–89.
Bokarev, Eŭgeno A., 1981. *Sravnitel'no-istoričeskaja fonetika vostočnokavkazskix jazykov*, Moscow: Nauka.
Carpelan, Christian & Parpola, Asko, 2001. 'Emergence, contacts and dispersal of Proto-Indo-European, Proto-Uralic and Proto-Aryan in archaeological perspective', *Mémoires de la Société finno-ougrienne* 242, 55–50.
Chirikba, Vyacheslav A., 1996. *Common West Caucasian*, Leiden: Research School CNWS.
Colarusso, John, 1981. 'Typological parallels between proto-Indo-European and North-West Caucasian languages' in Joel Arbeitman (ed.), *Bono Homini Donum. Essays in historical linguistics. In memoriam of Alexander Kerns*, Amsterdam: Benjamins, 475–557.
Colarusso, John, 1997. 'Phyletic links between Proto-Indo-European and Proto-Northwest Caucasian', *The Journal of Indo-European Studies* 25, 119–51.
Colarusso, John, 2003. 'Further etymologies between Indo-European and Northwest Caucasian', in Dee Ann Holisky & Kevin Tuite (eds), *Current Trends in Caucasian, East European and Inner Asian Linguistics, Papers in Honor of Howard Aronson*, Amsterdam: Benjamins, 41–60.
Collinder, Björn, 1960. *Comparative Grammar of the Uralic Languages*, Stockholm: Almqvist & Wiksell.
Corbett, Greville, 2006. *Agreement*, Cambridge: Cambridge University Press.
Cysouw, Michael, 2003. *The Paradigmatic Structure of Person Marking*, Oxford: Oxford University Press.
De Vaan, Michiel, 2008. *Etymological dictionary of Latin and the Other Italic Languages*, Leiden: Brill.
Dolgopolsky, Aaron, 1987. 'The Indo-European homeland and lexical contacts of Proto-Indo-European with other languages', *Mediterranean Language Review*, 3, 7–31.

DYBO, VLADIMIR A., 1989. 'Tipologija i rekonstrukcija paradimatičeskix akcentnyx sistem', in Radmila V. Bulatova and Vladimir A. Dybo (eds), *Istoričeskaja akcentologija i sravnitel'no-istoričeskij metod*, Moscow: Nauka, 7–45.
DYBO, VLADIMIR A., 2003. 'Balto-slavjanskaja akcentologičeskaja rekonstrukcija i indoevropejskaja akcentologija' *Slavjanskoe jazykoznanie. XIII meždunarodnyj s'jezd slavistov*. Moscow: Indrik, 131–61.
FRITZ, MATTHIAS, 2011. *Der Dual im Indogermanischen*, Heidelberg: Winter.
GAMKRELIDZE, TAMAZ V., 1966. 'A typology of Common Kartvelian', *Language* 42, 69–83.
GAMKRELIDZE, TAMAZ V. & IVANOV, VJAČESLAV V., 1984. *Indoevropejskij jazyk i indoevropejcy, I-II*, Tbilissi: Izdatel'stvo tbilisskogo universiteta.
GAMKRELIDZE, TAMAZ V. & MAČAVARIANI, GIVI I., 1982. *Sonantensystem und Ablaut in den Kartwelsprachen: eine Typologie der Struktur des Gemeinkartwelischen*, Tübingen: Günter Narr.
GIGINEJŠVILI, B. K., 1977. *Sravnitel'naja fonetika dagestanskix jazykov*, Tbilissi: Izdatel'stvo tbilisskogo universiteta.
GIMBUTAS, MARIJA, 1970. 'Proto-Indo-European culture: the Kurgan culture during the fifth, fourth, and third millennia B.C.', in George Cardona, Henry Hoenigswald and Alfred Senn (eds.), *Indo-European and the Indo-Europeans*, Philadelphia, PA: University of Pennsylvania Press, 155–98.
GIPPERT, JOST, 1995. 'Die Glottaltheorie und die Frage früher indogermanisch-kaukasischer Sprachkontakte', in Jens Elmegard Rasmussen (ed.), *In Honorem Holger Pedersen*, Wiesbaden: Reichert, 107–23.
GUY, GREGORY R., 1990. 'The sociolinguistic types of language change', *Diachronica* 7, 47–67.
HAARMANN, HARALD, 1994. 'Contact linguistics, archaeology and ethnogenetics: an interdisciplinary approach to the Indo-European homeland problem', *Journal of Indo-European Studies* 22, 265–88.
HARRIS, ALICE C., 1985. *Diachronic Syntax: The Kartvelian case* (Syntax and Semantics 18), New York: Academic Press.
HASPELMATH, MARTIN, DRYER, MATTHEW S., GIL, DAVID & COMRIE, BERNARD (eds), 2006. *The World Atlas of Language Structures*, Oxford: Oxford University Press.
HEWITT, GEORGE, 2004. *Introduction to the Study of the Languages of the Caucasus*, Munich: Lincom.
HOLISKY, DEE ANN, 1987. 'The case of the intransitive subject in Tsova-Tush (Batsbi)', *Lingua* 71, 103–32.
JENSEN, HANS, 1930. 'Bemerkungen zum urgeschlechtigen Personalpronomen des indogermanischen', *Indogermanische Forschungen* 48, 283–310.
KAISER, ELKE, 2010. Review of Anthony 2007, *Kratylos* 55, 35–44.
KALLIO, PETRI, 2004. *Studia Indo-Uralica. The Early Relations between Indo-European and Uralic*, PhD dissertation, University of Leiden.
KASSIAN, ALEKSEI, 2009. 'Hattic as a Sino-Caucasian Language', *Ugarit-Forschungen* 41, 309–448.
KHALILOVA, ZAIRA, 2009. *A Grammar of Khwarshi*, Utrecht: LOT.
KIBRIK, ALEKSANDR E., 1991. 'Organising principles of nominal paradigms in Daghestanian languages: comparative and typological observations', in Frans Plank (ed.), *Paradigms: The Economy of Inflection*, Berlin: Mouton de Gruyter, 255–74.
KLIMOV, GEORGIJ A., 1964. *Ètimologičeskij slovar' kartvel'skix jazykov*, Moscow: AN SSSR.
KLIMOV, GEORGIJ A., 1981. 'Neskol'ko kartvel'skix indo-evropeizmov', *Ètimologija* 11, 167–73.
KLIMOV GEORGIJ A. (ed.), 1978. *Strukturnye obščnosti kavkazskix jazykov*, Moscow: Nauka.
KLIMOV, GEORGIJ A., 1986. 'Ob areal'noj konfiguracii protoindoevropejskogo v svete dannyx kartvel'skix jazykov', *Voprosy drevnoj istorii* 3, 150–60.
KLIMOV, GEORGIJ A., 2010. *Drevnejšie indoevropeizmy kartvel'skix jazykov*, Moscow: KD Librokom.
KOHL, PETER L., 2007. *The Making of Bronze Age Eurasia*, Cambridge: Cambridge University Press.
KORENOVSKIJ, S. N., 2004. *Drevnejšie zemledelcy i skotovody Prekdavkaz'ja*, Moscow: KD Librokom.
KORENOVSKIJ, S. N., 2010. 'Fenomen bol'šix majkopskix kurganov: social'no-trudovoj i kul'tovyj aspekty stroitel'stva', *Kratkie soobščenija Instituta arxeologii RAN* 224, 149–70.
KORTLANDT, FREDERIK, 1989. *The spread of the Indo-Europeans*, Ms, Leiden University. Available online at http://www.kortlandt.nl.
KORTLANDT, FREDERIK, 2001. 'The Indo-Uralic verb'. Ms, Leiden University.
KORTLANDT, FREDERIK, 2009. 'C.C. Uhlenbeck on Indo-European, Uralic and Caucasian', *Historische Sprachforschung* 122, 39–47.
KULANDA, SERGEI V., 2008. '*Lošad'* v praindoevropejskom', *Aspekty komparativistiki* 3, 669–78.
LEHMANN, WINFRED P., 1993. *Theoretical Bases of Indo-European Linguistics*, London: Routledge.
MALLORY, JAMES, 1989. *In Search of the Indo-Europeans*, London: Thames and Hudson.
MALLORY, JAMES & DOUGLAS Q. ADAMS, 1997. *Encyclopaedia of Indo-European Culture*, London: Fitzroy Dearborn.
MANZURA, IGOR, 2005. 'Steps to the steppe. Or, how the North Pontic region was colonised', *Oxford Journal of Archaeology* 24, 313–38.
MARKEY, THOMAS & GREPPIN, JOHN A. C. (eds), 1990. *When Worlds Collide: The Indo-Europeans and the Pre-Indo-Europeans*, Ann Arbor, MI: Karoma Press.
MARTIROSYAN, HRACH K., 2010. *Etymological Dictionary of the Armenian Inherited Lexicon*, Brill: Leiden.
MATASOVIĆ, RANKO, 1992. 'Distribucija indoeuropskoga fonema *r: kavkasko-indoeuropska tipološka podudarnost?', *Suvremena lingvistika* 34, 201–7.
MATASOVIĆ, RANKO, 2000. 'Uses and misuses of typology in Indo-European linguistics', in Michaela Ofitsch & Christian Zinko (eds.), *125 Jahre Indogermanistik in Graz*, Graz: Leykam, 327–40.
MATASOVIĆ, RANKO, 2004. *Gender in Indo-European*, Heidelberg: Winter.
MATASOVIĆ, RANKO, 2010. 'Nominal agreement in Proto-Indo-European from the areal and typological points of view', paper presented at the Jena Conference on Gender in PIE, July 2010, to appear in the Proceedings.

MATASOVIĆ, RANKO, 2011a. 'Clause alignment in Proto-Indo-European', Ms, University of Zagreb. Available online at: http://www.ffzg.hr/~rmatasov.
MATASOVIĆ, RANKO, 2011b. 'A short grammar of East Circassian (Kabardian)', Ms, University of Zagreb. Available online at: http://www.ffzg.hr/~rmatasov.
MELIKIŠVILI, IRINA G., 1980. 'Struktura kornja v obščekartvel'skom i obščeindoevropejskom', *Voprosy jazykoznanija* 4, 60–70.
MOSCATI, SABATINO, 1964. *An Introduction to the Comparative Grammar of the Semitic Languages*, Wiesbaden: Harrassowitz.
NICHOLS, JOHANNA, 1992. *Linguistic Diversity in Space and Time*, Chicago, IL: The University of Chicago Press.
NICHOLS, JOHANNA, 2003. 'The Nakh-Daghestanian consonant correspondences', in Dee Ann Holisky & Kevin Tuite (eds), *Current Trends in Caucasian, East European and Inner Asian Linguistics*, Amsterdam: Benjamins, 207–64.
NICHOLS, JOHANNA, 2007. 'A typological geography for Proto-Indo-European', in Karlene Jones-Bley, Angela Della Volpe, Miriam Robbins Dexter, & Martin E. Huld (eds), *Proceedings of the 18th UCLA Indo-European Conference*, Washington, DC: Institute for the Study of Man, 191–211.
NIKOLAYEV, S. L. & STAROSTIN, SERGEI A., 1994. *North Caucasian Etymological Dictionary*, Moscow: Asterisk.
ONIANI, A. A., 1965. 'Otnositel'no kategorii èkskluziva-inkljuziva v kartvel'skix jazykax', *Vestnik OON* 1: 230–34.
POKORNY, JULIUS, 1959. *Indogermanisches etymologisches Wörterbuch*, Bern: Francke.
POPPE, NICHOLAS, 1964. *An Introduction to the Study of the Altaic Languages*, Wiesbaden: Harrassowitz.
PRONK, TIJMEN, 2010. 'On Indo-European tones, accentuation, and ablaut', Ms, Leiden University.
RASSAMAKIN, JURYJ J., 1999. 'The Eneolithic of the Black Sea Steppe: dynamics of cultural and economic development 4500–2300 BC', in M. Levine, Y. Rassamakin, A. Kislenko, & N. Tatarintseva (eds), *Late Prehistoric Exploitation of the Eurasian Steppe*, Cambridge: Cambridge University Press, 59–182.
RASSAMAKIN, JURYJ J., 2004. 'Die nordpontische Steppe in der Kupferzeit. Gräber aus der Mitte des 5. Jts. bis Ende des 4. Jts. v. Chr.', *Archäologie in Eurasien* 17, 1–67.
RENFREW, COLIN, 1987. *Archaeology and Language. The Puzzle of Indo-European Origins*, London: Jonathan Cape.
RIX, HELMUT, KÜMMEL, MARTIN, ZEHNDER, THOMAS, LIPP, REINER & SCHIRMER, BRIGITTE (eds), 2002. *Lexikon der indogermanischen Verben*, Wiesbaden: Reichert.
ROSS, MALCOLM D., 1991. 'Refining Guy's sociolinguistic types of language change', *Diachronica* 8, 119–29.
RUMSEY, ALAN, 1987. 'The chimera of Proto-Indo-European ergativity', *Lingua* 71, 297–318.
ŠAGIROV, A. K., 1977. *Ètimologičeskij slovar' adygskix (čerkesskix) jazykov, I, II*, Moscow: Nauka.
SALMONS, JOSEPH, 1994. *The glottalic theory: survey and synthesis*, Washington DC: The Journal of Indo-European Studies Monographs, Institute for the Study of Man.
SCHMALSTIEG, WILLIAM, 1981. 'Ergativity in Indo-European', in Y. L. Arbeitman & A. Bomhard (eds), *Bono Homini Donum. Studies in Memory of J. A. Kerns*, Amsterdam: Benjamins, 242–58.
SCHMIDT, KARL HORST, 1967. 'Beiträge zu einer typologisch-vergleichenden Grammatik der indogermanischen und südkaukasischen Sprachen', *Münchener Studien zur Sprachwissenschaft* 22, 17–34.
SCHULZE, WOLFGANG, 1997. Review of NCED, *Diachronica* 14, 149–61.
SHERRATT, ANDREW, 1983. 'The secondary exploitation of animals in the Old World', *World Archaeology* 15, 90–104.
SHERRATT, ANDREW, 1998. 'Echoes of the Big Bang: the historical context of language dispersal', in Karlene Jones-Bley, Angela Della Volpe, Miriam Robbins Dexter, & Martin E. Huld (eds), *Proceedings of the Tenth Annual UCLA Indo-European Conference*, Washington, DC: Institute for the Study of Man, 261–82.
SHIMOMIYA, TADAO, 1978. *Zur Typologie des georgischen (verglichen mit dem indogermanischen). Mit einem exkurs zur Sprachbundtheorie*, Tokyo: Gakushuin.
SILVERSTEIN, MICHAEL, 1976. 'Hierarchy of features and ergativity', in Robert M. W. Dixon (ed.), *Categories in Australian Languages*, Canberra: Australian Institutes of Aboriginal Studies, 112–71.
SKRIBNIK, ELENA, 2011. 'South Siberian Turkic languages. Past language contacts', paper presented at the conference 'Indo-European Linguistics in the 21st Century' in the University of Pavia, September.
SOLTA, GEORG RENATUS, 1980. *Einführung in die Balkanlinguistik mit besonderer Berücksichtigung des Substrats und des Balkanlateinischen*, Darmstadt: Wissenschaftliche Buchgesellschaft.
STAROSTIN, SERGEI A., 2007a. 'Indoevropejsko-severnokavkazskie izoglossy', in Georgij S. Starostin (ed.), *Trudy po jazykoznaniju*, Moscow: Jazyki slavjanskix kul'tur, 312–58.
STAROSTIN, SERGEI A., 2007b. Review of V. A. Chirikba's *Common West Caucasian*, in Georgij S. Starostin (ed.), *Trudy po jazykoznaniju*, Moscow: Jazyki slavjanskix kul'tur, 682–744.
STAROSTIN, SERGEI A. & D'JAKONOV, IGOR M., 2007. 'Hurrito-urartskie i vostočnokavkazskie jazyki', in Georgij S. Starostin (ed.), *Trudy po jazykoznaniju*, Moscow: Jazyki slavjanskix kul'tur, 359–406.
STEEVER, SANFORD B. (ed.), 1998. *The Dravidian Languages*, London: Routledge.
SZEMERÉNYI, OSWALD, 1989. *Einführung in die vergleichende Sprachwissenschaft*, Darmstadt: Wissenschaftliche Buchgesellschaft.
THOMASON, SARAH G., 2001. *Language Contact*, Edinburgh: Edinburgh University Press.
TRUBETZKOY, NIKOLAI S., 1939. 'Gedanken über das Indogermanenproblem', *Acta Linguistica* I, 81–89 (reprinted in A. Scherer (ed.), 1968, *Die Urheimat der Indogermanen*, Darmstadt: Wissenschaftliche Buchgesellschaft, 214–23).
TUITE, KEVIN, 1999. 'The myth of the Caucasian Sprachbund: the case of ergativity', *Lingua* 108: 1–26.
UHLENBECK, CHRISTIAN C., 1901. 'Agens und Patiens im Kasussystem der indogermanicshen Sprachen', *Indogermanische Forschungen* 12, 170–72.
VAILLANT, ANDRÉ, 1936. 'L'ergatif indo-européen', *Bulletin de la Société de Linguistique de Paris*, 37, 93–108.
VAN DEN BERG, HELMA, 2005. 'The East Caucasian Language family', *Lingua* 115, 147–90.

VINOGRADOV, V. V. (ed.), 1967. *Jazyki narodov SSSR, IV: Iberijsko-kavkazskie jazyki*, Moscow: Nauka.
WATKINS, CALVERT, 1969. *Indogermanische Grammatik, III: Formenlehre*, Heidelberg: Winter.
YAKUBOVICH, ILYA, 2009. *Sociolinguistics of the Luvian Language*, Leiden: Brill.
YAKUBOVICH, ILYA, 2010. 'Review of 'Aspekty komparativistiki 4', *Voprosy jazykovogo rodstva* 4, 207–11.

ABBREVIATIONS

Alb.	Albanian
Arm.	Armenian
Av.	Avestan
CS	Church Slavonic
Croat.	(Standard) Croatian
Germ.	German
Goth.	Gothic
Gr.	Greek
Hitt.	Hittite
Hom.	Homeric
Lat.	Latin
Lith.	Lithuanian
NEC	North-East Caucasian
NWC	North-West Caucasian
OCS	Old Church Slavonic
OE	Old English
OHG	Old High German
OIc.	Old Icelandic
OIr.	Old Irish
OPr.	Old Prussian
PIE	Proto-Indo-European
PNC	Proto-North Caucasian
Russ.	Russian
ToA	Tocharian A
ToB	Tocharian B
W	Welsh