

Serhii V. Borshchevskyi

ORCID 0000-0002-5538-8036

Taras Shevchenko National University of Kyiv

VASCO-CAUCASIAN VECTOR IN THE STUDY OF GERMANIC VOCABULARY OF UNKNOWN ORIGIN

The heterogeneous characteristic features of the Indo-European languages in general and hybrid nature of the Germanic languages in particular have both become a popular object of multidisciplinary research and a highly disputable issue in recent scholarship.

As a thorough analysis of etymological dictionaries of modern and classical European languages shows, they contain a rather significant vocabulary layer that does not have satisfactory Indo-European etymology. In most cases, it is presupposed that such lexemes of uncertain or unknown origin were inherited from the ancient non-Indo-European languages of the Eastern Mediterranean as well as Northern and Western Europe, which is an argument for the thesis of the Indo-European languages emerging as a result of the imposition of the Indo-European proto-language or its dialects onto different Pre-Indo-European substrata (sublayer languages) [Mosenkis 2002: 6; Portsig 1964: 53, 62]. The solution to the problem of substratum donor influence upon the recipient language requires careful and comprehensive study through the synthesis of linguistics with history, archaeology and other related sciences [Tkachenko 1989: 4].

An overview of earlier and recent publications on the research problem. Analyzing non-Indo-European lexical relics of Indo-European languages, researchers have chosen different etymological strategies. Some scholars confined themselves to general hypothetical statements about the “non-Indo-European”, “Pre-Aryan”, “Proto-European”, “Aegean”, “Mediterranean”, “Caucasian”, “Alpine”, “Asia Minor”, etc. origin of such lexemes in the Slavic [Brückner 1974: 53], Baltic [Machek 1934], Germanic [Dyakonov 1984: 18], Celtic [Kalygin, Korolyov 1989: 21], Latin [Ernout, Meillet 1938] and Greek [Kretschmer 1896: 401–420] languages. Thus, the Polish Slavist Aleksander Brückner suggests the

Pre-Indo-European (or, according to the scientist's own terminology, "Pre-Aryan") origin of the Slavic nominations for *bull*, *goat*, *ram*, *Vistula* etc. [Brückner 1974: 53], the Czech etymologist Vaclav Machek surmises that a large part of the names of higher plants in the Slavic languages belong to the so-called "Proto-European" substratum [Machek 1954]. Numerous hypotheses of other linguists [Oštir 1930] proved to be insufficiently founded, since they were based on mostly random semantic coincidences of phonetically incompatible lexical comparisons that did not form reliable word-building rows. However, there were also such scholars who subjected non-etymologized lexical units to complex systematic analysis, by dividing the corpus into thematic groups and widely involving the material of non-Indo-European languages.

Since the end of the 19th century, when the theory of language "crossing" and "mixing" was gaining popularity, attempts have been made to substantiate the Basque-Slavic unity based on lexical (phonetic-semantic) correlations [Topolovšek 1894]. The Pre-Indo-European layer in the vocabulary of the Germanic group was generally associated with the languages of the Caucasian [Marr 1922] and Finno-Ugric families [Bubrich 1926], non-etymologized elements of the insular Celtic and Romance languages – with Afro-Asiatic (Semitic-Hamitic) [Pokorný 1927; Wagner 1931], or even more ancient (Pre-Berber-Libyan or Euro-African) influence [Hubschmid 1953].

On the one hand, the Pre-Indo-European Substratum Theory has long been regarded as a rather dubious realm of research due to the almost total absence of systemic approach to substratum studies, namely the lack of a comprehensive evidence base in the form of a reliable correlation between linguistic, anthropological and archaeological hypotheses, as well as the considerable complexity and chronological depth of the investigated material, insufficient development of methodology of its analysis, which led to scepticism concerning substratum researches and, consequently, to depopularisation of the Substratum Theory in scientific circles. This situation resulted in the refusal to further develop the substratum perspective on a concrete linguistic material.

On the other hand, we are now witnessing the revival of the Pre-Indo-European Substratum Theory in, to some extent, tendentious and biased studies of modern Ukrainian linguists Yurii Mosenkis, Roman Synyshyn, Dmytro Pereverziev who are influenced by the so-called "Japhetic" theory advanced in the early 20th century by Academician Nikolay Marr but now considered as pseudo-scientific and discredited¹.

The author of the article is deeply convinced that in relation to Marr's scientific legacy it is necessary to defy a number of firmly established stereotypes. This is primarily about the rehabilitation of the early publications of this, we can say without exaggeration, outstanding expert in the Caucasian languages who definitely

¹ <https://www.britannica.com/biography/Nikolay-Yakovlevich-Marr>

cannot be referred to as an “armchair” scientist. Hereby we mean the works written before his false “paleontological” method, based on the notorious “four-element” analysis that spoiled Marr’s later achievements and placed a significant part of his legacy beyond the course of science. Early works had been buried under the method and notoriety surrounding Marr’s name in general and, according to some orientalists, unfairly forgotten. As Academician Shalva Dzidziguri points out, “... even in Marr’s later works there are etymologies that are forgotten but have been repeated and are still being repeated by other scholars without reference to Marr [...] Marr’s many etymological ideas just await their further development” [Dzidziguri 1987: 7].

It is undeniable that Academician Marr outpaced contemporary linguists, historians, ethnographers and archaeologists who assume the appearance of the Caucasian tribes with a cultural mission in prehistoric Europe [Abaev 1986: 10; Lingvisticheskiy, red. 1990: 97; Mosenkis 2007: 8; Safronov 1989: 273], as evidenced in Modern English by traces of the Iberian language of the first settlers on the British Isles [Origin 2010: 8–9]. On our part, we should like to emphasize that both tendencies – total rejection, on the one hand, and unconditional adoption of the substratum on the other – are unacceptable in contemporary linguistic scholarship. Given the foregoing, we can suppose that the science of the 21st century will only benefit from the balanced attitude to the use of N. Ya. Marr’s etymological legacy.

Unfortunately, not all researchers were critical enough of Marr’s later “scientific” heritage – a multitude of investigations of the above-mentioned Ukrainian linguist Yu. L. Mosenkis is a case in point.

One cannot disregard the fact that Mosenkis and his followers, in addition to Slavic studies performed under the heavy influence of Marr’s “new doctrine on language”, resort to tendentious attempts at interpreting lexemes of unclear origin in the Germanic languages. A showcase example is one of Mosenkis’ latest publications “The Caucasian Languages in Ancient Britain and Germany” [Mosenkis 2007], in which the above scholar makes bold comparisons that are often in disagreement with the available linguistic data. In particular, we cannot agree with the proposed questionable Basque-Caucasian parallels to those English words that have a rather reliable Indo-European etymology, e.g. *path*, *thick*, *walk* [Mosenkis 2007: 9; cf. Klein 1971: 540, 761, 820]. There are no particular objections to Kartvelian-Basque parallels to English *girl*, *body* and *dog*. The approximation of the last word with Basque *u-dagara* “otter” (literally “water dog”) does not explain the nature of the Basque consonant *-r-*, which is not attested by English forms, but is really worth attention because the word *dog* does not have a satisfactory Indo-European explanation.

As far as we are concerned, the suggested morphological parallels between the Germanic and Kartvelian languages [Mosenkis 2007: 10] seem to be exceptionally interesting and require further study, as well as plausible North Caucasian

correlations of pre-Celtic substratum words [Mosenkis 2007: 10–11], which is a serious argument for taking into account the hypothesis of the common origin of the Pre-Germanic and Pre-Celtic substrata or at least one of their sublayers [Origin 2010: 47].

At the same time, we do not agree with the proposed Mosenkis' etymology of the English word *puss* "cat", which is approximated to the Basque lexeme *potzo* "big dog" [Mosenkis 2007: 9]. The impossibility to draw a comparison between English *puss* and Basque *potzo* is explained by several reasons. Firstly, the integral seme 'domestic animal', in our opinion, is not enough because the cat and the dog differ fundamentally from the point of view of their appearance, physical parameters, behaviour, domestication period and function in the home of the prehistoric man [Morgan 1926: 168–169]. Secondly, in order to assert the legitimacy of the proposed semantic development, the scholar should have provided examples of such semantic shifts as "dog" > "cat", "cat" > "mouse", or, let us say, "wolf" > "hare/rabbit" in languages of different groups. We cannot help but follow Vasilii Abaev who remarks that "[...] for the etymologist, the repetition of certain semantic shifts in different languages is the most important, if not the only, guide in the complex labyrinth of historical semasiology" [Abaev 1986: 22]. Thirdly, as convincingly shown in etymological sources, English *puss* is a well-attested onomatopoeic formation with numerous detailed correspondences, all of which occur abundantly and share the same meaning "cat" in the Indo-European languages of different groups and even in other language families: Dutch *poes* / *i puus*; Irish, Scottish and Swedish *pus*; Albanian *piso*, Romanian *pisică*, Lithuanian *puž*, Afghan *pusha*, Tamil *pusie*, *pusei* etc. [Klein 1971: 604; Skeat 1984: 422]. In other words, in the case of English *puss* we can state either independent elementary or distant (interfamilial) genetic relationship – if we assume that we are dealing with the reflexion of the ancient imitative base whose formation dates back to the time of the Nostratic proto-language unity – but not at all substratum influence.

Turning back to the Substratum Hypothesis, it seems safe to assume that today scientists have at their disposal considerable linguistic proof supported by the evidence of related disciplines which is an argument for the former spread of non-Indo-European ethnic groups in prehistoric Europe – the native speakers of Basque, Finno-Ugric, and Caucasian languages – that have survived in peripheral regions of the Continent.

First and foremost, the progress in substratum studies is associated with the latest research findings in the field of craniological anthropology and genetics which have established the affiliation of the Celts of the British Isles to the Fenno-Nordic anthropological type and also found direct evidence of the genetic proximity of the Basques, Sards, Celtic peoples of Britain and ethnic groups of the Caucasus [Cavalli-Sforza, Menozzi, Piazza 1994: 272–276]. Some Ukrainian researchers draw attention to the correlation of lexical data with the traces of substratum

phonetic influences of the Finnic idioms in the Slavic, Baltic, Germanic, and even Romance languages [Tyshchenko 2006: 78], which are known to have been formed on the basis of vulgar Latin. The Danish substratologist Guus Kroonen relates the Germanic substratum to the non-Indo-European layer of words in Greek that together represent the linguistic residue of the first European farmers [Kroonen 2012: 241]. Thus the linguistic substratum is linked with the introduction of agriculture in North-West Europe during the period of the Neolithic Revolution between the eighth and the fourth millennium BC, and is strongly associated with the Central European Linear Pottery Culture of ca. 5500–4500 BC [ibid.].

Since the 1980s, it has been increasingly argued by many that there is a significant Caucasian (especially Kartvelian) component in the Basque language, which is “[...] the only living continuation of the linguistic world of Pre-Indo-European Western Europe” [Lingvisticheskiy, red. 1990: 70]. In this way, the prominent Indo-Europeanist Abaev is inclined to accept the hypothesis of the affinity of the Basque language with the Caucasian family and considers the return to the theory of a linguistic continuum that extended from the Pyrenees to the Caucasus and served as a substratum for the Indo-European languages of the Mediterranean [Abaev 1986: 10].

A serious argument for the Pre-Indo-European substratum in the Indo-European languages is onomastics whose etymology does not find a satisfactory Indo-European explanation. Thus, the results of toponymic studies of the German onomast Hans Krahe give grounds to regard Indo-Europeans as a non-autochthonous (immigrant) population on the territory of Greece, the Apennine and the Iberian Peninsulas [Obschee 1973, red. B. A. Serebrennikov: 165].

It is worthy of note that Krahe’s “ancient European” hydronymy correlates with the “Vasconian” onomastics according to the terminology of the contemporary German substratologist Theo Vennemann [Vennemann 2003: 205]. The latter genealogically links the Indo-European languages of Western Europe with two families – “Vasconian”, whose only living representative is the Basque language, and “Atlantic”, whose descendants are the Semito-Hamitic (in modern use Afroasiatic, Afro-Asiatic or Afrasian) languages in North Africa and in the Middle East [Vennemann 2003: 203]. In this case, the “Vasconian” family is considered as the western branch of the extinct “Vasco-Caucasian” family, which in prehistoric times (ca. 4000–3500 B. C.) was disintegrated as a result of the expansion of the Indo-European tribes [Vennemann 2003: 205; cf. Gimbutas 1973: 1–21]. Here, supporting Vennemann and sharing the terminology of the above researcher, we acknowledge the existence of the “Vasco-Caucasian” proto-linguistic unity in the distant past.

The following information is an attempt by the author of this article to present further linguistic and extralinguistic evidence in favour of the Vasco-Caucasian vector in the Germanic Substratum study and to draw etymologists’ attention to

the striking Basque-Caucasian parallels to the lexemes of unknown or disputable origin of Modern English and German. No doubt, in view of the above findings, these are the languages of peoples whose history is closely connected with the prehistory of the European continent (and adjacent islands) in general and is inextricably intertwined with the Pre-Indo-European substratum donor influence in particular, and therefore, to the etymologist, is of great interest in the context of the multifaceted problem of continuity of cultures and heredity of languages.

We are deeply convinced that the lexical residue or substratum has the potential to shed light on the nature of interaction between Indo-Europeans and non-Indo-Europeans and will definitely help to penetrate the mystery of prehistoric times and to better understand the origins of modern civilization.

The **aim** of the article is partial genetic identification of hypothetically Pre-Indo-European substratum elements in the lexicon of modern English and German languages, mainly based on Common Germanic, Proto-Germanic or Proto-West-Germanic reconstructions².

The achievement of the above aim presupposes the fulfilment of the following **tasks**:

1) to identify words of unknown origin by analyzing lexicographic sources (in this publication, the coverage of results was continued on the basis of the lexical data obtained from the *Oxford dictionary of English etymology* (ODEE), edited by Charles Onions [Oxford, ed. 1966], as well as *Etymologisches Wörterbuch der deutschen Sprache* (EWDS) by Friedrich Kluge et al. [Kluge et al. 1989]);

2) to divide the identified lexemes into thematic groups;

3) to find parallels to the English and German words that do not have reliable Indo-European correspondences outside the Germanic languages, in the languages of other families – primarily in the Caucasian languages which are linked by most modern scholars with the genetically isolated Basque language – and with the help of the *Descriptive (Thematic Grouping)*, *Comparative Historical (Linguogenetic)*, *Areal (Linguogeographic)*, *Structural (Componential Analysis) Methods* and *Method of Mass Comparison* to scientifically substantiate the possibility of German-Basque and German-Caucasian approximations in the context of the Germanic (Pre-Indo-European) Substratum Theory.

Research results. An analysis of ODEE, for the purposes of discovering the lexicon of unclear etymology, allowed us to reveal almost 600 words without distinct Indo-European correspondences outside Germanic. On the other hand, according to the register of the EWDS, we have identified over 500 lexical units of unclear origin, with almost 240 of them having no reliable Indo-European correspondences outside the Germanic languages. Withdrawing from this corpus indisputable cases of sound symbolism and jargon, we arrive at 200 lexemes that

² Further abbreviated in this article as CG, PG or PWG respectively.

can be arranged into 10 main thematic groups, of which the first four turned out to be the most representative (the number of words is indicated in parentheses below):

- 1) names of plants (16): *Beere* “berry”, *Traube* “grapes”;
- 2) wildlife (21): *Aal* “eel”, *Dachs* “badger”;
- 3) the environment, objects and phenomena of nature (22): *Esch* “Field”, *Kees* “glacier”, *Regen* “rain”, *See* “sea”;
- 4) social (family) life, household utensils and products (48): *Boot* “boat”, *Ruß* “soot; ash(es)”.

For a more detailed analysis of the registers of the two above-mentioned dictionaries and the full division into thematic groups, see [Borshchevskiyi 2009: 78–82, 89–90].

As with ODEE, the given vocabulary is marked by a high degree of indigenosity (especially the first three thematic groups). One of the EWDS authors’ many merits is a well-developed procedure for the synchronisation of proto-forms: for 85 lexical units (accounting for over 40% of the 200-word list), there were reconstructed CG, PWG or pre-German forms, the fact that greatly facilitated the consideration of such a vocabulary in the context of the Pre-Indo-European Substratum Theory.

With regard to the Germanic lexicon, we can state that a number of English words of unknown origin (both with and without correspondences in other Germanic languages and reconstructed CG proto-forms) find quite reliable parallels in the basic vocabulary of the Basque language (16 correlations).

1. English *ail* “pain; trouble, worry” – Basque *(h)il* “death”.
2. English *barrow* “castrated male pig (wild boar)” (PG **barguz / -waz*) – Basque *bargo* “young boar; piglet”.

The above approximation in fact involves the entire lexeme and there seem to be no objections as to the form and semantics. On the one hand, there is an opinion that Basque *bargo* belongs to the basic native lexicon of unknown origin [Trask 2008: 127], which allegedly excludes the possibility of a late cultural borrowing and is an additional argument for prehistoric distribution of the autochthonous Basque component on the territories that confronted the expansion of Proto-Germanic tribes. On the other hand, the above PG reconstruction proves the archaic nature of the English lexeme, for which Viktor Levitskiy offers an alternative PG **barugaz* from the PIE. **bher-* “to cut”, the former having the meaning “circumcised” [Levitskiy 2010: 91]. The origin of the aforementioned Indo-European base is considered by the Ukrainian Germanist in a broader, namely Nostratic (following Vladislav Illich-Svitych), context [ibid: 96].

3. English *black* “black” – Basque *beltz (baltz)* “black” < Proto-Basque **belc³*; here also belongs *bele (bela)* “crow”. Robert Trask derives the above-mentioned

³ This proto-Basque reconstruction is given by: <http://starling.rinet.ru/cgi-bin/query.cgi?root=-config&morpho=0&basename=/data/sinocauc/basqet>

parallels from the reconstructed Basque **bel* “dark” of unknown origin [Trask 2008: 134]. The archaic and autochthonous character of the Basque lexemes is confirmed by positive onomastic evidence [ibid.].

4. English *cur* “watchdog” – Basque (*h*)*or* “dog”.

5. English *gaunt* “tall and slim” – Basque (*h*)*andi*, (*h*)*aundi* “1) big; 2) tall; 3) main, chief; 4) strange”.

6. English *gavel* “hammer (used by a judge or chairman of the meeting)” – Basque *gabi* “sledgehammer, blacksmith hammer”. English *gavel* is predominantly used in American English, which can serve as an additional argument for the early formation of the word, since the above dialect, along with numerous regional innovations (first of all, we mean late borrowings from indigenous languages of Native American tribes) still retains a significant amount of archaic features at all levels of linguistic structure that have been lost by British English over the past centuries.

7. English *geld* “castrate” – Basque *galdu* (*galtu*) “1) get lost; disappear; 2) to lose”; here also belongs *galte* “loss”; *kalte* (*kalta*) “damage; losses”.

8. English *glade* – Basque *garate* “passage in the mountains; gorge; mountain crossing”. We do not rule out the connection with Basque *goialde* “top”, which is a case of word composition on the basis of autochthonous **goi* “height” of unknown origin [Trask 2008: 207] and unclear *alde* “1) side; 2) terrain, space” [ibid.: 85]. The lost intermediate semantic link between the English and Basque lexemes could have been the meaning “open place or space”.

9. English *hame* “the bent part of a clamp (in a horse’s harness), a clutch” – Basque *kima* (*kuma*, *khuma*) “mane (of a horse)”.

10. English *haze* “haze, fog” – Basque *hese(a)* “wet, damp, humid”.

11. English *lack* 1) lack; shortage; absence; 2) fault” – Basque *akal* “empty (about a chestnut)”. In the case of the above approximation, we are obviously dealing with the metathesis of consonants [k] ↔ [l].

12. English *lass* (*lassie*) “girl or young woman” – Basque *lotsa* “1) shame; 2) dishonour; 3) modesty”. English *lass* (*lassie*) is used predominantly in Northern England and Scotland [Longman 2003: 739], i.e. on the territories that in the first millennium BC were colonized by the Celts who experienced a powerful ethnic and linguistic influence on the part of the indigenous non-Indo-European population of the British Isles, which most likely included the Basque component.

Presumably, the original meaning of the lexical unit was “shy” or “the one that should behave modestly.”

13. English *sorrow* “1) sorrow, misfortune, disaster; 2) sadness” – Basque *zori* “bad sign; omen”. There is no particular objection to the approximation of the English word to Basque *sari* “1) fee; tax; toll; 2) price, cost; 3) award; 4) punishment”. For semantic reasons, however, the above Basque *zori* seems to be a more archaic formation. It is interesting to compare English *sorrow* and Ukr. *zhurba* “sadness” with Finnish *suru* “mournful”, *surra* “fade”, *surma* “death” as proposed

by Kostiantyn Tyshchenko who considers the English-Ukrainian isogloss to be a Finnish borrowing in Germanic and Slavic [Tyshchenko 2006: 77]. It is our firm conviction that this isogloss can be expanded by including the revealed Basque parallel. Recent studies in the field of craniological anthropology and genetics suggest Basque-Finnish ethno-linguistic contacts in the distant past that were possible before the arrival of Indo-Europeans at the Danube Basin [ibid: 56–58].

14. English *weed* “(1) an unwanted wild plant; 2) algae” – Basque *idoi* “(1) puddle; 2) dirt; mud; sludge; ooze; 3) swamp”. The bilabial [w] in English *weed*, most likely, has a prosthetic nature.

15. English *yarrow* “a type of plant (*Achillea millefolium*)” – Basque (*h*)*orri* “(big) leaf”. We do not rule out the connection with Basque *orre* “juniper” or *ira* “fern”, but both formally and semantically the correlation with Basque (*h*)*orri* seems more substantiated and reasonable. Firstly, for English *yarrow* it is possible to surmise the development of the prosthetic mediolingual [j], and to state almost the same phonetic process in the proposed Basque correlate (*h*)*orri* – the emergence of the optional inspiratory [h]. Secondly, the Russian correspondence of the name of the plant *tysiachelistnik* is rather compelling if we take a closer look at its internal form, which is typologically correlated with Basque (*h*)*orri* “(large) leaf” or “a thousand leaves” (cf. Latin *millefolium*).

16. English *womb* “uterus; stomach” – Basque (*h*)*ume* (from the reconstructed **unbe* [Trask 2008: 355], Aquitanian *ombe*) “(1) the child; 2) young animal”; herein also *umoi* “(1) uterus; 2) cradle”. As in the previous comparison, in both components we assume the development of prosthetic consonants.

Etymologically valuable are striking parallels between a number of English and German words of unknown origin (7 correlations), which are found in the basic vocabulary of the Ibero-Caucasian languages of the southern (Kartvelian) and other (Abkhazo-Adyghean and Nakho-Dagestanian) groups⁴.

1) English *dill* – Georgian *dol*- “winter crop(s)”, Megrelian *ndolo* “grass, verdure, greenery; plants” (< Proto-Kartvelian⁵ **dol*). Almost perfect phonetically and compelling semantically is proto-Altaic **dilu* “juice; resin” with numerous reflexes in Turkic, Mongolian, and Tungus-Manchurian languages [SEB]. It is known that dill is native to Mediterranean countries and southeastern Europe⁶, and one of the three types of dill – fragrant dill – is used as an aromatic seasoning for dishes, because it contains, in particular, essential and fatty oils [Ukrainskyi 1986–1987, t. 2: 198]. This characteristic feature could be the basis for the original nomination of dill. We consider the above etymological approximation to be reliable, since

⁴ All the proto-language reconstructions of the Caucasian part of the comparisons are given by S. A. Starostin’s Scientific School Etymological Base [SEB]. Access mode: <http://starling.rinet.ru/cgi-bin/main.cgi?flags=wygnnnl>

⁵ Further abbreviated as PK in this article.

⁶ <https://www.britannica.com/plant/dill>

it indeed encompasses the whole lexeme as well as suggests close formal and semantic match. Obviously, the culture of cultivating dill was gradually spreading from the Pre-Indo-European Mediterranean to the regions of Western Europe, that is, onto the territories of later settlement of the proto-Germanic tribes.

2) English *heather* (Old English **hæddre*) – Megrelian *od*, Svan *ad-ra*, *had-ra* meaning “(bot.) 1) nightshade; 2) azalea” (< PK *(h)ad). Perhaps, herein belongs Dagestanian (Akhvakh) *šidira* “young grass”. The proposed comparison does not call into question the semantic aspect: the development of meaning could have resulted from narrowing the semantics according to the scheme “grass > young (juicy) grass > the name of a specific young grassy plant or bush (heather, nightshade, etc.)”. Given the variability of the PK reconstructions, the difficulties of the formal and phonetic nature seem to be resolved, especially under the assumption that the English lexeme developed the prosthetic [h] in the process of inheriting the Caucasian prototype. What is more, English *heather* belongs to the basic vocabulary of unknown origin, which excludes the possibility of a late cultural borrowing and is an additional argument for the former spread of the autochthonous Caucasian component on the territories that confronted the expansion of the Proto-Germanic tribes.

3) English *lark* – Proto-West-Caucasian **la(r)q^{wa}* with numerous reflexes in various Abkhazian-Adygean languages meaning “1) raven; 2) rook; 3) pigeon”, which indicate that the Proto-West-Caucasian prototype served as a nomination of a large or medium-sized bird.

4) German *Laum* “(dialect, dated) moisture; steam; evaporation” – PK **lumb* “wet, damp”, Proto-Altaiic **lĕmo* “raw, fresh” (< Paleo-Eurasian **IVmV*). In the process of inheriting the Caucasian prototype there must have taken place a widespread phonetic phenomenon of lag (regressive) assimilation of labial consonants (cf. German *dumm* < PG **dumba* “dumb; stupid”).

5) German *Mage* (PG **mæga-*) “relative” – Proto-North-Caucasian **zāmV* / **māzV* “relative” with reflexes in different North Caucasian languages [SEB]. The proposed comparison is semantically flawless and phonetically compelling. In the case of the North Caucasian prototype, we are obviously dealing with the metathesis of consonants [m] ↔ [ʒ]. The presence of the reconstructed PG form confirms the archaic nature of the German lexeme.

6) German *Miere* “(bot.) starwort; chickweed” – Proto-North-Caucasian **mihwVrV* “a kind of fragrant grass”; Proto-Dargin **mura* “hay” and Proto-Lezgian **mihur(aj)* “1) mint; 2) barberry” are phonetically very close to the German lexeme.

7) German *Molch* “(zool.) salamander” – Proto-Lezgian **mulaql^w* “1) worm; 2) snake”. On the other hand, a connection with the etymological cluster, centred around the Proto-Lezgian **mV[l]qlV* and Proto-Nakh **mšēlqu* “lizard”, seems very compelling and highly probable; herein is Chechen *mōlqa* having the same meaning. For more detailed typological parallels of semantic development in the names of amphibians, spiders and reptiles, see [Borshchevskiy 2013: 207].

Conclusions and prospects for further research. The above-mentioned linguistic facts that have a reliable extra-linguistic base in the form of the latest systemic data of related disciplines indicate that in Common Germanic as well as in the modern Germanic languages there are traces of ancient *non-Indo-European substrata* whose characteristic features at the lexico-semantic level are the belonging of words of unknown origin to thematic groups that are marked by a high degree of *indigenesness*. The analyzed lexical material allows us to assume the former spread of *autochthonous Basque and Ibero-Caucasian languages* on the territory of the ancient settlement of the Germanic tribes (including the British Isles). More research is needed to confirm these observations. For now though the results seem to prove the Germanic Substratum Hypothesis and provide grounds to assert the participation of the *Vasco-Caucasian component in the linguoethnogeny of the Proto-Germans*, which, however, does not rule out the *heterogeneity* of the Pre-Indo-European substrata of Proto-Germanic and needs further scientific substantiation.

References

- Abaev V. I., 1986, *Kak mozhno uluchshit etimologicheskie slovari // Etimologiya 1984*. M.: Nauka, s. 7–27.
- Borshchevskiy S., 2010, *Baskski paraleli hermanskoi leksyky nevidomoho pokhodzhennia // Mova ta istoriia : Zb. nauk. prats. K., Uman : PP Zhovtyi, № 128, s. 46–54*.
- Borshchevskiy S. V., 2009, *Slovnnyk apeliatyvnoi leksyky nevidomoho pokhodzhennia (na materialy praslovianskoi, polskoi, anhliiskoi, nimetskoi ta latynskoi mov)*. K.: Vyd-vo NPU im. M. P. Drahomanova.
- Borshchevskiy S., 2013, *Slovianska, hermanska ta romanska leksyka nevidomoho pokhodzhennia v konteksti kontaktolohichnoi peredistorii mov yevropeiskoho arealu : Monohrafiia*. K. : Vyd-vo NPU imeni M. P. Drahomanova.
- Brückner A., 1974, *Słownik etymologiczny języka polskiego*. Warszawa: Wiedza powszechna.
- Bubrih D. V., 1926, *O yazykovykh sledakh finskikh tevtonov-chudi // Yazyk i literatura*. – Leningrad, t. 1, s. 53–92.
- Cavalli-Sforza L. L., Menozzi P., Piazza A., 1994, *The History and Geography of Human Genes*. – Princeton : Princeton University Press.
- Dzidziguri Sh. V., 1987, *Baskoniya glazami N. Ya. Marra (vstupitel'naya statya) // Marr N.Ya. Basko-kavkazskie leksicheskie paralleli*. Tbilisi: Metsniereba, s. 3–8.
- Dyakonov I. M., 1984, *Sravnitel'noe yazykoznanie, istoriya i drugie smezhnyie nauki // Lingvističeskaya rekonstruktsiya i drevneyshaya istoriya Vostoka: Tezisy i doklady konferentsii*. M.: Nauka, cz. 2, s. 3–20.
- Ernout A., Meillet A., 1932, *Dictionnaire étymologique de la langue latine. Histoire des mots*. Paris: Librairie C. Klincksieck.
- Gimbutas M., 1973, *Old Europe since 7000 – 3500 B. C.: the earliest European civilization before the infiltration of the Indo-European peoples // The Journal of the Indo-European studies*, vol. 1, № 1, p. 1–21.

- Hubschmid J., 1953, *Sardische Studien. Das mediterrane Substrat des Sardischen, seine Beziehungen zum Berberischen und Baskischen sowie zum eurafrikanischen und hispano-kaukasischen Substrat der romanischen Sprachen // Romanica Helvetica*. Vol. 41.
- Kalygin V. P., Korolev A. A., 1989, *Vvedenie v keltskuyu filologiyu*. M.: Nauka.
- Klein E., 1971, *A Comprehensive Etymological Dictionary of the English Language*. Amst.; L.; N.Y.: Elsevier Publishing Company.
- Kluge F., 1989, *Etymologisches Wörterbuch der deutschen Sprache*. Unter Mithilfe von Max Bürgisser u. Bernd Gregor völlig neu bearb. von Elmar Seebold. 22. Aufl. Berlin; N. Y.: de Gruyter.
- Kretschmer P., 1896, *Einleitung in die Geschichte der griechischen Sprache*. Göttingen: Vandenhoeck und Ruprecht.
- Kroonen G., 2012, *Non-Indo-European root nouns in Germanic: evidence in support of the Agricultural Substrate Hypothesis // A Linguistic Map of Prehistoric Northern Europe. Mémoires de la Société Finno-Ougrienne 266*. Helsinki, p. 239–260.
- Levitskiy V. V., 2010, *Etimologicheskii slovar germanskikh yazykov : v 2 tomakh*. Vinnitsa : Nova Knyga, t. 1.
- Lingvisticheskii entsiklopedicheskii slovar*, 1990, gl. red. V. N. Yartseva. M.: Sovetskaya entsiklopediya.
- Longman Dictionary of English Language and Culture*, 2003. Harlow: Pearson Education Limited.
- Machek V., 1934, *Recherches dans le domaine du lexique balto-slave // Spisy filosoficke*. Brno, číslo 37.
- Machek V., 1954, *Česká a slovenská jména rostlin*. Praha: Nakladatelství Československé akademie věd.
- Marr N. Ya., 1922, *K voprosu ob yafetidizmah v germanskikh yazykah // Yafeticheskii sbornik*. – Petrograd, nr. 1, s. 43–56.
- Morgan Zh. de, 1926, *Doistoricheskoe chelovechestvo*. Moskva, Leningrad: Gosizdat.
- Mosenkis Yu. L., 2002, *Problema rekonstruktsii movy trypil'skoi kultury: avtoref. dysertatsii na zdobuttia nauk. stupenia dokt. filol. nauk: spets. 10.02.01 «Ukrainska mova», 10.02.15 «Zahalne movoznavstvo»*, Kyiv.
- Mosenkis Yu. L., 2007, *Kavkazskie yazyki v Drevney Britanii i Germanii // Proishozhdenie yazyka i kultury: drevnyaya istoriya chelovechestva*. T. 1, nr 3, s. 8–12.
- Obschee yazykoznanie : Metody lingvisticheskikh issledovaniy*, 1973, red. B. A. Serebrennikov. M.: Nauka.
- Oštir K., 1930, *Drei vor-slavisches-etruskische Vogelnamen*. Ljubljana: Znanstveno društvo.
- Pokhodzhennia anhliskoi movy v konteksti istorii brytanskoi kultury*, 2010, [Boloban M. A., Bura Kh. V., Hrabovyi P. M., Mosenkis Yu. L., Pereverziev D. I., Synyshyn R. I.] / Akademiia nauk vyshchoi osvity Ukrainy. Rada molodykh uchenykh. – Kyiv; Uman: PP Zhovtyi.
- Pokorny J., 1927, *Das nicht-indogermanische Substrat im Irischen // Zeitschrift für celtische Philologie*. Bd. XVI–XVII, S. 95–127.
- Portsig V., 1964, *Chlenenie indoevropeyskoy yazykovoy oblasti*. – M.: Progress.
- Safronov V. A., 1989, *Indoevropeskie prarodiny*. Gorkiy: Volgo-Vyatskoe knizhnoe izd-vo.
- S. A. Starostin's Scientific School Etymological Base [Electronic source]. – Access Mode : <http://starling.rinet.ru/cgi-bin/main.cgi?flags=wygnnnl>
- Skeat W., 1984, *A Concise Etymological Dictionary of the English Language*. Oxford: Oxford University Press.
- The Oxford Dictionary of English Etymology*, 1966, ed. by C. T. Onions. Oxford; N.Y.
- Tkachenko O. B., 1989, *Ocherki teorii yazykovogo substrata*. K.: Naukova dumka.

- Topolovšek J., 1894, *Die basko-slavische Spracheinheit*. Wien: Commissions Verlag von K. Gerold's Sohn, Bd. 1.
- Trask R. L., 2008, *Etymological Dictionary of Basque* / [edited for web publication by Max W. Wheeler]. University of Sussex. Access mode : <http://www.sussex.ac.uk/linguistics/1-4-1-2.html>.
- Tyshchenko K., 2006, *Movni kontakty: svidky formuvannia ukrainsiv*. K.: Akvilon-Plius.
- Ukrainskyi Radianskyi Entsyklopedychnyi Slovnyk : v 3-kh tomakh*, 1986–1987, red. kol. : A. V. Kudrytskyi (vidp. red.) ta in. 2-e vyd. K.: Holov. red. URE, t. 2.
- Vennemann Th., 2003, *Europa Vasconica – Europa Semitica (Trends in Linguistics. Studies and Monographs; 138)*; Patrizia Noel Aziz Hanna (Ed.). Berlin: Mouton de Gruyter.
- Wagner M. L., 1931, *Über die vorrömischen Bestandteile des Sardischen // Acta Romanica*. Vol. 15, S. 207–247.

VASCO-CAUCASIAN VECTOR IN THE STUDY OF GERMANIC VOCABULARY OF UNKNOWN ORIGIN

Summary

The present article reflects the results of research into the lexicon of unknown origin of the modern Germanic (English and German) languages. It is presupposed that the vocabulary specified was inherited from prehistoric idioms of ancient Europe and therefore belongs to heterogeneous Pre-Indo-European substrata. Original etymologies on the basis of revealed Germanic-Basque and Germanic-Caucasian parallels are suggested. The author postulates the participation of the Vasco-Caucasian component in ethnogeny of the Germans and differentiation of the Germanic languages.

Key words: vocabulary of unknown origin, Pre-Indo-European substratum, heterogeneity, Germanic languages, Vasco-Caucasian component